OPTIMAL ENERGY MANAGEMENT IN DECENTRALIZED SYSTEMS

Thokozile Fortunate Mazibuko

20519007

A dissertation submitted to
Durban University of Technology,
Faculty of Engineering and the Built Environmen,
in fulfillment of the requirements for the degree of

Doctor of Engineering in Electrical Engineering

Supervisor: Professor K. Moloi
Co-Supervisor: Professor K. Akindeji

Department of Electrical Power Engineering,

Durban University of Technology

March 2025

DECLARATION

I, Thokozile Fortunate Mazibuko , declare that:

(i) The research reported in this dissertation, except where otherwise indicated, is my original research.

(ii) This dissertation has not been submitted for any degree or examination at any other university.

(iii) This dissertation does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.

(iv) This dissertation does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

a) their words have been re-written but the general information attributed to them has been referenced:

b) where their exact words have been used, their writing has been placed inside quotation marks, and referenced.

(v) This dissertation does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the dissertation/thesis and in the References sections.

Candidate: Thokozile Fortunate Mazibuko

Signature: ______________________________

As the candidate's supervisors we agree to the submission of this dissertation for examination.

Supervisor: Professor K. Moloi

Signature

Co-Super

Signatur

ABSTRACT

South Africa is currently grappling with a significant energy crisis, marked by increasing electricity demand, deteriorating coal-based infrastructure, and frequent instances of load shedding. With coal accounting for more than 80% of electricity production, this dependency results in elevated carbon emissions, power outages, and economic instability. Shifting towards renewable energy offers a viable solution to these issues by harnessing South Africa's plentiful solar and wind resources, thereby promoting sustainability and cost-effectiveness. Nevertheless, critical challenges such as energy intermittency, management of excess energy, and high infrastructure expenses must be overcome to facilitate a dependable and equitable energy transition.

This research introduces a comprehensive and optimized energy-sharing framework aimed at addressing these challenges, in alignment with Sustainable Development Goal 7 (SDG7), which seeks to ensure affordable, reliable, sustainable, and modern energy for all. A hybrid renewable energy system is modeled using HOMER Pro, integrating solar, wind, and battery storage systems to achieve cost-effective energy generation. To enhance energy demand forecasting, a hybrid Long Short-Term Memory (LSTM) and eXtreme Gradient Boosting (XGBoost) model is utilized, allowing for data-driven planning and minimizing supply-demand discrepancies.

The sizing of system components is further refined through a hybrid Genetic Algorithm-Particle Swarm Optimization (GA-PSO) method, aimed at reducing costs and enhancing performance. An energy-sharing model based on Linear Programming (LP) guarantees equitable energy distribution among consumers, with a focus on prioritizing surplus renewable energy. This framework is further strengthened by incorporating Game Theory (GT) to encourage cooperative energy trading, enhance equitable energy access, and maximize cost savings.

Key performance indicators, such as reduced grid dependence, cost savings, integration of renewable energy, and decreased carbon emissions, are assessed to validate the framework's effectiveness. The results indicate that this hybrid, data-driven strategy significantly improves energy efficiency, resilience, and sustainability.

Keywords: Cost Efficiency, Energy Sharing, Energy Sizing and Allocation, Energy Transfer, Optimization

List of Abbreviations

ARIMA: Autoregressive Integrated Moving Average

COE: Cost of Energy

DG: Distributed Generation

EE:Excess Energy

GA: Genetic Algorithm

GT: Game Theory

LCOE: Levelized Cost of Energy LP:Linear programming

LSTM: Long short-term memory

PSO: Particle Swarm Optimization

SDG7: Sustainable Development Goal 7

SG: Smart Grid

XGBoost: eXtreme Gradient Boosting

LIST OF PUBLICATIONS

1. T. Mazibuko, K. Moloi and K. Akindeji, ”An Assessment to Determine Energy Surplus in an Isolated Hybrid Energy System,” 2024 IEEE PES/IAS PowerAfrica, Johannesburg, South Africa, 2024, pp. 1-6, doi: 10.1109/Power Africa61624.2024.10759522. keywords: Renewable energy sources;Electric potential; Costs; Electricity; Hybrid power systems;Energy storage;Energy excess; Isolated hybrid system; optimized energy system; Renewable energy system.

2. T. F. Mazibuko, K. Moloi and K. T. Akindeji, ”The Cost-effectiveness of Photovoltaic Systems to Ease Energy Financial Burdens on Underprivileged Households,” 2024 7th International Conference on Renewable Energy and Power Engineering (REPE), Beijing, China, 2024, pp. 123-128, doi: 10.1109/REPE62578.2024.10810125. keywords: Photovoltaic systems; Renewable energy sources; Electric potential; Costs; Software packages; Electricity; Simulation; Solar energy; Hybrid power systems;Load modeling; Energy Prices; Grid; HOMER;Hybrid Energy; Renewable Energy,

3. Mazibuko, Thokozile, Katleho Moloi, and Kayode Akindeji. 2024. ”Techno Economic Design and Optimization of Hybrid Energy Systems” Energies 17, no. 16: 4176. https://doi.org/10.3390/en17164176

ACKNOWLEDGMENTS

I extend my heartfelt gratitude to my creator, the wellspring of my strength and determination. Throughout my doctoral studies, I was fortunate to receive his careful guidance and protection. I would like to express my sincere appreciation to the Durban University of Technology for fostering a conducive academic environment that facilitated my learning, research, and adherence to sound principles. I am particularly thankful to my supervisors **Prof. Moloi**, the Deputy Dean and **Prof. Akindeji** the Head of the Department: Electrical Engineering,Durban University of Technology, for their patience and support during the course of the program. I am also indebted to the Goot Program for providing financial assistance for my studies, for which I am profoundly grateful.To my mother, who overcame numerous challenges to ensure my success, you are truly my inspiration.

I would also like to acknowledge my sons, **Iswelihle** and **Esihle**, for their understanding and unwavering faith in me. Thank you for your loving support; together, we will continue to achieve great things in life.

Place: Durban (KwaZulu Natal)

Date: 31/03/2025

Contents

Table of Contents **xi**

List of Figures **xix**

List of Tables **xxiii**

1 Introduction **1**

1.1 Research Background 1

1.2 Problem Statement 5

1.3 Research Aim and Objectives 7

1.3.1 Objectives 7

1.4 Research Motivation 8

1.5 Research Contributions 8

1.6 Research Hypothesis 11

1.7 Limitations of the Study 11

1.8 Thesis Layout 12

1.9 Conclusion 13

2 Literature Review **15**

2.1 Introduction 15

2.2 Energy Status in South Africa 17

2.3 Energy Management in Decentralized Systems 20

2.3.1 Decentralized Systems Advantages 21

2.4 Role of Energy Storage and Renewable Integration 23

2.4.1 Storing surplus energy for Future Utilization. 25

2.4.2 Grid Reliance and Peak Demand Reduction by Energy Storage 29

2.4.3 Enhancing grid stability through load balancing 33

2.5 Forecasting Techniques in Energy Systems 36

2.5.1 Importance of Load Forecasting 37

2.5.2 Machine Learning and Deep Learning Approaches 39

2.6 Optimization Algorithms for Energy Storage Sizing 53

2.6.1 Optimization Techniques for Energy Storage Sizing 55

2.6.2 Particle Swarm Optimization (PSO) Sizing Technique : 60

2.6.3 PSO Pseudocode 61

2.7 Conclusion 67

2.8 Energy Sharing and Trading in Distributed Systems 67

2.9 Game Theory Appproaches to Energy sharing: 68

2.9.1 Nash Equilibrium in Energy Sharing : 71

2.9.2 Stackelberg Game for Energy Pricing 73

2.9.3 Leader's Problem 74
2.10 Conclusion 75
2.10.1 Summary of Research Gaps 75

3 Surplus Energy Assessment in a Hybrid Energy System Network 77
3.1 Executive Summary 77
3.2 Introduction 78
3.2.1 Study Objective 79
3.3 Methodology 80
3.3.1 Aim 80
3.3.2 Meteorological Data 81
3.3.3 Solar PV Modeling 82
3.3.4 Battery Modeling 84
3.3.5 Wind Energy 85
3.3.6 Grid Modeling 87
3.4 Results 89
3.5 Conclusion 92

4 Energy Consumption Forecast by XGBoost-XGBoost Hybrid Model 95
4.1 Introduction 95
4.2 Data Analysis 97
4.3 Methodology 104

4.3.1 Data Collection and Description 105
4.4 Energy Consumption Forecasting 107
4.4.1 Data Understanding and Exploration 109
4.5 Long-Short-Term-Memory Model 112
4.6 Problems and Limitations of XGBoost Model 117
4.7 Hybrid XGBoost-XGBoost Model Overview 117
4.7.1 XGBoost Model 118
4.7.2 XGBoost Model 118
4.7.3 Evaluation of XGBoost-XGBoost Model 119
4.8 Results 119
4.8.1 Electrical Household Behavior Analysis 120
4.8.2 Comparison and performance Evaluation of Models 120
4.8.3 Hybrid XGBoost-XGBoost Model 123
4.9 Discussion 131

5 GA-PSO Hybrid Approach for Sizing and Allocation of Renewable Energy Sources 133

5.1 Introduction 133
5.2 Energy Sizing 134
5.2.1 Methodology of Sizing PV Solar and Battery System 135
5.2.2 PV System Sizing 136
5.2.3 Sizing the Battery Storage 137
5.2.4 Methodology of Sizing Wind Energy System 139

5.3 Final Turbine Specifications 143
5.4 Integration of PV Solar and Wind Energy and the Grid System 143
5.5 Energy Generation Source and Cost Data 145
5.5.1 Summary of Energy Production and Costs 148
5.5.2 Balance of Contributions 151
5.5.3 Allocation of Generation Sources 156
5.6 Methodology 160
5.7 Optimization Problem 162
5.7.1 GA-PSO Hybrid Algorithm 164
5.8 Cost Analsyis 167
5.9 Objective of the Analysis 167
5.9.1 Feature Importance Analysis 168
5.9.2 Model Accuracy 168
5.9.3 Insights from Visualizations 168
5.10 Results 169
5.10.1 Generation Sizing 169
5.10.2 Feature Importance 170
5.10.3 Energy Usage Behavior 172
5.10.4 Solar Energy Potential 172
5.10.5 Battery Cost Impact 172
5.10.6 Allocation of Generation Sources Analysis 172

5.10.7 Step 1: Consumer Allocations of Wind, Solar and Battery 173

5.10.8 Step 2: Shortfall Calculation 173

5.10.9 Step 3: Total Grid Energy Demand 173

5.10.10 Step 4: Grid Cost Calculation 174

5.10.11 Optimal Sizing of System Components 174

5.10.12 Inputs for LCOE Calculation 175

5.11 Payback Time 177

5.11.1 Inputs for Payback Time Calculation 177

5.11.2 Summary of Results from GA-PSO Energy Allocation Optimization 179

5.12 Conclusion 183

6 Optimal Energy Sharing Amongst Households 185

6.1 Introduction 185

6.2 Methodology Based on Linear Optimization (LP) Approach 187

6.3 Results Obtained from LP Model Approach 191

6.3.1 Summary Of LP Optimization Model 197

6.4 Methodology Based on Game Theory Optimization (GT) Approach 198

6.4.1 **Problem Constraints** 200

6.5 Cost Analysis Comparison Between LP and GT 202

6.6 Results Obtained from Game Theory Optimization Model 207

6.7 Conclusion 214

7 Conclusions and Recommendations **217**

7.1 Conclusion . . . 217

7.2 Recommendations and Future Work . . . 219

References **223**

Appendices **235**

A Raw Data and Preprocessing **237**

A.1 Data Source and Accessibility . . . 237

A.2 Criteria for data reduction and selection . . . 237

A.2.1 Summary Statistics of Selected Data . . . 238

B Load Prediction Models and Performance Metrics **249**

B.1 Specified Hyperparameters Used . . . 249

B.2 LSTM Model Code . . . 250

List of Figures

2.1 Classification of Metaheuristic Optimization Techniques 67

3.1 Network Configuration Map under Investigation 82

3.2 Network Configuration under Investigation 82

3.3 Monthly Electricity production from Durban South Region 90

3.4 Excess Energy Production from Renewables, Grid Sales, and Purchases (kW) Over Time (Monthly. 91

4.1 Framework for Prediction Modeling. 110

4.2 Forecasting results from different models 122

4.3 Comparison of Mean Absolute Error (MAE) for different models 123

4.4 Actual vs Predicted Energy Consumption (XGBoost-XGBoost Model) 126

4.5 Short-term and Medim-term Load Forecast 128

4.6 XGBoost-XGBoost Model Training and Validation Loss in Energy Demand Forecasting 129

4.7 Histogram of Residuals (Difference between Predicted and Actual Values) 130

5.1 Flowchart of the Algorithm Steps 165

5.2 GA PSO Optimized Hybrid Energy Sytsem 176

5.3 GA-PSO Energy Distribution 180

6.1 Energy Sharing Network Graph showing energy allocations from Battery, Surplus, to Consumers (LP). 193

6.2 LP Optimal Energy Transfer (LP) 194

6.3 Heatmap Transfer Matrix (LP) 195

6.4 Energy Balance Pie Chart (LP) 195

6.5 (a) Surplus Storage Efficiency (LP) and (b) Distribution of Surplus Energy and Battery Storage (LP) 196

6.6 Pie Chart Mean Battery Per Consumer (LP) 196

6.7 Energy Sharing Network Graph (GT). 209

6.8 Optimal Energy Transfer Per Consumer (GT). 210

6.9 Mean Battery and Surplus Distribution (GT). 211

6.10 Battery Surplus Efficieny. 212

A.1 Statistical Sampling 239

A.2 Seasonal Decomposition Consumer 1 240

A.3 Seasonal Decomposition Consumer 2 241

A.4 Seasonal Decomposition Consumer 3 242

A.5 Seasonal Decomposition Consumer 4 243

A.6 Seasonal Decomposition Consumer 6 244

A.7 Seasonal Decomposition Consumer 7 245

A.8 Seasonal Decomposition Consumer 9 246

A.9 Seasonal Decomposition Consumer 247

B.1 LSTM Model 251

List of Tables

2.1 Comparison of Advantages of PSO over GA 65

3.1 Summary of Case Study Parameters for Durban South Coast 89

3.2 Monthly Energy Data 91

4.1 Energy Consumption Statistical Analysis 100

4.2 Comparison of Model Evaluation Metrics 125

4.3 Short-term Forecasts (Next 3 Days) 127

4.4 Medium-term Forecasts (Next 7 Days) 127

4.5 Daily Forecast (kWh) per Consumer 127

5.1 Feature Coefficients and Interpretations 171

5.2 Consumer Energy Demand, Allocation, and Shortfall 174

5.3 Energy Source Contributions 181

5.4 Summary of Hybrid Energy System Performance and Economic Metrics 182

6.1 Comparison of LP and GT Energy Economics 213

Chapter 1

Introduction

1.1 Research Background

South Africa operates an energy-intensive economy that consumes more than 277.78 GWh of primary energy annually [1]. This high energy consumption is driven by the country's energy-intensive sectors, particularly the mining and metal processing industries, which rely heavily on low-cost electricity and coal as their primary energy sources

Although the growth in energy consumption has been slower than anticipated since the late 1970s, the recent economic growth and enhanced power delivery to consumers have led to a notable rise in electricity demand [2]. This increased demand has put a significant strain on the country's electrical supply system. Unfortunately, South Africa now faces a critical challenge due to an unreliable electrical supply, which severely hampers the nation's economic performance. Existing power plants, many of which are aging and overburdened, are increasingly prone to failures due to inadequate generation capacity. The ability to generate electricity has been consistently significantly diminished, and the construction of new power facilities is significantly delayed.

As a result, the country is now facing regular load shedding (planned power outages) and blackouts, highlighting the urgent need to analyze the underlying issues and explore potential solutions to address this energy crisis. To tackle environmental issues and reduce pollutant emissions into the atmosphere, various governments, including South Africa, have enacted regulations and incentives aimed at enhancing renewable energy capacity. Additionally, since 2007, South African consumers have faced escalating electricity prices and system instability due to insufficient electricity generation [3]. Furthermore, since 2007, South African consumers have experienced rising electricity costs [4], and system instability resulting from inadequate electricity generation [5]. Recent trends indicate a notable increase in these initiatives, driven by the impending depletion of various fossil fuels that traditional energy sources rely upon.

This situation, coupled with the substantial emissions of greenhouse gases resulting from the combustion of fossil fuels, has prompted institutions and authorities globally to invest in local generation from renewable energy sources [6]. This strategy involves seeking alternatives or at least diminishing reliance on fossil fuels. Moreover, a review of the literature on rural electrification demonstrates that renewable energy sources represent one of the most effective means to supply power and decrease dependence on the grid, thereby lowering electricity costs in South Africa. The generation of renewable energy sources, such as wind and solar, is inherently characterized by intermittency, a phenomenon that arises from their dependence on various meteorological factors.

These factors include solar radiation levels, which fluctuate throughout the day and across seasons, and wind patterns that can vary significantly in both intensity and direction. This variability can lead to the production of surplus energy, a situation where the amount of electricity generated exceeds the immediate demand. This surplus energy is often referred to as excess electricity or dumped energy, and it represents a critical challenge within hybrid renewable energy systems (HRESs). The presence of surplus power can have profound implications for the stability, cost, and reliability of the overall energy system. When renewable energy generation outpaces consumption, it can lead to grid instability, necessitating the curtailment of energy production from renewable sources.

This not only results in wasted potential energy, but also affects the economic viability of renewable energy projects, as producers may not be compensated for the energy generated but not utilized. Surplus power typically arises from the variable nature of renewable energy generation, particularly in scenarios where energy storage systems, such as batteries, are fully charged or when the minimum output of a generator exceeds the current energy demand. To effectively address the challenges posed by surplus energy, the implementation of energy storage systems emerges as a viable and strategic solution. These systems provide rapid response capabilities, allowing the immediate adjustment of the energy supply to match fluctuations in demand.

In addition, energy storage systems offer geographic flexibility, allowing the storage of excess energy generated at one location for use in another, thereby enhancing the overall efficiency and reliability of the power system. Among the various types of energy storage solutions, battery energy storage systems (BESS) stand out as particularly beneficial. These systems can capture excess renewable energy production during periods of high generation and low demand, and store it for later use when demand increases or generation decreases. This capability not only helps mitigate the issues associated with surplus energy, but also contributes to reducing consumer electricity expenses. By facilitating price arbitrage, where energy is stored when prices are low and released when prices are high, battery energy storage systems can optimize the economic performance of both consumers and energy producers. However, the successful implementation of these storage solutions hinges on careful planning and optimization.

Researchers have conducted extensive studies to determine the optimal sizing of battery storage systems, with the aim of aligning their capacity with the specific needs of renewable energy resources. This process involves analyzing factors such as energy consumption patterns, generation profiles, and the characteristics of the storage technology itself. Despite these efforts, many consumers, individuals or entities that both produce and consume energy, and designers often find themselves purchasing battery storage solutions that do not perfectly match the requirements of their renewable energy systems. This misalignment can lead to a range of issues, including underperformance where the anticipated benefits of energy storage are not fully realized and overinvestment where the costs associated with oversized systems outweigh the advantages. As a result, there is growing recognition of the need for innovative approaches to energy management, particularly in contexts where grid reliance is high, such as in South Africa. A promising strategy is the concept of energy sharing among households or neighbors.

This approach allows for the utilization of surplus energy generated by one household once their battery is fully charged, allowing others to draw from this stored energy when needed. This type of system not only maximizes the use of renewable energy, but also fosters community resilience and reduces overall demand on the grid.In this context, energy sharing can serve as a complementary solution to traditional energy storage, providing a more flexible and efficient means of managing energy resources. Using local networks, households can optimize their energy consumption and production, leading to a more sustainable and economically viable energy ecosystem. This collaborative approach not only improves the reliability of the energy supply, but also promotes a sense of community and shared responsibility in the transition toward a greener future. Ultimately, the objective is to merge sophisticated energy storage technologies with pioneering energy approaches. The primary goal of this initiative is to establish a hybrid energy system that incorporates a communal storage framework that effectively utilizes the surplus energy generated by renewable sources within a residential neighborhood.

This innovative approach aims to create a collaborative energy environment in which residents can maximize the benefits of renewable energy. Those who produce substantial excess energy, such as families with solar panels or other renewable energy installations, will have the opportunity to use their own storage units. Households that produce a considerable amount of surplus energy -such as those with solar panels or wind energy systems will have the opportunity to utilize their excess energy or stored energy and provide it to those in need during times of peak demand. This capability for personal energy storage allows them to manage their energy resources more effectively, ensuring that they can meet their energy requirements even during high-consumption periods within the community. In addition, residents with higher energy demands will benefit from the communal storage infrastructure, which serves as a reliable energy source when their individual production is low. By participating in this shared energy system, residents not only improve their energy security, but also contribute to a more sustainable and interconnected energy framework.

1.2 Problem Statement

South Africa is currently grappling with a significant energy crisis, characterized by soaring electricity demand, outdated infrastructure, and a substantial dependence on fossil fuels, especially coal, which constitutes over 80% of the country's electricity production. This reliance results in frequent load shedding, adversely affecting industries, communities, and overall economic development. The situation is further complicated by issues related to energy intermittency, challenges in managing surplus energy, and the high expenses associated with deploying renewable energy infrastructure.

In alignment with Sustainable Development Goal 7 (SDG7), which promotes access to affordable, reliable, sustainable, and modern energy for everyone, this research highlights the pressing need for an optimized energy-sharing model focused on renewable sources, specifically designed for South Africa's changing energy environment.

The study introduces a hybrid energy system that incorporates HOMER Pro for system design, Long Short-Term Memory (LSTM) networks for energy forecasting, Genetic Algorithm-Particle Swarm Optimization (GA-PSO) for optimal sizing, Linear Programming (LP) for cost-effective energy distribution, and Game Theory (GT) to encourage fair energy sharing among consumers. This comprehensive framework is designed to:

- **Minimize dependence on non-renewable energy sources** by focusing on the distribution of excess renewable energy.
- **Increase energy affordability** through optimal system size and equitable energy distribution.
- **Enhance energy reliability** by incorporating battery storage and utilizing grid energy as a backup option.
- **Encourage sustainable energy practices** by promoting collaborative energy sharing among users.

By adhering to the fundamental principles of SDG7, this research supports both technological advancement and social equity, paving the way for a cleaner and more resilient energy future for South Africa and beyond.

1.3 Research Aim and Objectives

The main objective of this study is to develop and assess an optimized energy-sharing framework tailored for distributed energy consumers. The framework seeks to improve cost-effectiveness, and ensure equitable energy distribution, especially within the context of South Africa's ongoing energy challenges.

1.3.1 Objectives

To achieve this aim, the study will focus on the following key objectives:

- To model a hybrid renewable energy system using simulation and optimization tools to assess sustainability, reliability, and cost performance.

- To implement advanced optimization algorithms aimed at determining the optimal sizing and distribution of components within distributed energy systems, with an emphasis on reducing costs and enhancing energy efficiency.

- To develop a resource-sharing framework utilizing mathematical optimization techniques aimed at enhancing energy distribution among consumers while adhering to multiple constraints.

- To integrate game-theoretic principles into the energy-sharing strategy in order to promote collaboration, enhance equity in energy access, and diminish reliance on the central grid.

1.4 Research Motivation

South Africa's ongoing energy crisis, characterized by frequent power outages, load shedding, and a significant dependence on aging coal plants, has profoundly affected both industries and households. The increasing need for reliable, affordable, and sustainable energy solutions is the catalyst for this research.

Renewable energy presents a viable alternative; however, challenges such as intermittency, surplus energy management, and high initial costs hinder its complete integration. In addition, current energy sharing frameworks are not optimized, resulting in inefficiencies, excessive dependence on the grid, and unequal distribution of electricity costs among consumers.

This research is driven by the objective of developing an optimized, cost-effective, and fair energy sharing framework that incorporates advanced forecasting, optimization techniques, and Game Theory to improve energy distribution and enhance consumer participation in energy trading.

1.5 Research Contributions

This research presents an innovative, data-driven framework for energy sharing that aims to tackle the energy crisis in South Africa while also promoting Sustainable Development Goal 7 (SDG7), which focuses on providing affordable, reliable, sustainable, and modern energy for everyone. The main contributions are detailed below:

1. **Development of a Hybrid Renewable Energy System (SDG7: Sustainability & Access**

- **Contribution:** This research utilizes HOMER Pro to develop a hybrid system that combines solar, wind, and battery storage, aiming to enhance the use of renewable energy sources.
- **SDG7 Alignment :** This effort aligns with target 7.2, which focuses on increasing the proportion of renewable energy within the global energy framework, by decreasing dependence on coal-based power generation and advocating for cleaner and sustainable options.

2. **LSTM-XGBoost Based Energy Forecasting for Grid Reliability (SDG7: Reliability)**

   - **Contribution:** (LSTM-XGBoost) model has been created to effectively forecast energy consumption and generation trends, thereby improving the reliability and stability of the system.
   - **SDG7 Alignment:** This initiative aligns with target 7.1, which aims to guarantee universal access to reliable energy, by facilitating proactive grid management, minimizing outages, and balancing fluctuations in supply and demand.

3. **GA-PSO Optimization for Economical Energy Sizing (SDG7: Cost Efficiency)**

   - **Contribution:** A combined approach utilizing Genetic Algorithm and Particle Swarm Optimization (GA-PSO) is employed to enhance the sizing of energy system components, aiming to reduce costs while increasing the capacity for renewable energy generation.
   - **SDG7 Alignment:** This initiative directly supports Target 7.3, which focuses on enhancing energy efficiency, by lowering both capital and operational expenses, ensuring affordability, and aiding energy consumers with low incomes.

4. **Linear Programming for Energy Distribution (SDG7: Equity)**

- **Contribution:** A Linear Programming (LP) model has been developed to facilitate the efficient and fair distribution of energy among various consumers, with a focus on maximizing the use of surplus renewable energy.
- **SDG7 Alignment:** This initiative aligns with the Sustainable Development Goal 7, specifically Target 7.a, by improving energy infrastructure and technology in developing countries, thus ensuring a fair distribution of energy resources, particularly in marginalized communities.

5. **Game Theory-Based Energy Sharing Framework (SDG7: Cooperation)**

   - **Contribution:** Utilizing a Game Theory (GT) framework encourages collaborative energy sharing among prosumers (those who both produce and consume energy), promoting joint energy trading and decreasing dependence on the grid.
   - **SDG7 Alignment:** This approach aligns with Target 7.b, which aims to extend modern energy services to everyone by establishing an energy ecosystem that enables consumers to actively engage, trade excess energy and collectively reduce expenses.

Although this research is focused on South Africa, the energy sharing framework can be scaled and tailored to address energy access issues in other developing regions. The integration of HOMER Pro, LSTM, GA-PSO, Linear Programming, and Game Theory provides a flexible model for:

- **Energy-deprived communities** that are in search of cost-effective and decentralized energy solutions.
- **Developing economies** aiming to achieve a balance between sustainability and energy reliability.

- **Future smart grids** that incorporate IoT, real-time data analytics, and blockchain-based energy trading.

This initiative not only supports South Africa's shift towards a low carbon economy but also serves as a replicable, data-driven framework that promotes global efforts towards achieving SDG7.

1.6 Research Hypothesis

1. The integration of hybrid forecasting techniques, advanced optimization approaches (including GA-PSO), and game-theoretic models for energy sharing into a cohesive energy management framework can greatly enhance energy efficiency, guarantee equitable resource distribution among consumers, lessen reliance on conventional grid systems, and encourage sustainable energy practices.

1.7 Limitations of the Study

- **Scalability:** The research is based on a limited sample of 10 consumers, and expanding this to a larger group (e.g., 300 consumers or more) may present computational and algorithmic difficulties.

- **Real-World Application:** The assumptions made in the modeling process, such as ideal energy transfer and consistent storage capacity, may not accurately represent real-world limitations, including energy transfer inefficiencies and battery wear.

- **Dynamic Market Conditions:** The analysis does not take into account fluctuating pricing or changing market conditions, which could influence the viability and profitability of energy trading strategies.

1.8 Thesis Layout

This thesis is organized into seven chapters, each of which examines a vital component of the research.

- Chapter 2: Literature Review explores existing studies on energy management, renewable energy integration, forecasting techniques, optimization methods, and energy-sharing frameworks. Critically examines the limitations of current approaches and identifies the research gap that this study aims to address.
- Chapter 3, The chapter focuses on the design of hybrid Energy Systems employing HOMER Pro, which combines renewable energy sources, specifically solar PV and wind energy, with traditional energy sources like the grid. It provides an assessment and analysis of the surplus energy derived from renewable sources and its influence on the overall cost of energy.
- Chapter 4, is based on Energy Demand Forecasting Utilizing Long Short-Term Memory (XGBoost). Its explores the application of XGBoost neural networks for predicting energy consumption trends. Details the procedures for data preprocessing, the structure of the model, the training methodology, and the accuracy assessment to guarantee reliable demand predictions.
- Chapter 5, presents a comprehensive examination of the optimization of system component sizing using GA-PSO. It details the optimization objectives, parameter tuning, and perfor-

mance evaluation to ensure cost-effective and efficient energy generation.

- Chapter 6, which focuses on the "Energy Sharing Framework Through Linear Programming and Game Theory," presents the establishment of an energy-sharing model that leverages Linear Programming (LP) alongside Game Theory. It clarifies how Game Theory is utilized to improve equitable energy distribution, support cooperative trading, and achieve cost efficiency. The chapter also analyzes the performance of the framework by considering aspects such as grid dependence, cost savings, and the benefits afforded to consumers.

- Chapter 7, Conclusion and Future Work encapsulates the principal discoveries of the research, emphasizing its contributions to energy management and sharing strategies. In addition, it addresses practical implications and limitations and offers suggestions for future research aimed at improving energy-sharing mechanisms.

This systematic framework facilitates a clear and coherent development of ideas, addressing the technical, economic, and strategic elements related to energy optimization and sharing.

1.9 Conclusion

This chapter has outlined the structure of the thesis, providing a roadmap for the study's approach to addressing South Africa's energy crisis through an optimized energy-sharing framework. The introduction presented the problem statement, the purpose of the research, the objectives, motivation and key contributions, establishing the foundation of the research. A thorough review of existing literature follows in Chapter 2, critically analyzing current energy management and sharing models, identifying their limitations, and defining the research gap that this study seeks to fill.

The methodological approach is spread over multiple chapters, each focusing on a crucial component of the energy-sharing framework. Chapter 3 develops a hybrid energy system using HOMER Pro, integrating renewable energy sources to reduce the dependency on fossil fuels. Chapter 4 focuses on XGBoost-based demand forecasting, ensuring accurate energy predictions to minimize supply-demand imbalances. Chapter 5 optimizes system component sizing using GA-PSO, ensuring cost-effective and efficient renewable energy generation. Chapter 6 introduces a Linear Programming (LP)-based energy sharing framework, enhanced with Game Theory, to ensure fair, cooperative, and cost-efficient electricity trading among consumers.

The final chapter of the thesis (Chapter 7) will present a comprehensive summary of findings, discussing the contributions and impact of the study, along with its limitations and recommendations for future work. Following this structured approach, the thesis aims to propose an innovative, data-driven and optimization-based energy sharing solution that improves grid stability, cost efficiency, and equitable energy distribution in the evolving energy landscape of South Africa.

Chapter 2

Literature Review

2.1 Introduction

In a hybrid renewable energy system, surplus power refers to the portion of energy generated by power-producing components that remains unused. This situation can arise in any energy system designed to meet a particular demand. When the battery reaches full capacity or when the generator's minimum output exceeds the load requirements, renewable energy sources may produce excess electricity that cannot be utilized by either the load or the batteries. Consequently, this surplus power must be discarded or curtailed. The amount of surplus power generated is a critical factor influencing the voltage and frequency stability of the hybrid energy system, and this quantity needs to be minimized to ensure the system operates reliably and delivers electricity to consumers with a high level of dependability [7]. This parameter has a significant impact on the economic practicality of the hybrid energy system. Energy cost is defined as the ratio of the total cost of the energy system to the amount of usable power produced. A lower level of surplus electricity in each specific hybrid configuration results in a greater amount of usable electricity, thereby improving energy efficiency and reducing overall energy costs.

In off-grid hybrid renewable energy systems (HRESs), any power that is not utilized is deemed wasted. During the process of power wastage, excess electricity is sent to a nonproductive dump load, typically a resistor bank. The employment of such a nonproductive dump load effectively squanders surplus production from the HRES, which can be likened to the practice of curtailment [8]. The integration of the Hybrid Renewable Energy System (HRES) with the utility grid allows surplus power generated after charging storage units to be fed back into the grid, resulting in minimal excess electricity production [9]. However, this practice has not yet been adopted in South Africa. In such systems, procuring electricity from the grid can facilitate peak-shaving, thereby reducing the amount of surplus electricity generated by the HRES [10].

If selling surplus electricity to the grid is not feasible for any reason, it becomes necessary to implement a dummy load to manage the unused electricity. Currently, the design of HRESs that can effectively mitigate excess electricity generation is a critical area of research. In [11] it is indicated that the allowed limit of excess electricity in standalone hybrid renewable energy systems (HRESs) should be maintained below 5%. According to the existing literature, an energy surplus that exceeds 10% signifies a suboptimal level of energy efficiency within hybrid energy systems. This mitigation can be achieved by creating a more adaptable energy system or employing strategies for the efficient utilization of surplus power, such as the energy export method in large-scale renewable systems or advanced energy storage solutions [12].

As a result, consumers often opt for energy storage systems that closely correspond to the size of their renewable energy sources, which can lead to an excess or a deficiency in storage capacity [13]. This misalignment can diminish the potential benefits of energy storage systems. This study proposes a design methodology that integrates the challenges of surplus energy management with those of storage system optimization, with the aim of improving the effective use of surplus energy generated from renewable sources while simultaneously improving the efficiency of storage systems. The methodology is designed to be versatile and applicable in various contexts of installation, electricity pricing structures, and funding models. Furthermore, the study shares selected findings on optimally sized residential renewable storage systems.

2.2 Energy Status in South Africa

In South Africa, coal-fired power plants are a dominant source of electricity, generating around 78% of the total energy supply of the country. The majority of these power facilities were constructed during the apartheid era, specifically in the 1960s and 1970s, a time when the nation was heavily reliant on fossil fuels for its energy needs. This historical context has had lasting implications for the country's energy landscape. As a direct consequence of this dependence on coal, South Africa has become the seventh highest country in the world for per capita carbon dioxide emissions. This is particularly concerning given that the nation contributes approximately 40% of the total carbon dioxide emissions produced on the African continent, highlighting its significant role in the global climate crisis [14].

Load shedding has emerged as a significant issue in South Africa, primarily attributed to the deterioration of infrastructure and corruption in the procurement processes for coal and power plants. This situation has contributed to Eskom's substantial debt, which stands at USD 24 billion. The challenge of meeting the increasing demand for electricity is exacerbated by the high costs and greenhouse gas emissions linked to conventional energy sources. Enhancing the adoption of renewable energy systems could play a vital role in mitigating carbon emissions resulting from these greenhouse gases. In addition, there are numerous challenges associated with electricity generation, transmission of power from generation sites to consumers, and distribution of electricity in rural regions, particularly in light of rising costs. The manner in which rural communities adapt during energy crises is a critical concern that deserves attention. Energy poverty is prevalent in these areas, defined by an insufficient supply of energy to fulfill basic human energy needs. In terms of affordability, energy poverty signifies a condition where individuals are unable to access electricity from either traditional or modern sources. In particular, the proportion of South African households connected to the electrical grid has increased from 76.7% in 2002 to 89.6% in 2022 [15].

The observed increase in the use of alternative electricity sources, such as generators, corresponded with a rise in the percentage of families utilizing these options. The reliance on this type of energy for cooking escalated from 1.2% in 2014 to 7.8% in 2019, subsequently declining to 4.8% by 2022. Similarly, the proportion of households that use gas, predominantly liquefied petroleum gas, increased from 2.2% in 2002 to 6.7% in 2022. Since 2002, there has been a notable reduction in the consumption of paraffin, coal, and firewood. The percentage of homes that use paraffin decreased from 16.1% in 2002 to 2.8% in 2022, while the use of firewood decreased from 20% to 7.7%. The highest percentage of families relying on electricity as their primary cooking energy is found in the Free State (86.2%) and KwaZulu-Natal (82.0%), whereas Limpopo records the lowest at 63.5%. Alternative electricity sources, including generators, are most prevalent in Gauteng (9.0%) and Northwest (6.8%) [16].

In Gauteng, paraffin remains the most commonly used fuel at 5.3%, with the Western Cape exhibiting the least usage at 0.2%. The reliance on wood and coal is most significant in Limpopo (34.0%), Mpumalanga (18.0%), the Eastern Cape (9.3%), Northwest (8.4%), and KwaZulu-Natal (8.2%). In contrast, less than 1% of families in the Western Cape and Gauteng utilize wood for cooking, at 0.5% and 0.7%, respectively. Gas is predominantly used by households in the Western Cape (19.9%) and Northern Cape (12.0%) [17]. South Africa's electricity infrastructure is primarily composed of major power plants situated inland, close to Gauteng's mining and industrial regions. Eskom has undertaken various initiatives aimed at integrating Independent Power Producers (IPPs) into the national energy landscape [18]. These efforts were designed to diversify the energy supply, enhance competition, and ultimately improve the reliability and sustainability of electricity generation in South Africa. However, despite these intentions, the integration of IPPs has faced significant challenges within the existing single-buyer framework. This framework, which centralizes the purchasing of electricity through Eskom, has often limited the ability of IPPs to operate effectively and compete on equal footing with Eskom's own generation capacity.

The constraints of the single-buyer model have led to delays in project approvals, regulatory hurdles, and a lack of clarity regarding pricing and procurement processes. As a result, many potential IPPs have struggled to enter the market, stifling innovation and the growth of renewable energy sources. This situation has highlighted the need for a more flexible and inclusive approach to energy procurement that can accommodate the diverse range of energy producers and technologies available today.

Looking ahead, it is imperative that South Africa's development of energy policy prioritize the inclusion of renewable energy sources as a fundamental component of the national energy mix. This shift is essential not only for meeting the country's climate commitments but also for ensuring energy security and reducing reliance on fossil fuels. Policymakers must create a regulatory environment that encourages investment in renewable technologies, facilitates the entry of new players into the market, and supports the development of a decentralized energy system. Incorporating renewable energy into the energy policy framework will require a comprehensive strategy that addresses issues such as grid integration, energy storage solutions, and the establishment of fair pricing mechanisms. In addition, fostering collaboration between Eskom, IPPs, and other stakeholders will be crucial in creating a cohesive and sustainable energy ecosystem. By embracing these changes, South Africa can move towards a more resilient and environmentally friendly energy future, ultimately benefiting both the economy and the well-being of its citizens. The subsequent section focuses on the shift from centralization to decentralization, exploring the various factors that have contributed to this transformation in organizational and governance structures. Centralization, characterized by a concentration of decision-making authority and control within a single central entity, has traditionally been the dominant model in many sectors, including government, business, and technology. However, recent trends indicate a significant movement towards decentralization, where power and decision making are distributed across multiple levels or entities.

2.3 Energy Management in Decentralized Systems

Historically, the electricity sector has relied on centralized power generation, which predominantly uses fossil fuels like coal, oil, and gas to supply electricity to consumers via a vast transmission network. In 2022, fossil fuels represented 60% of global electricity generation. The combustion of these fuels emits greenhouse gases, including carbon dioxide and methane, which contribute to global warming and climate change. To mitigate carbon emissions, various strategies have been proposed to enhance carbon efficiency while meeting the heating needs of residential buildings [19]. While numerous researchers have concentrated on reducing demand, an alternative approach involves gradually transitioning away from high-carbon fossil fuels or implementing grid-connected hybrid systems. Countries such as South Africa remain heavily reliant on the grid, making it essential to shift from natural gas and oil to low-carbon alternatives like green electricity.

This transition can reduce electricity purchases from the grid, decrease grid dependency, and promote environmental health. Notably, the adoption of decentralized systems can lead to improved conversion efficiency and facilitate the integration of locally generated renewable energy, such as the installation of photovoltaic panels on rooftops.

2.3.1 Decentralized Systems Advantages

- **Self-Generation:** Individuals can harness renewable energy sources such as solar panels and wind turbines to produce their own electricity. This self-generation reduces reliance on traditional energy grids and fossil fuels, promoting sustainability.

- **Energy Storage:** To achieve true energy autonomy, people must also have the means to store the energy they produce. This can be accomplished through battery systems, thermal storage, or other innovative technologies that allow energy retention and use during periods of low generation or high demand.

- **Energy Efficiency:** Alongside production and storage, energy autonomy involves optimizing energy consumption. Individuals can implement energy-efficient appliances, smart home technologies, and sustainable practices to minimize their energy needs, thereby improving their overall autonomy.

- **Decentralization:** Energy autonomy promotes a decentralized energy model, where power generation is distributed rather than concentrated in large, centralized facilities. This shift can lead to increased resilience to power outages and fluctuations in energy prices.

- **Economic Benefits:** By producing and managing their own energy, people can potentially reduce their energy bills, increase their financial independence, and even generate income by selling excess energy back to the grid or participating in community energy programs.

- **Environmental Impact:** Energy autonomy contributes to a reduction in carbon emissions and environmental degradation by encouraging the use of clean and renewable energy sources. This shift supports global efforts to combat climate change and promotes a more sustainable future.

- **Community Empowerment:** Energy autonomy can foster a sense of community and collaboration, as individuals and neighborhoods come together to share resources, knowledge, and technologies. This collective approach can improve social ties and promote local economies.

- **Technological Innovation:** The pursuit of energy autonomy drives innovation in energy technologies, including advancements in renewable energy systems, energy storage solutions, and smart grid technologies. These innovations can further enhance the efficiency and accessibility of energy production and management.

- **Policy and Regulation:** Achieving energy autonomy may require supportive policies and regulations that encourage the adoption of renewable energy, provide incentives for self-generation, and facilitate the integration of decentralized energy systems into existing infrastructure.

The variability in power generation resulting from fluctuations in renewable energy sources poses a significant challenge to the capacity of the electrical power grid. This variability can lead to periods of excess energy production when conditions are optimal, such as during sunny or windy days, followed by times of reduced generation when the weather is less favorable. Such inconsistencies can strain the grid, making it difficult to maintain a balance between supply and demand. Consequently, alongside efforts to optimize energy production from renewable sources, such as improving the efficiency of solar panels and wind turbines and enhancing grid infrastructure, it is essential to implement battery energy systems. These systems can store surplus energy generated during peak production times and release it during periods of low generation, ensuring that the stability of the grid is not compromised. By integrating battery energy storage solutions, we can create a more resilient and reliable power grid that can effectively accommodate the inherent fluctuations of renewable energy sources. Therefore, it is essential to appropriately size energy storage systems to ensure the effective management of surplus energy generated from renewable sources.

2.4 Role of Energy Storage and Renewable Integration

The recent surge in efforts to reduce greenhouse gas emissions and improve the security of electric energy has been nothing short of remarkable. This movement reflects a growing global awareness of the urgent need to combat climate change and transition to more sustainable energy systems. Over the past decade, we have witnessed substantial growth in the integration of intermittent renewable energy sources (RESs), such as photovoltaic (PV) systems and wind power, into the existing electrical grid. This shift is driven by technological advancements, policy incentives, and collective commitment to a cleaner energy future.

However, the integration of these renewable energy sources is not without challenges. The reliable and stable operation of the electrical grid is increasingly tested by the unique characteristics of RESs, which introduce a variety of operational and control issues. One of the primary challenges is generation uncertainty. Unlike traditional fossil fuel power plants that can produce a consistent output, the power generated from RESs is inherently variable and dependent on unpredictable weather conditions. For example, fluctuations in wind speed can lead to significant changes in wind power output, while cloud cover can affect the efficiency of solar panels. This variability can create difficulties in maintaining a balanced supply and demand for electricity [20].

In addition to the uncertainty of the generation, the integration of RESs raises concerns about voltage and angular stability. The electrical grid relies on a delicate balance of voltage levels and phase angles to function effectively. The intermittent nature of RESs can disrupt this balance, leading to potential instability in the grid. Power quality is another critical issue, as fluctuations in renewable generation can result in voltage sags, swells, and frequency deviations, which can adversely affect sensitive electrical equipment and overall grid performance. Reactive power support is also a significant concern. Traditional power plants can provide reactive power, which is essential for maintaining voltage levels and ensuring the stability of the grid. However, many RESs, particularly those that are not equipped with advanced control systems, may struggle to provide adequate reactive power support, further complicating grid management.

In addition, fault ride-through capability is a vital aspect of grid reliability. In the event of a fault, such as a short circuit, traditional power plants can remain online and support the grid. However, many RESs may disconnect from the grid during such events, which can exacerbate the situation and lead to larger-scale outages. To address these challenges, energy storage systems (ESSs) have emerged as a crucial component in the modern energy landscape. ESSs play a pivotal role in managing the fluctuations associated with RESs by capturing excess generated power during periods of high production and making it available when demand exceeds supply. Below are some key requirements for ESS within DG systems, which include the capacity to absorb excess energy during periods of high generation, the ability to discharge stored energy during peak demand, Reducing reliance on the grid during peak demand and to enhance grid stability through load balance.

2.4.1 Storing surplus energy for Future Utilization.

In a hybrid renewable energy system, surplus power is a crucial concept that refers to the portion of electricity generated by various power generation components that remains unutilized. This phenomenon can occur in any energy system that is designed to meet a specific demand, particularly when the generation capacity exceeds the immediate consumption needs. Surplus power typically arises under certain conditions. For example, when energy storage systems, such as batteries, reach their full capacity, they cannot accept additional electricity. Similarly, if the output of a generator, such as a wind turbine or solar panel, exceeds the load requirements, the excess energy produced cannot be consumed by the electrical load or stored in batteries. This situation can lead to a scenario where renewable energy sources, which are often variable and intermittent, generate more electricity than is needed at a given time.

When surplus electricity is generated, it presents a challenge for the management of the hybrid energy system. The excess power must be either discarded or curtailed to prevent potential damage to the system and to maintain operational stability. Discarding surplus power can involve shutting down generation units or diverting the excess energy to resistive loads, which can be inefficient and wasteful. On the other hand, curtailment strategies may include adjusting the output of renewable sources or implementing demand response measures to align generation with consumption. The quantity of excess power generated is a critical factor that significantly influences the voltage and frequency stability of the hybrid energy system. Voltage stability refers to the ability of the system to maintain a constant voltage level, while frequency stability pertains to the system's ability to keep the frequency of the electrical supply within a specified range. Both of these parameters are essential for the reliable operation of the electrical grid and for ensuring that consumers receive a consistent and stable electricity supply.

To achieve optimal performance, it is essential for the amount of surplus power generated to be minimized, ideally approaching zero [21]. This can be achieved through various strategies, such as improving renewable generation forecasting techniques, improving energy storage capabilities, and implementing advanced grid management technologies. By effectively managing surplus power, hybrid renewable energy systems can improve their reliability, reduce waste, and provide a more stable and reliable electricity supply to consumers, ultimately contributing to a more sustainable energy future. Today, the design of Hybrid Renewable Energy Systems (HRESs) that can prevent overproduction of electricity is a critical area of research. As global demand for clean energy continues to increase, the integration of various renewable sources, such as solar, wind, hydro, and biomass, into cohesive systems has become increasingly important. The challenge lies in managing the intermittent nature of these energy sources, which can lead to periods of excess electricity generation that, if not properly managed, can result in wasted resources and potential damage to the grid.

To address this issue, researchers and engineers are focusing on developing more adaptable energy systems that can respond dynamically to fluctuations in energy production and consumption. This adaptability can be achieved through advanced forecasting techniques, real-time monitoring, and smart grid technologies that allow better communication between energy producers and consumers. Using data analytics and machine learning, these systems can predict energy demand more accurately and adjust production accordingly, thus minimizing the risk of overproduction. In addition to creating more adaptable systems, it is essential to implement strategies for the efficient utilization of surplus energy. An effective approach is the establishment of energy export mechanisms, particularly in larger renewable installations. By connecting these systems to regional or national grids, excess electricity can be sold or transferred to areas experiencing higher demand, thus ensuring that generated energy is used instead of wasted [22]. This not only helps stabilize the grid, but also provides an additional revenue stream for energy producers.

For smaller HRESs, the incorporation of water storage solutions presents another viable strategy for managing surplus energy. By using excess electricity to pump water into elevated storage tanks or to convert it into hydrogen through electrolysis, these systems can store energy in a form that can be easily converted back into electricity when needed. This not only helps balance supply and demand, but also improves the overall resilience of the energy system, allowing greater integration of renewable sources [23].

In addition, the development of innovative technologies, such as battery storage systems, thermal energy storage, and demand response programs, can complement these strategies. By investing in research and development, policymakers and industry stakeholders can create a more sustainable energy landscape that maximizes the benefits of renewable resources while minimizing the challenges associated with overproduction. Surplus electricity is not a pressing concern in certain regions of the world, particularly in Europe, where a significant number of areas are interconnected through a synchronous electrical grid. This interconnectedness allows for a more stable and efficient distribution of electricity across vast distances. However, despite this advantage, there is still considerable interest in promoting on-site electricity consumption. This approach aims to reduce the energy losses that typically occur with long transmission lines and to address the imbalances that can arise between power generation and demand, especially during peak usage periods. Even in countries with robust grid systems, these issues remain relevant.

In grid-connected systems, the presence of a substantial amount of excess electricity can create challenges, particularly for low voltage (LV) and medium voltage (MV) grids that host large renewable energy installations. The variability of renewable energy sources, such as solar and wind, can lead to periods of overproduction, which, if not managed properly, can strain the grid infrastructure. To effectively handle these situations, several researchers have proposed several viable solutions [24]:

Short-term and Long-term Energy Storage :

One of the most effective strategies involves the use of energy storage systems. Short-term storage solutions, such as battery banks, can quickly absorb excess electricity and release it back into the grid when demand increases. These battery systems are known for their high roundtrip efficiency, making them an attractive option for managing surplus energy. On the other hand, long-term energy storage solutions, such as hydrogen tanks, offer a different set of advantages. Hydrogen can be produced through electrolysis during periods of excess electricity and stored for later use, either as a fuel for fuel cells or reintegrated into the grid. While hydrogen tanks provide greater energy density, they typically have lower roundtrip efficiency compared to battery systems.

Fast-Charging Stations for Electric Vehicles :

Another promising solution is the establishment of fast charging stations for battery electric vehicles (BEVs) and hydrogen-powered vehicles. These stations can serve as a buffer for excess electricity, allowing for rapid charging of vehicles during periods of surplus. This not only helps stabilize the grid, but also promotes the adoption of electric vehicles, which can further reduce the reliance on fossil fuels.

Exporting Surplus Stored Energy :

Finally, exporting surplus stored energy presents an opportunity to optimize resource utilization. This can be achieved by selling excess electricity to neighboring grids or converting it into hydrogen for local markets. By exporting energy, regions can capitalize on their renewable resources while also contributing to a more interconnected and resilient energy landscape.

2.4.2 Grid Reliance and Peak Demand Reduction by Energy Storage

The reliance on energy resources derived from fossil fuels, such as coal, oil, and natural gas, has long been a cornerstone of global energy production. However, this dependence has significant environmental repercussions, particularly in terms of global warming and the emission of greenhouse gases (GHGs). The combustion of fossil fuels releases large quantities of carbon dioxide (CO2) and other harmful pollutants into the atmosphere, which trap heat and contribute to the greenhouse effect. This phenomenon has led to rising global temperatures, extreme weather events, and a host of ecological disruptions that threaten both natural ecosystems and human societies.

In response to the urgent need to mitigate climate change and reduce our carbon footprint, there has been a growing focus on the incorporation of renewable energy sources into the energy mix. Among these, renewable energy systems such as solar photovoltaic (PV) and wind energy systems have emerged as a particularly promising solution, which generates energy without emitting greenhouse gases, making it a clean and sustainable alternative to fossil fuels.The integration of solar PV systems into the distribution network is seen as a vital component of the transition towards clean energy. By decentralizing energy production and allowing for localized generation, solar PV can enhance energy security and resilience. Additionally, the widespread adoption of solar technology can lead to significant reductions in energy costs for consumers and businesses alike, as the price of solar panels continues to decline.

Moreover, the shift towards renewable energy sources like solar PV and wind turbines aligns with global efforts to meet international climate targets [25]. By investing in renewable technologies, countries can not only reduce their reliance on fossil fuels but also stimulate economic growth through the creation of green jobs and the development of new industries.This shift not only helps to mitigate global warming and reduce greenhouse gas emissions, but also paves the way for a more sustainable and resilient energy future. The intermittent nature of renewable energy sources, such as solar and wind, presents significant challenges for utilities in countries like Sri Lanka and Brazil. In these regions, peak electricity demand typically occurs at night, following sunset, when solar energy generation ceases. This mismatch between energy supply and demand creates a pressing need for utilities to operate additional power plants, often relying on fossil fuels, to meet the heightened demand during these hours [26–29].

As a result, the reliance on these supplementary power sources not only increases operational costs but also contributes to rising electricity prices. This economic strain can have far-reaching consequences, adversely affecting critical services such as healthcare and education. For instance, hospitals may struggle to maintain essential operations during peak hours due to high energy costs, while schools may face budget constraints that limit their ability to provide quality education. The challenge of balancing supply and demand in the context of renewable energy integration thus poses a significant barrier to sustainable development in these countries.BESS is capable of delivering a range of auxiliary services to the power grid, including load shifting, frequency regulation, and stabilization of the power supply. These functions play a crucial role in mitigating the effects of fluctuations in energy supply and demand. By enhancing the stability of the power system, these capabilities alleviate some of the difficulties encountered by utilities in ensuring consistent power delivery. Nevertheless, it is essential to implement an effective energy management strategy to ensure the harmonious operation of both the storage system and the solar photovoltaic system[30], [31].

The research carried out in [32] introduces an innovative Energy Management System (EMS) algorithm specifically designed for a standalone photovoltaic (PV) system. This algorithm aims to optimize the efficiency of power utilization in various operational modes, ensuring that the energy generated by the photovoltaic system is used in the most effective manner possible. By intelligently managing the distribution and storage of energy, the EMS seeks to maximize the benefits of solar power, reduce waste, and improve overall system performance. However, it is important to note that the EMS has limitations, particularly in situations where photovoltaic power generation is not available. For instance, during periods of low sunlight, such as cloudy days or at night, the system relies on the Battery Energy Storage System (BESS) to provide the necessary power to meet the load demands. If the BESS is completely depleted, the EMS faces a significant challenge: it is unable to meet the load requirements of the system. This scenario highlights a critical vulnerability in the energy management strategy, as the absence of both solar power and stored energy can lead to power shortages and interruptions in service.

To address this issue, more research and development may be necessary to explore alternative energy sources or backup systems that can be integrated into the EMS. In addition, strategies could be investigated to improve the efficiency of energy storage and enhance the resilience of the system during periods of low energy availability. By tackling these challenges, the EMS could be refined to ensure a more reliable and consistent power supply, even in the absence of photovoltaic energy. In the research conducted in [33], the daily time frame is segmented into three distinct modes to manage the charging and discharging of the Battery Energy Storage System (BESS) and the procurement of energy from the grid. The Energy Management System (EMS) plays a crucial role in enhancing the efficiency and effectiveness of the Battery Energy Storage System (BESS) by strategically managing the charging process. During times when electricity demand is low, the EMS identifies opportunities to source electricity from the grid at reduced rates or during off-peak hours. This proactive approach not only helps minimize energy costs but also ensures that the BESS is charged optimally.

Using the excess capacity of the grid during these low-demand periods, EMS maximizes the utilization of the photovoltaic (PV) and BESS configurations. The integration of solar energy generation with battery storage allows for a more sustainable energy solution, as EMS can store excess solar energy generated during the day for use during peak demand times or when solar generation is insufficient [34]. However, these research methodologies emphasize the importance of optimizing energy management systems, particularly through the strategic use of power sourced from the electrical grid to charge Battery Energy Storage Systems (BESS).

This approach is designed to improve the efficiency and reliability of energy usage, allowing a better integration of renewable energy sources and improved overall system performance. Although it is crucial to consider the financial implications of this practice. Charging the BESS from the grid can lead to increased demand charges on the user's electricity bill. Demand charges are fees imposed by utility providers based on the maximum amount of power drawn from the grid during peak usage periods. These charges can significantly impact the overall cost of electricity for users, especially if the pricing structure established by the utility provider is not favorable. The methodologies suggest that, while the use of grid power to charge BESS can provide operational benefits, users must carefully analyze their energy consumption patterns and associated costs. By understanding the pricing structure and demand charges, users can make informed decisions about when to charge their BESS, potentially shifting their energy usage to off-peak times when electricity rates are lower. This strategic approach can help mitigate the financial impact of demand charges while still reaping the benefits of better energy management.

In summary, while the methodologies presented in the referenced studies offer valuable insight into optimizing energy management through BESS, they also highlight the need for users to be aware of the potential cost implications associated with demand charges. A comprehensive understanding of both the technical and financial aspects of energy management is essential to maximize the benefits of BESS while minimizing costs.

In conclusion, the design of Hybrid Renewable Energy Systems that effectively prevent the overproduction of electricity is a multifaceted challenge that requires a combination of adaptable technologies and strategic energy management practices. By focusing on these areas, we can pave the way for a more efficient, reliable, and sustainable energy future.

2.4.3 Enhancing grid stability through load balancing

Load balancing is a critical process that ensures the equitable distribution of electrical demand among various energy sources, thereby optimizing resource efficiency and preventing the overloading of any single source. This practice is particularly important in the context of renewable energy systems, where the generation of power can be highly variable and subject to rapid changes [35]. In renewable energy systems, effective load balancing becomes essential for managing the inherent fluctuations in power generation that arise from factors such as weather conditions, time of day, and seasonal variations. For instance, solar energy production is heavily dependent on sunlight availability, which can vary significantly due to cloud cover, atmospheric conditions, and the angle of the sun throughout the day.

Similarly, wind energy generation is influenced by wind speed and direction, which can change unpredictably.To address these challenges, load balancing strategies must be implemented to ensure that energy supply meets demand consistently. This can involve a combination of techniques, such as integrating energy storage solutions, utilizing demand response programs, and employing advanced forecasting methods to predict power generation and consumption patterns. Energy storage systems, such as batteries, can store excess energy generated during peak production times and release it during periods of low generation, thus smoothing out the supply. Moreover, demand response programs can incentivize consumers to adjust their energy usage during peak demand periods, further alleviating pressure on the grid. By encouraging users to shift their consumption to times when renewable energy generation is high, these programs help to balance the load and enhance the overall efficiency of the energy system.

In addition to these strategies, the integration of smart grid technologies plays a vital role in load balancing. Smart grids utilize advanced communication and control systems to monitor and manage energy flows in real-time, allowing for more responsive and adaptive load balancing. This technology enables grid operators to respond quickly to fluctuations in both supply and demand, ensuring a stable and reliable energy supply.

Overall, effective load balancing in renewable energy systems is not only crucial for maintaining grid stability but also for maximizing the utilization of renewable resources. By addressing the challenges posed by variable power generation, load balancing contributes to a more resilient and sustainable energy future, ultimately supporting the transition towards cleaner energy sources and reducing reliance on fossil fuels. In the lack of effective load balancing and demand response measures, the power grid faces heightened susceptibility to instability, which may lead to power outages and various disruptive incidents that can have extensive economic and social consequences [36].The authors sought to tackle these challenges by developing an extensive simulation model using MATLAB/Simulink, a powerful tool widely used for modeling, simulating, and analyzing dynamic systems. This model integrates various modules that are meticulously designed to replicate the behavior and characteristics of renewable energy sources, such as solar panels, wind turbines, and hydroelectric systems, as well as energy storage systems like batteries and pumped hydro storage.

Furthermore, the model encompasses grid infrastructure components, including transmission lines, substations, and distribution networks, allowing for a comprehensive representation of the entire energy ecosystem. To enhance the model's functionality, the authors incorporated advanced algorithms specifically designed for load balancing and demand response. These algorithms are crucial for managing the fluctuations in energy supply and demand that often accompany the integration of renewable energy sources. By simulating different scenarios, the model enables a detailed analysis of how these strategies can influence grid stability, particularly during peak demand periods or when renewable generation is low. their objective was to offer significant insights into the effectiveness of various strategies for managing the complexities that arise from the incorporation of renewable energy into the existing grid infrastructure. By examining the interactions between renewable energy generation, energy storage, and demand-side management, the authors aimed to identify best practices and innovative solutions that could enhance the reliability and resilience of the power grid.

Ultimately, the findings from this study are expected to contribute to the development of more sustainable energy systems, facilitating a smoother transition towards a low-carbon future while ensuring that energy supply remains consistent and reliable for consumers. The vast array of research concerning load balancing and demand response within renewable energy systems highlights the ever-evolving characteristics of this domain. As the global energy landscape shifts towards sustainability, the integration of renewable energy sources such as solar, wind, and hydroelectric power has become increasingly prevalent. However, these sources are inherently variable and unpredictable, leading to challenges in maintaining a stable and reliable energy supply. Traditional centralized load balancing techniques, which rely on a consistent and predictable energy output, encounter significant difficulties in this context. The sporadic and decentralized attributes of renewable energy sources necessitate a rethinking of conventional strategies.

Recent studies have explored novel approaches that integrate advanced technologies such as machine learning, which can analyze vast datasets to predict energy generation and consumption patterns, optimization methods that can efficiently allocate resources in real-time, and game theory, which can model the interactions between various stakeholders in the energy market. These innovative strategies aim to effectively navigate the complexities of this changing environment, ensuring that energy supply and demand are balanced even in the face of fluctuating renewable energy outputs [37, 38]. Although significant advances have been made, considerable research deficiencies remain. Several studies concentrate exclusively on load balancing or demand response, failing to incorporate both elements, which are essential for the effective functioning of renewable energy grids.

2.5 Forecasting Techniques in Energy Systems

Energy forecasting plays a crucial role in the effective operation of smart grid (SG) systems, as it supports a variety of applications that are vital to maintain grid stability and efficiency. These applications include demand-side management, which involves adjusting consumer demand for energy through various strategies to ensure that supply meets demand; load shedding, which is the intentional reduction of load on the grid to prevent overload and maintain system reliability; and optimal dispatch, which refers to the process of determining the most efficient way to allocate generation resources to meet demand while minimizing costs and emissions [39]. One of the primary challenges currently facing the grid is the need to achieve accurate forecasting while minimizing prediction errors. This challenge is exacerbated by the inherent uncertainties associated with energy consumption patterns, which can be influenced by a multitude of factors, such as weather conditions, economic activity, and consumer behavior. Furthermore, the detailed nature of smart grid data, which includes high-resolution time series data from various sources such as smart meters, renewable energy sources, and grid sensors, adds complexity to the forecasting process.

To address these challenges, advanced forecasting techniques and models are being developed that take advantage of machine learning, artificial intelligence, and big data analytics. These technologies can help improve prediction accuracy by identifying patterns and trends in historical data and incorporating real-time information. Furthermore, integrating forecasting models with decision-making processes can enhance the grid's responsiveness to fluctuations in demand and supply, ultimately leading to a more resilient and efficient energy system. The importance of energy forecasting in smart grid systems cannot be overstated. It is essential to enable effective demand-side management, ensure reliable load shedding, and facilitate optimal resource distribution. However, achieving accurate forecasts in the face of uncertainties and the complexity of SG data remains a significant challenge that requires ongoing research and innovation.

2.5.1 Importance of Load Forecasting

Effective forecasting of load consumption is essential to achieve equilibrium between supply and demand in decentralized energy grids. Accurate predictions allow energy providers and consumers to make informed decisions, optimize resource allocation, and improve overall system performance. In the absence of precise forecasting, systems can face several significant challenges that can undermine their efficiency and sustainability.

- **Inefficient use of storage capabilities:** When load consumption is not accurately predicted, energy storage systems can be under- or over-utilized. This can lead to unnecessary energy loss, as stored energy may be wasted if it is not deployed at the right time. For example, if energy is stored during periods of low demand but is not released during peak usage times, it results in a missed opportunity to effectively balance the grid. Furthermore, frequent cycling of storage systems due to erratic demand can lead to accelerated wear and tear, reducing the lifespan of these assets and increasing maintenance costs.

- **Excessive Dependence on Grid-Supplied Energy :** Inaccurate load forecasting can lead to an overreliance on energy supplied by the traditional grid. When decentralized systems cannot accurately predict local energy needs accurately, they may default to drawing power from the grid, especially during peak demand periods. This dependence can significantly increase operational expenses, as grid-supplied energy is often more expensive than locally generated renewable energy. Furthermore, this dependence can strain the central grid, particularly during high-demand periods, leading to potential outages or higher energy prices.

- **Ineffective Energy Distribution Choices within Peer-to-Peer Networks :** Decentralized energy grids often rely on peer-to-peer (P2P) networks for energy distribution, where consumers can buy and sell energy among themselves. However, without accurate load forecasting, these networks may struggle to make effective energy distribution choices. For example, if a community misjudges its energy needs, it may either overproduce energy that goes unsold or underproduce, leading to shortages. This inefficiency can hinder the growth of P2P energy markets, discourage participation, and ultimately limit the potential benefits of decentralized energy systems.

In summary, the importance of effective load consumption forecasting cannot be overstated. Optimizing energy storage, reducing operational costs, and ensuring efficient energy distribution in decentralized energy grids is crucial. Addressing the challenges posed by inaccurate predictions will not only enhance the resilience of these systems but also contribute to a more sustainable and economically viable energy future.

2.5.2 Machine Learning and Deep Learning Approaches

Over the past decade, there has been a notable surge in the interest and application of methods that leverage artificial intelligence (AI), especially those grounded in machine learning (ML) and deep learning (DL) algorithms. This growing fascination can be attributed to the impressive capabilities of these technologies in generating highly accurate predictions and insights, particularly within the context of smart grid (SG) systems [40].

Smart grids represent a modernized electrical grid that utilizes digital communication technology to monitor and manage the transport of electricity from all generation sources to meet the varying electricity demands of end users. The integration of AI into these systems has revolutionized the way energy is produced, distributed, and consumed. Machine learning algorithms, which enable systems to learn from data and improve their performance over time without explicitly programming, have been instrumental in optimizing various aspects of smart grid operations. Specifically, sequence-oriented deep learning algorithms, including recurrent neural networks (RNN) and long-short-term memory (XGBoost) models, have demonstrated significant effectiveness in addressing the non-linear characteristics of energy data that involve extended sequences [41].

The generation of smart grid (SG) data faces a multitude of challenges, particularly due to the inherent stochastic uncertainties that arise in the context of energy systems. One of the primary difficulties stems from the intermittent nature of renewable energy sources, such as solar and wind power. These sources are not only variable but also unpredictable, leading to significant fluctuations in energy generation. For example, solar energy production can be heavily influenced by weather conditions, such as cloud cover, while wind energy generation can vary based on wind speed and direction. This variability complicates the task of accurately forecasting the energy supply.

In addition to the challenges posed by renewable energy generation, fluctuations in energy consumption patterns further exacerbate the situation. Consumer demand for electricity can change rapidly due to various factors, including time of day, seasonal variations, and unexpected events such as extreme weather conditions or economic changes. These unpredictable consumption patterns create additional layers of complexity for energy forecasting, making it difficult to achieve reliable predictions [42–44]. Moreover, addressing parametric or model uncertainties presents a significant obstacle to the forecasting techniques that have been previously employed. Many traditional forecasting methods are inherently deterministic, which means that they provide single-point forecasts without accounting for the range of possible outcomes. This limitation can lead to suboptimal decision making, as stakeholders may not fully understand the risks associated with their forecasts.

In this context, probabilistic forecasting methods emerge as a more effective alternative. Unlike traditional point-prediction approaches, probabilistic methods are designed to produce prediction intervals (PIs), which offer a range of possible outcomes along with their associated probabilities. This capability allows decision-makers to better manage uncertainties and make more informed choices based on the likelihood of various scenarios. For example, a probabilistic forecast might indicate that there is a 70% chance that energy demand will fall within a specific range, providing valuable information for grid operators and energy suppliers [45]. Recent advances in probabilistic deep learning have further enhanced the capabilities of these forecasting methods. Using complex neural network architectures, probabilistic deep learning models can capture intricate patterns in data and provide more accurate and reliable predictions. These models can incorporate a wide array of variables, including historical data, real-time measurements, and external factors, to generate robust forecasts that reflect the uncertainties inherent in energy systems.

In this research, a comprehensive comparative analysis of energy consumption is performed using three distinct predictive modeling techniques. Autoregressive Integrated Moving Average (ARIMA), Extreme Gradient Boosting (XGBoost), and Random Forest and Long-Term-Short-Memory. The primary objective of this analysis is to identify the most effective method for analyzing and predicting energy consumption patterns, which is crucial to optimize energy management and enhance energy systems efficiency. Generation of smart grid data introduces a myriad of challenges, particularly due to inherent stochastic uncertainties associated with energy production and consumption.

One of the most significant factors contributing to these uncertainties is the intermittent nature of renewable energy sources, such as solar and wind power. These sources are subject to variability based on weather conditions, time of day, and seasonal changes, leading to unpredictable levels of energy generation. Consequently, this variability can create difficulties in maintaining a stable energy supply and demand balance. In addition, fluctuations in energy consumption patterns further complicate the analysis. Consumer behavior can be influenced by a variety of factors, including economic conditions, technological advances, and changes in lifestyle. These factors can cause sudden spikes or drops in energy demand, making it challenging to develop accurate predictive models.

To address these complexities, this study employs ARIMA, XGBoost, Random Forest, and XGBoost as analytical tools. ARIMA is a traditional time series forecasting method that is particularly effective in capturing linear trends and seasonality in historical data. In contrast, XGBoost is a powerful machine learning algorithm that excels in handling nonlinear relationships and interactions among variables, making it suitable for complex datasets. Random Forest, another machine learning technique, utilizes an ensemble of decision trees to improve prediction accuracy and robustness against overfitting, and long-short-term memory (XGBoost) networks play a crucial role in forecasting solar energy by adeptly capturing and modeling the temporal dynamics inherent in the data. Their architecture is specifically designed to handle sequences of data, making them particularly effective for time series analysis.

By comparing the performance of these three methods, the study aimed to determine which approach produces the most reliable and accurate predictions of energy consumption. This analysis not only contributes to the existing body of knowledge in energy forecasting but also provides valuable information for policymakers, energy providers, and researchers seeking to improve the efficiency and reliability of smart grid systems. Ultimately, our findings may facilitate better decision-making processes in energy management, leading to more sustainable and resilient energy systems in the face of increasing demand and the increasing integration of renewable energy sources.

An Overview of ARIMA Model:

The ARIMA models represent a sophisticated synthesis of three key components: Autoregressive (AR), Integrated (I), and Moving Average (MA). Each of these components plays a vital role in the overall functionality of the model. The autoregressive part relies on the relationship between an observation and a number of lagged observations (previous time points), allowing the model to capture the influence of past values on current outcomes. The integrated component addresses the need for stationarity in time-series data by differencing the data, which helps to stabilize the mean of the time series by removing trends or seasonality. Finally, the Moving Average component incorporates the dependency between an observation and a residual error from a moving average model applied to lagged observations.

Together, these elements create a powerful framework for forecasting time-series data that exhibit relative stability, making ARIMA a widely used and trusted method in various fields such as economics, finance, and environmental science. The acronym ARIMA stands for Autoregressive Integrated Moving Average, and it encapsulates the ability of the model to leverage historical data effectively to predict future patterns, providing valuable information for decision-making processes [46].The generalized autoregressive moving average (ARMA) model is a powerful statistical tool that combines two fundamental components: autoregressive (AR) and moving average (MA). The AR component captures the relationship between an observation and a specified number of lagged observations (previous time points), allowing the model to leverage past values to predict future outcomes. However, the MA component focuses on the relationship between an observation and a residual error from a moving average model applied to lagged observations. This duality enables the ARMA model to effectively account for both the inherent trends and patterns in the data, as well as the random shocks or noise that may influence the time series.

The ARIMA (p, d, q) framework further enhances the ARMA model by incorporating an integration (I) component, which is crucial for handling non-stationary time-series data. In this context, the parameters p, d, and q represent the order of the autoregressive part, the degree of differencing required to achieve stationarity, and the order of the moving average part, respectively. The integration aspect (I) specifically addresses the need to transform a nonstationary time series into a stationary one, which is essential for the validity of many time-series forecasting methods. By applying differencing, the model can remove trends and seasonality, allowing more accurate predictions based on the underlying patterns in the data [47, 48].

Autoregression (AR) utilizes the relationship between a current observation and its preceding values to make predictions. In the context of electric load forecasting, this implies that there is an expectation of future consumption in the upcoming hour or the possibility of a peak in demand during that same period. The integrated component (I) of the model (d) encompasses the terms that determine the degree of differentiation applied to the time series data [49]. Additionally, the moving average aspect of the model (q) enables the characterization of the model error as a linear combination of errors recorded at previous time points. The ARIMA (p,d,q) model, which incorporates the lag polynomial L, is represented by Equation (2.1) [50].

$$\left(1-\sum_{i=1}^{p}\phi_i L_i\right)(1-L)^d=\left(1-\sum_{j=1}^{q}\theta_j L_j\right)\varepsilon_t \tag{2.1}$$

Where:

- L_i is the lag operator,
- ϕ_i represents the autoregressive model parameters,
- θ_j represents the moving average parameters,
- ε_t are the error terms.

Overview of Random Forest Model:

Random forest is a supervised learning technique in machine learning. It stands out as one of the most effective techniques in the realm of machine learning. This method operates as an ensemble of classifiers, specifically utilizing decision trees that are created through two distinct forms of randomization. The initial step involves training each decision tree on a random sample drawn with replacement from the original dataset, ensuring that the sample size matches that of the training set. In the realm of energy management, Random Forest regression is often favored for load forecasting due to its distinctive benefits. As an ensemble learning approach, it generates multiple decision trees during the training phase and merges their predictions.

This process significantly mitigates the likelihood of overfitting while improving the model's generalization capabilities, which is particularly beneficial when dealing with the intricate datasets typically associated with load forecasting. The enhancement of classification accuracy in machine learning can be significantly attributed to the development of ensemble methods, particularly those that utilize decision trees. These ensembles operate on the principle of combining the predictions of multiple individual trees to arrive at a more robust and accurate classification. The collective decision-making process is often facilitated by a voting mechanism, in which each tree casts a vote for a particular class and the class with the majority of votes is selected as the final prediction. This approach helps mitigate the weaknesses of individual trees that may be prone to overfitting or high variance.

To effectively construct these ensembles, random vectors are frequently employed to introduce variability in the growth of each individual tree. This randomness is crucial, as it ensures that the trees are diverse, which in turn enhances the overall performance of the ensemble. One of the most notable techniques in this domain is bagging, short for bootstrap aggregation, which was introduced in [51].Bagging involves creating multiple decision trees by randomly selecting subsets of examples from the training dataset. This selection is done with replacement, meaning that the same example can be chosen multiple times for a single tree. By training each tree on a different subset of data, bagging reduces the variance of the model and improves its generalization to unseen data.

Another influential method in the realm of ensemble learning is random split selection, proposed in [52]In this approach, the decision for each node's split in a decision tree is made randomly from the top K potential splits, rather than selecting the best split based on a specific criterion. This randomness in the splitting process contributes to the diversity of the trees in the ensemble, further enhancing the model's ability to generalize. Moreover, Breiman's work in [53]introduced an innovative technique that involved generating new training datasets by randomizing the outputs of the original training set. This method, known as "random forests," combines the principles of bagging with the idea of introducing randomness in the output labels, thereby creating a more varied set of training examples for each tree. The result is a powerful ensemble model that leverages the strengths of multiple decision trees while minimizing their individual weaknesses.In a random forest model, two critical parameters play a significant role in determining the model's performance and accuracy: the total number of trees, commonly referred to as ntree, and the number of variables considered for partitioning at each node, known as mtry.

Total Number of Trees (ntree) :

- The parameter (ntree) specifies how many individual decision trees will be constructed in the random forest. Each tree is built using a random subset of the training data, which helps to ensure diversity among the trees.
- A higher number of trees generally leads to better model performance, as it reduces the variance and helps to mitigate overfitting. However, there is a trade-off, as increasing the number of trees also increases the
- computational cost and time required for training and prediction,

- It is common practice to experiment with different values of ntree to find an optimal balance between performance and computational efficiency. In many cases, a value between 100 and 500 trees is often used, but this can vary depending on the specific dataset and problem.

Number of Variables for Partitioning (mtry)

- The mtry parameter determines how many features (or variables) are randomly selected to consider for splitting at each node of the decision trees. This randomness is a key aspect of the random forest algorithm, as it helps to create diverse trees that capture different patterns in the data.
- The choice of mode can significantly impact the model's performance. If mtry is set too low, the trees may not capture enough information, leading to underfitting. Conversely, if mtry is set too high, the trees may become too similar to each other, resulting in over-fitting.
- A common approach is to set mtry to the square root of the total number of features for classification tasks and to one-third of the total number of features for regression tasks. However, tuning this parameter through cross-validation can lead to improved model performance.

The key parameters in a random forest model are the total number of trees, known as ntree, and the number of variables considered for partitioning at each node, known as mtry shown in equation (2.2):

$$Y = h(X) = \frac{1}{n_{\text{tree}}} \sum_{i=1}^{n_{\text{tree}}} h_i(X) \tag{2.2}$$

Where:

- Y is the predicted output,
- $h(X)$ is the function representing the ensemble of trees,
- n_{tree} is the number of trees in the ensemble,
- $h_i(X)$ is the prediction from the i-th tree.

Overview of XGBoost Model:

XGBoost is an ensemble learning technique that uses decision trees and operates within a gradient-boosting framework as its core data structure. It is frequently used in the field of load forecasting, where accurate predictions are essential. Its ability to handle large datasets and complex relationships makes it a preferred choice among data scientists and machine learning practitioners[54]. The primary objective of ensemble classification is to improve the accuracy of classification outcomes by constructing multiple classifiers that work together to make predictions. This approach takes advantage of the strengths of various models to mitigate the weaknesses of individual classifiers, thereby improving overall performance. However, a significant challenge in this classification process is the issue of data missingness. Many classification algorithms, including traditional ones, rely on complete datasets for effective performance. When data is missing, it can lead to biased estimates, reduced model precision, and ultimately poor classification results [55].

To address this challenge, advanced methods such as XGBoost have been developed. XGBoost, which stands for Extreme Gradient Boosting, is a powerful gradient-boosted decision tree method that incrementally trains decision trees using the available training data. This means that instead of building a single model from the entire dataset, XGBoost constructs a series of decision trees sequentially. At each iteration, a new decision tree is introduced that builds on the predictions of the previously established trees. This iterative process allows the model to refine its predictive capabilities, as each new tree focuses on correcting the errors made by the ensemble of trees that preceded it. The optimization process in XGBoost is guided by an objective function that the method aims to minimize. This objective function consists of two key components: the loss term (l) and the regularization term (omega).

The loss term quantifies how well the model's predictions align with the actual outcomes, while the regularization term helps to prevent overfitting by penalizing overly complex models. By optimizing this combined function, XGBoost seeks to maximize the overall value of the goal function, which ultimately leads to improved classification results. This dual focus on accuracy and model complexity allows XGBoost to achieve high performance even in the presence of missing data, making it a robust choice for ensemble classification tasks. Equation (2.3) shows the objective function of the t^{th} iteration L^t, yi is the actual class label of instance I and $\hat{y}_i$ is the predicted class label of instance i, fk denotes the tree function, n and Ω, represents the number of instances in the training set and the regularization term, respectively. The model complexity penalizes by $\Omega(f_t\ )$ to prevent overfitting. Gamma and lambda are the hyperparameters, where T and w are the number of leaves of the trees and the weight of the leaves, respectively [56].

$$L^t = \sum_{i=1}^{n} \left[l(y_i, \hat{y}_i^{(t-1)}) + f_t(x_i) \right] + \Omega(f_t) \tag{2.3}$$

Where:

- L^t is the loss at iteration t,
- $l(y_i, \hat{y}_i^{(t-1)})$ is the loss function comparing the true value y_i and the predicted value $\hat{y}_i^{(t-1)}$,
- $f_t(x_i)$ is the prediction function at iteration t for input x_i,

- $\Omega(f_t)$ is the regularization term.

The regularization term is defined as shown in (2.4):

$$\Omega(f_t) = \gamma T_{t+1} + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \tag{2.4}$$

Where:

- γ is a parameter controlling the complexity of the model,
- T_{t+1} is the number of leaves in the tree at iteration $t+1$,
- λ is the regularization parameter,
- w_j are the weights associated with the leaves.

An Overview of XGBoost Model:

Long-Short-Term Memory (XGBoost) networks represent a significant advancement within the domain of deep learning, particularly in the context of processing sequential data. These networks are a specialized form of Recurrent Neural Networks (RNNs) that have been meticulously designed to tackle the challenges posed by the vanishing gradient problem, a common issue that hampers the performance of traditional RNNs. The vanishing-gradient problem occurs when gradients used in the training of neural networks become exceedingly small, effectively preventing the network from learning long-range dependencies in the data. This limitation is particularly pronounced in tasks that require the retention of information over extended periods, such as language modeling, time series prediction, and speech recognition.

XGBoosts are specifically engineered to overcome these limitations, allowing them to effectively capture long-term dependencies in sequential data. Unlike traditional RNNs, which struggle to maintain relevant information over time, XGBoosts utilize a unique architecture that enables them to remember information for longer durations. This capability is crucial for applications where context and historical data play a vital role in making accurate predictions. In addition, XGBoosts are particularly adept at processing continuous values and mitigating noise, which enhances their performance in real-world scenarios where data can be unpredictable and variable. One of the key advantages of XGBoosts is their ability to eliminate the need to maintain a fixed set of states, allowing greater flexibility in handling varying input sequences and adapting to different contexts [57]. The architecture of XGBoosts is characterized by three essential components: input gates, forget gates, and output gates. These gates work in concert to regulate the flow of information from, through, and within the XGBoost unit. The input gate determines which information from the current input should be added to the cell state, while the forget gate decides which information should be discarded from the cell state. The output gate, on the other hand, controls which information from the cell state should be output to the next layer in the network. The management of unit storage within the XGBoost is facilitated by a mechanism known as the cell state, which acts as a conveyor of information over time steps.

The operation of the XGBoost at any given time step, denoted as t, can be articulated through a series of mathematical processes that govern the flow of information. These processes involve the application of various activation functions, such as the sigmoid and hyperbolic tangent functions, to compute the gates and cell state values [58]:

$$
\begin{aligned}
i_t &= \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \\
f_t &= \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \\
o_t &= \sigma\left(W_o \cdot [h_{t-1}, x_t] + b_o\right) \\
\overline{C_t} &= \tanh\left(W_c \cdot [h_{t-1}, x_t] + b_c\right) \\
C_t &= f_t \cdot C_{t-1} + i_t \cdot \overline{C_t} \\
h_t &= o_t \cdot \tanh\left(C_t\right)
\end{aligned}
$$

Where:

- i_t is the input gate,
- f_t is the forget gate,
- o_t is the output gate,
- C_t is the cell state,
- h_t is the hidden state,
- W are weight matrices,
- b are bias terms,
- σ is the sigmoid activation function,
- tanh is the hyperbolic tangent activation function

Our findings align with the existing literature, suggesting that long- and short-term memory (XG-Boost) models exhibit enhanced predictive capabilities compared to other forecasting methods. Specifically, our analysis reveals that these models are adept at anticipating a modest increase in energy consumption. This observation highlights the shifting trends in consumer behavior, suggesting that individuals and households adapt their energy use patterns in response to various factors, such as economic conditions, technological advancements, and environmental awareness. The ability of XGBoost models to capture these evolving dynamics not only reinforces their effectiveness in energy consumption forecasting, but also emphasizes the importance of understanding consumer behavior in the context of energy management and sustainability efforts. In continuing to explore these patterns, it becomes increasingly clear that leveraging advanced predictive modeling techniques such as XGBoost can provide valuable insights for policymakers, energy providers, and researchers aiming to address the challenges associated with energy demand and consumption.

2.6 Optimization Algorithms for Energy Storage Sizing

The integration and management of an electrical power system are inherently complex due to several factors, particularly the elevated costs and unpredictable characteristics associated with intermittent renewable energy (RE) resources. Unlike traditional energy sources, which provide steady and controllable output, renewable energy sources such as solar and wind are subject to fluctuations based on environmental conditions. This variability can present challenges in maintaining a consistent power supply, necessitating advanced integration strategies. As a result, careful planning and coordination among various components of the power system are essential to ensure that it operates both effectively and efficiently. This involves not only the technical aspects of energy generation and distribution but also considerations related to economic factors, regulatory frameworks, and the evolving landscape of energy demand [59].

To maximize the economic advantages of a storage system, it is crucial to determine the optimal size of the battery. This calculation is inherently complex as it must take into account various factors such as battery efficiency, its lifecycle, and the prevailing electricity tariffs, all of which are influenced by the dimensions of the battery. The efficiency of a battery refers to the way well it can convert and store energy, directly affecting the overall cost-effectiveness of the system. Higher efficiency means less energy is lost during the charging and discharging processes, leading to greater savings over time. In addition, the battery life cycle, which encompasses its lifespan and the number of charge-discharge cycles it can endure before its performance degrades, plays a significant role in determining the long-term economic viability of the storage solution. Furthermore, electricity tariffs can fluctuate based on demand, time of day, and other market dynamics, so it is essential to consider these variables when sizing the battery. The interplay of these factors creates a multifaceted challenge, as the optimal battery size must strike a balance between initial investment costs, operational efficiency, and potential savings from energy arbitrage [60–63].

The capital expenditure associated with a Battery Energy Storage System (BESS) within a specific microgrid is influenced by two primary cost components: the power rating and the energy rating. The power rating refers to the maximum output the system can deliver at any given moment, while the energy rating indicates the total amount of energy the system can store and subsequently release. When a BESS is oversized, it has a higher power and energy capacity than necessary; this leads to increased investment costs, as larger systems typically require more expensive components and infrastructure.

Conversely, if the BESS is undersized, it may not be able to meet the operational demands of the Microgrid, resulting in potential energy shortages, reliability issues and frequently reliance on the grid. As the size of the BESS increases, the investment costs tend to rise linearly, reflecting a direct correlation between size and cost. However, operating costs, which include maintenance, management, and energy losses, tend to decrease in a non-linear manner as the system scales. This is often due to economies of scale, where larger systems can operate more efficiently. Therefore, the optimal configuration of a BESS is achieved at a specific point where the rising investment costs are balanced against the declining operating costs, ensuring that the system is both economically feasible and capable of meeting the energy demands of the microgrid [64].

2.6.1 Optimization Techniques for Energy Storage Sizing

Numerous researchers explore and implement various methodologies to determine the dimensions and positioning of a Battery Energy Storage System (BESS). These approaches encompass a variety of analytical and computational techniques, each tailored to address specific requirements and constraints associated with the deployment of BESS. The diversity in methods reflects the complexity of the factors that influence optimal size and site selection for energy storage solutions, highlighting the ongoing advances in this field of study [65, 66].

These studies indicate that optimal sizing techniques can be categorized into two primary types: single-objective and multi-objective optimal sizing methods. The focus of single-objective optimal methods is primarily on cost reduction, with the aim of achieving the lowest possible costs associated with the system. These methods typically involve analyzing various components of the power system, such as generation capacity, transmission infrastructure, and distribution networks, to identify the most cost-effective configurations. By focusing solely on minimizing costs, single-objective methods can overlook other critical factors, such as system reliability, environmental impact, and operational efficiency. Consequently, while they can lead to significant savings, they may not always result in the most sustainable or resilient power systems. This study examined various methods for optimizing energy storage. Among the approaches evaluated, several techniques were identified as significant for improving the efficiency and effectiveness of energy storage systems.

Linear Programming (LP) Sizing Technique:

Linear programming is a powerful mathematical optimization technique widely used in various fields, including energy management, to determine the most effective configuration of an energy storage system. The primary goal of this method is to identify the optimal dimensions of the storage system that can minimize costs or maximize benefits, depending on the specific objectives of the project.

To achieve this, a mathematical model is constructed that incorporates various factors and constraints relevant to the energy storage system. These constraints may include power demand, which refers to the amount of energy required by consumers at any given time; energy generation, which encompasses the amount of energy produced by renewable sources such as solar or wind; and storage capacity, which defines the maximum amount of energy that can be stored in the system. By carefully defining these parameters, the model can accurately reflect the operational environment of the energy storage system [67].

The linear programming approach involves formulating an objective function, which is a mathematical expression that quantifies the optimization goal, whether it is to minimize costs associated with energy procurement, maintenance and operation or to maximize the benefits derived from energy sales, efficiency and reliability. The optimization process then seeks to find the values of decision variables, such as the size of the storage system, that will yield the best possible outcome while satisfying all the established constraints.

This systematic analysis not only helps determine the most efficient storage size, but also provides insights into the trade-offs between energy storage capacity and operational costs. By exploring different scenarios and adjusting the parameters within the model, stakeholders can assess how changes in demand, generation patterns, or storage technologies could impact the overall performance and economics of the energy storage system.

Genetic Algorithm (GA) Sizing Technique :

he genetic algorithm, as defined within the realm of artificial intelligence, functions as a search heuristic that emulates the principles of natural selection. This approach embodies a sophisticated method of conducting a random search to address optimization challenges. It mirrors the concept of survival of the fittest, where individuals compete over successive generations within a dynamic environment to find solutions to specific problems. Each individual is represented by a chromosome, which is composed of genes that denote its characteristics.

In this context, each individual signifies a point within a search space, while each generation comprises a population of strings (chromosomes) that encode potential solutions (individuals) to an optimization problem, progressively evolving toward more effective solutions [68]. Genetic algorithms (GAs) are a class of optimization techniques inspired by the principles of natural evolution and selection. At their core, these algorithms mimic the process of natural selection , where the fittest individuals are more likely to survive and reproduce. This fundamental concept is crucial for understanding how GAs operate and evolve a population of potential solutions to a given problem.

In a genetic algorithm, a population consists of a set of candidate solutions, often represented as strings of data, such as binary strings or real-valued vectors. Each individual in this population is subjected to various genetic operators that simulate biological processes. Two primary operators are mutation and crossover (or recombination) [69].

- Mutation: Introduces random changes to an individual's genetic makeup, which helps maintain genetic diversity within the population and allows the algorithm to explore new areas of the solution space. For example, in a binary representation, mutation might involve a flipping a bit from 0 to 1 or vice versa. This operator is essential to prevent premature convergence to suboptimal solutions by ensuring that the algorithm does not become trapped in local optima.

- Crossover: on the other hand, combines the genetic information of two parent individuals to produce one or more offspring. This process is similar to biological reproduction, in which traits from both parents are mixed to create new individuals. Crossover can take various forms, such as one-point crossover, where a single crossover point is selected, or multipoint crossover, where multiple points are chosen to exchange segments of the parent strings. This operator is vital for exploiting existing solutions in the population and generating potentially superior offspring.

A fitness function is used to assess the effectiveness of each individual. This function quantifies how well a solution solves the problem at hand, assigning a fitness score based on predefined criteria. The fitness levels of individuals determine their likelihood of being selected for reproduction in the next generation. Higher fitness scores correspond to a greater chance of being chosen, thus ensuring that the most promising solutions contribute to the next generation.

The fundamental Key Steps procedure of a genetic algorithm:

- Initialization : A random population of individuals is generated, each representing a potential solution to the problem.

- Evaluation : The fitness of each individual is assessed using the fitness function.

- Selection : Individuals are selected for reproduction based on their fitness scores. Various selection methods can be used, such as roulette wheel selection, tournament selection, or rank-based selection, each with its own advantages and disadvantages.

There are notable distinctions associated with genetic algorithms. Firstly, these algorithms function on an entire population of solutions, which enhances the likelihood of achieving a global optimum and diminishes the possibility of being confined to a local stationary point. Secondly, genetic algorithms work with encoded representations of the problem parameters rather than the parameters in their original form. Additionally, standard genetic algorithms can be utilized for a wide range of optimization problems, whether discrete or continuous, as they do not rely on supplementary information regarding the objective function's value, such as derivatives. Lastly, genetic algorithms incorporate probabilistic approaches in their transition operators, in contrast to traditional methods that utilize deterministic transition operators.

Genetic Algorithm (GA) Pseudocode

Algorithm 1 Genetic Algorithm

```
Initialize population with random chromosomes
Evaluate fitness of each individual while termination condition not met
Select parents from population
Apply crossover to generate offspring
Apply mutation to offspring
Evaluate fitness of new population
Select individuals for next generation (elitism can be used)
```

2.6.2 Particle Swarm Optimization (PSO) Sizing Technique :

The particle swarm optimization (PSO) algorithm is a stochastic optimization method inspired by the collective behavior of swarms introduced by Eberhart and Kennedy in 1995. This algorithm mimics the social interactions observed in various animal groups, such as insects, flocks of birds, and schools of fish. Within these swarms, individuals work collaboratively to locate food sources, with each member continuously adapting its search strategy based on personal experiences and the insights gained from other group members [70]. The primary design concept of the Particle Swarm Optimization (PSO) algorithm is fundamentally linked to two areas of research. The first is evolutionary algorithms; similar to these algorithms, PSO employs a swarm approach that enables it to explore extensive areas within the solution space of the optimization objective concurrently. The second area pertains to artificial life, which examines systems that exhibit characteristics akin to living organisms.

2.6.3 PSO Pseudocode

Algorithm 2 Particle Swarm Optimization (PSO)

%State Evaluate fitness of each particle

for all particle $p \in \mathscr{P}$ **do**

  $v_p \leftarrow w v_p + c_1 r_1 \big(pBest_p - position_p\big) + c_2 r_2 \big(gBest - position_p\big)$

  $position_p \leftarrow position_p + v_p$

  Evaluate fitness of p

  if fitness(p) better than $pBest_p$ **then**

    $pBest_p \leftarrow position_p$

  end if

end for

Update $gBest$ based on all $pBest$

In the context of artificial life theory, which investigates the behaviors of social animals, Millonas introduced five essential principles aimed at developing swarm artificial life systems that exhibit cooperative behavior through computational means [71]. The subsequent section provides a comprehensive overview of the Particle Swarm Optimization (PSO) algorithm. In the context of a continuous space coordinate system, the PSO can be mathematically articulated in the following manner when assuming that a swarm consisting of N particles, with each particle represented by a position vector in a D-dimensional space, is: The position vector of the *i*-th particle is given by the equation in (2.5):

$$X_i = (x_{i1}, x_{i2}, \cdots, x_{id}, \cdots, x_{iD}) \tag{2.5}$$

The velocity vector of the *i*-th particle is given as in (2.6):

$$V_i = (v_{i1}, v_{i2}, \cdots, v_{id}, \cdots, v_{iD}) \tag{2.6}$$

Where:

- X_i represents the position vector of the i-th particle in a D-dimensional space,
- V_i represents the velocity vector of the i-th particle in the same D-dimensional space.

The individual's optimal position (2.7) (i.e., the optimal position that the particle has experienced) is:

$$P_i = (p_{i1}, p_{i2}, \cdots, p_{id}, \cdots, p_{iD}) \tag{2.7}$$

Where:

- P_i represents the position vector of the i-th particle in a D-dimensional space,
- p_{id} denotes the position of the i-th particle in the d-th dimension.

Without loss of generality, taking the minimizing problem as an example, in the initial version of the PSO algorithm, the update formula for the optimal position of the individual is given as (2.8):

$$p_{i,t+1}^{d} = \begin{cases} x_{i,t+1}^{d}, & \text{if } f(X_{i,t+1}) < f(P_{i,t}) \\ p_{i,t}^{d}, & \text{otherwise} \end{cases} \tag{2.8}$$

Where:

- $p_{i,t+1}^d$ is the updated position of the i-th particle in dimension d at time $t+1$,
- $x_{i,t+1}^d$ is the new candidate position of the i-th particle in dimension d at time $t+1$,
- $f(X_{i,t+1})$ is the fitness function evaluated at the new position,
- $f(P_{i,t})$ is the fitness function evaluated at the current position,
- $p_{i,t}^d$ is the current position of the i-th particle in dimension d at time t.

And the update for the position is given by equation 2.9:

$$x_{i,t+1}^d = x_{i,t}^d + v_{i,t+1}^d \tag{2.9}$$

Where:

- $x_{i,t+1}^d$ is the updated position of the i-th particle in dimension d at time $t+1$,
- $x_{i,t}^d$ is the current position of the i-th particle in dimension d at time t,
- $v_{i,t+1}^d$ is the velocity of the i-th particle in dimension d at time $t+1$.

The original iteration of the Particle Swarm Optimization (PSO) algorithm demonstrated limited effectiveness in addressing optimization challenges, prompting the development of a modified PSO algorithm by [72] shortly after the initial proposal. This revised algorithm incorporated an inertia weight into the velocity update equation, resulting in an enhanced velocity update formula.The velocity update equation (2.10) for the i-th particle in dimension d is given by:

$$v^d_{i,t+1} = \omega \cdot v^d_{i,t} + c_1 \cdot \text{rand} \cdot (p^d_{i,t} - x^d_{i,t}) \\ + c_2 \cdot \text{rand} \cdot (p^d_{g,t} - x^d_{i,t}) \tag{2.10}$$

Where:

- $v^d_{i,t+1}$ is the updated velocity of the i-th particle in dimension d,
- ω is the inertia weight,
- c_1 and c_2 are the cognitive and social coefficients, respectively,
- rand is a random number in the range [0, 1],
- $p^d_{i,t}$ is the individual best position of the i-th particle,
- $p^d_{g,t}$ is the global best position.

The following are some advantages of PSO over GA:

In this section a comparative analysis of two well-known metaheuristic algorithms is provided: Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) (Table 2.1). PSO is a heuristic search technique that draws inspiration from the collective behavior observed in natural swarms. In figure 2.1, the Classification of Metaheuristic Optimization Techniques chart is shown. Both PSO and GA share similarities, as they are population-based search methods that rely on the exchange of information among their members to improve their search efficiency through a blend

Table 2.1 Comparison of Advantages of PSO over GA

Genetic Algorithm	**Particle Swarm Optimization**
Requires genetic operators such as crossover, mutation, and selection.	Only requires a few parameters to adjust, making it easy to implement.
High computational cost.	Minimizes computational cost.
Checks only the present fitness function.	Checks both local and global fitness functions.
Does not have memory to store previous fitness values.	Has memory to store previous fitness values.
Difficult to implement.	Easy to implement.

of deterministic and probabilistic strategies. Over time, both algorithms have been employed to address a variety of problems and have established themselves as robust methodologies with numerous variations and applications. Consequently, integrating these two approaches could yield significant benefits.

Hybrid GA-PSO Models:

The limitations associated with current numerical methods in optimization have led researchers to increasingly rely on computational intelligence techniques. Genetic algorithms (GAs) and particle swarm optimization (PSO) algorithms offer a powerful and effective means of addressing intricate real-world challenges. This research employs a hybrid approach that combines GAs and PSO algorithms to determine the optimal cost of hybrid energy systems (HES). Both GA and PSO are population-based optimization techniques, each possessing distinct advantages and disadvantages.

For instance, while GA is particularly adept at tackling multi-objective problems and demonstrates considerable robustness, it suffers from slow convergence rates due to its reliance on evolutionary operators such as selection, crossover, and mutation, which necessitate numerous function evaluations.

In contrast, PSO is characterized by rapid convergence, as it utilizes mathematical operators for solution generation, resulting in simpler coding compared to GA. However, PSO's primary limitation is its tendency for premature convergence, which can lead to a lack of diversity. Therefore, an ideal algorithm would integrate the strengths of both methods while mitigating their weaknesses, achieving rapid convergence and high diversity. Consequently, this study implements a hybrid algorithm that merges GA and PSO for effective allocation and sizing of renewable energy sources.The hybrid Genetic Algorithm-Particle Swarm Optimization (GA-PSO) introduced in this work represents a novel integration of two well-established optimization techniques, which have been extensively discussed in the literature, particularly in references [73]. This innovative approach leverages the strengths of both algorithms to improve the efficiency and effectiveness of the optimization process.

At its core, the hybrid GA-PSO employs a randomized global search mechanism that combines the exploratory capabilities of the Genetic Algorithm (GA) with the exploitative strengths of Particle Swarm Optimization (PSO). The GA is known for its ability to explore a vast search space through mechanisms such as selection, crossover, and mutation, which mimic the process of natural evolution. However, PSO is inspired by the social behavior of birds and fish, utilizing a swarm of particles that move through the search space, adjusting their positions based on their own experiences and those of their neighbors.

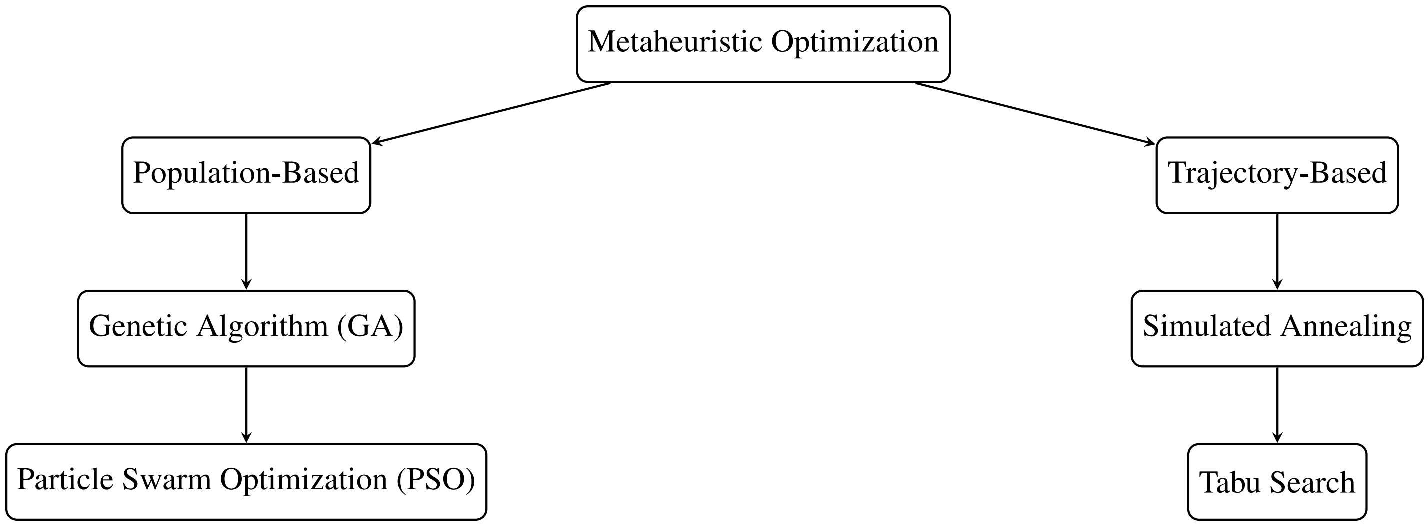

Figure 2.1 Classification of Metaheuristic Optimization Techniques

2.7 Conclusion

The hybrid GA-PSO algorithm effectively combines the strengths of both Genetic Algorithms and Particle Swarm Optimization, leading to improved optimization performance in various applications.

2.8 Energy Sharing and Trading in Distributed Systems

The integration of renewable energy sources, such as solar and wind, plays a crucial role in this transformation. These sources are inherently variable, which requires the use of battery storage solutions to ensure a stable and reliable energy supply. By storing excess energy generated during peak production times, consumers can utilize this stored energy during periods of high demand or low generation, thus optimizing their energy usage and reducing costs.

Advanced trading frameworks, including those based on game-theoretic models, facilitate efficient allocation of energy resources among consumers. These models analyze strategic interactions between participants in the energy market, allowing the development of pricing structures that reflect the dynamics of supply and demand in real time. Using these models, decentralized energy systems can achieve optimal pricing that benefits both producers and consumers, fostering a more equitable energy distribution.

2.9 Game Theory Appproaches to Energy sharing:

The future success of a decentralized energy framework is significantly dependent on the active participation of various stakeholders. These stakeholders can be classified into three primary roles: energy buyers, who are consumers; energy sellers, who are the suppliers; or individuals who fulfill both roles [74]. Energy buyers encompass a wide range of participants, including residential households, commercial enterprises, and industrial facilities, all of whom seek reliable and cost-effective energy solutions. However, energy sellers can include traditional utility companies, renewable energy producers, and even individual consumers, those who consume and generate energy. As these participants engage with one another, the decentralized energy framework evolves into a dynamic marketplace where diverse entities interact strategically to facilitate energy transactions, whether through buying or selling. This interaction not only enhances the efficiency of energy distribution but also encourages innovation in energy technologies and services, ultimately leading to a more resilient and sustainable energy ecosystem.

Game theory is a sophisticated mathematical discipline that delves into the intricacies of strategic interactions among participants, commonly referred to as players [75], [76]. These players, who can represent individuals, companies, or even nations, engage with each other in various contexts, such as the trading of specific products such as energy. In this environment, they are faced with the critical task of determining both the quantity of the product to be traded and the price at which the transaction will occur. The decisions made by these players are not made in isolation; rather, they are heavily influenced by a range of offers or strategies that they propose. These strategies are formulated on the basis of the players' expectations of the benefits they will derive from the trade, which are often quantified in terms of utility or payoff. This interplay of strategies and expected outcomes forms the foundation of game theory, allowing for a structured analysis of decision-making in competitive environments [77].

In the context of a game, a player can be defined as an entity, whether an organization or an individual, that has the capacity to determine a strategic approach and make informed decisions. Numerous research efforts have shown that game-theoretic models play a crucial role in enhancing various important aspects of the energy sector. These models provide a structured framework for analyzing and understanding the complex strategic interactions that occur between different stakeholders, including consumers, grid operators, and energy producers. By applying game theory, researchers and practitioners can better comprehend how these entities make decisions in response to one another, leading to more informed and effective strategies for energy management.

One of the key benefits of game-theoretic models is their ability to enhance the strategic interactions between consumers and grid operators. In traditional energy markets, these interactions can often be adversarial, with consumers seeking to minimize their costs, while grid operators aim to maximize efficiency and reliability. However, game theory fosters a more collaborative environment by encouraging stakeholders to consider the potential benefits of cooperation. For instance, through mechanisms such as demand response programs, consumers can adjust their energy usage in response to real-time pricing signals from grid operators, leading to a win-win situation where both parties benefit from reduced costs and improved grid stability.Moreover, game-theoretic models promote the efficient distribution of surplus energy, which is particularly important in the context of renewable energy sources. As more consumers generate their own energy through solar panels or wind turbines, there is often a surplus of energy that can be shared with the grid or other consumers. Game theory provides insights into how this surplus can be allocated in a way that maximizes overall welfare, reducing reliance on the grid, and minimizing waste.

In addition to fostering collaboration and efficient energy distribution, game-theoretic models also introduce dynamic pricing strategies that encourage greater consumer participation. Traditional fixed pricing structures can lead to inefficiencies and disincentivize consumers to adjust their energy usage based on supply and demand fluctuations. However, by implementing dynamic pricing models informed by game theory, grid operators can signal to consumers when to use or conserve energy, thereby optimizing overall consumption patterns. This not only helps balance supply and demand, but also incentivizes consumers to participate more actively in energy management, leading to a more responsive and adaptive energy system.

In summary, applying game-theoretic models in the energy sector offers significant advantages by enhancing strategic interactions, promoting efficient energy distribution, and introducing dynamic pricing strategies.

The set of players participating in a game process is represented in (2.11).

$$N = \{1, 2, \ldots, n\} \tag{2.11}$$

This expression denotes the participation of n players in the game. Strategies are essential in the game. The viability of a strategy, as well as the number of tactics available to players, is determined by the knowledge and resources gathered during the game.

The set of strategies (2.12) available to a player is represented as:

$$s = \{s_1, s_2, \ldots, s_n\} \tag{2.12}$$

The entire strategy set produced by all players' separate strategies (2.13)is denoted as follows:

$$S = \{S_1, S_2, \ldots, S_n\} \tag{2.13}$$

The entire set of strategies produced by all players' separate strategies is denoted as follows: Profit refers to the rewards that players receive after the game. Typically, this is the maximum advantage that someone can obtain, though it may also take negative values.

The benefits of all players in the game can be expressed as in (2.14)

$$u = \{u_1, u_2, \ldots, u_n\} \tag{2.14}$$

This expression captures the profit for each player in the game.

2.9.1 Nash Equilibrium in Energy Sharing :

Game models can be categorized into two primary types: classical games and evolutionary games. Within the realm of classical games, further distinctions can be made based on various criteria, leading to classifications such as cooperative versus non-cooperative games, as well as static versus dynamic games. The differentiation between cooperative and noncooperative games hinges on whether the players establish a cooperative relationship during the game [78].

In cooperative games, the focus lies on two fundamental aspects: the allocation of benefits arising from cooperation and the adherence to mutually agreed-upon conditions throughout the cooperative process. These elements are crucial to understanding how players interact and achieve collective outcomes. The dynamics of cooperation are essential for fostering collaboration among participants, which can lead to more favorable results for all involved [79].

Conversely, non-cooperative games are primarily characterized by the concept of Nash equilibrium, where players make decisions independently, often leading to competitive rather than collaborative outcomes. In this framework, the strategies used by the players are influenced by the anticipated actions of others, resulting in a scenario where cooperation is not a central feature. This distinction highlights the varying approaches and results that can arise from different types of game models [80].The Nash equilibrium algorithm is a fundamental concept in the realm of non-cooperative game theory, frequently employed to analyze and predict the behavior of players in competitive environments [81]. In these settings, players are often motivated by self-interest, and the Nash equilibrium provides a critical framework for identifying a stable state where no player has an incentive to unilaterally change their strategy. This equilibrium point is essential for achieving balance among competing players as it reflects a situation where each participant's strategy is optimal given the strategies of others. In such environments, it becomes imperative for all players to carefully evaluate their own strategies in relation to the anticipated actions of their competitors, ensuring that their decisions are informed by the competitive landscape [82].

2.9.2 Stackelberg Game for Energy Pricing

Here, the strategies employed by the Energy Market Trading System (EMTS) and the DGs are represented through the buying and selling prices, alongside the energy consumption metrics. These elements serve as critical components of the game strategies, facilitating a structured approach to energy trading and distribution within the microgrid environment. The mathematical formulations that underlie these strategies are integral to understanding the interactions and outcomes of the game. The mathematical formulations that underlie these strategies are integral to understanding the interactions and outcomes of the game [83]:

The Stackelberg game represents a strategic framework in which companies engage in price competition. Within this model, a primary firm, called the leader, establishes its price initially, causing subsequent firms, known as followers, to adjust their pricing strategies in response [84].

Consider a market with one leader firm and n follower firms. Let:

- P_L: Price set by the leader. - P_F: Price set by the follower firms. - Q_L: Quantity produced by the leader. - Q_F: Quantity produced by the follower firms. - $D(P)$: Demand function (2.15), where P is the market price.

$$D(P) = a - bP \tag{2.15}$$

where a and b are positive constants.

2.9.3 Leader's Problem

The leader firm maximizes its profit, which is given in (2.16):

$$\Pi_L = P_L \cdot Q_L - C_L(Q_L) \tag{2.16}$$

where $C_L(Q_L)$ is the leader's cost function. The quantity produced by the leader (2.17)can be expressed as:

$$Q_L = D(P_L) + n \cdot D(P_F) \tag{2.17}$$

Substituting the demand function into the profit function, equation (2.18) is formulated.

$$\Pi_L = P_L \cdot (a - bP_L + n \cdot (a - bP_F)) - C_L(Q_L) \tag{2.18}$$

To find the optimal price P_L^* (2.19), the first derivative of the profit function is taken with respect to P_L and is set to zero:

$$\frac{d\Pi_L}{dP_L} = 0 \tag{2.19}$$

The following companies also maximize their profits, which is given by equation (2.20):

$$\Pi_F = P_F \cdot Q_F - C_F(Q_F) \tag{2.20}$$

The quantity produced by the follower can be expressed as in (2.21).

$$Q_F = D(P_F) \tag{2.21}$$

2.10 Conclusion

This study examines essential significant challenges related to the optimization of renewable energy, energy sharing, and the reduction of dependence on the grid. The increasing demand for sustainable energy solutions has made it imperative to explore innovative methods to harness renewable resources effectively. By focusing on energy sharing, the study highlights the potential for decentralized energy systems that empower consumers to both produce and consume energy, thereby reducing reliance on traditional grid infrastructures. Although it offers important insights into these areas, additional research is required to enhance real-time adaptability, which is crucial for responding to fluctuations in energy supply and demand. Furthermore, the economic viability of these systems must be thoroughly examined to ensure that they are not only sustainable but also financially feasible for widespread adoption. Cybersecurity measures are another critical aspect that needs to be addressed, as the integration of digital technologies in energy systems can expose them to various cyber threats. Lastly, scalability is essential for broader implementation in smart energy systems, as solutions must be adaptable to different contexts and capable of supporting larger populations and diverse energy needs.

2.10.1 Summary of Research Gaps

- Absence of a cohesive framework for forecasting, optimization, and energy-sharing.

- Insufficient research validating models with real consumption data from South Africa.
- A limited number of studies that integrate GA-PSO with LP/GT for both sizing and fair energy distribution.

Chapter 3

Surplus Energy Assessment in a Hybrid Energy System Network

3.1 Executive Summary

The increase in electricity costs in South Africa has profoundly affected low-income households, resulting in significant economic hardships. A substantial portion of their income is allocated to electricity bills, and the increasing tariffs have increased their vulnerability to energy poverty. This research investigates the feasibility and effectiveness of incorporating renewable energy systems into the energy framework of economically disadvantaged communities to mitigate this challenge. The study encompasses both technological and economic dimensions, with the goal of promoting fair access to clean and affordable energy for all. In addition, it highlights potential barriers and challenges associated with the implementation of solar energy in low-income residences, while also offering practical strategies to address these concerns. The analysis assesses the economic practicality of grid-connected hybrid power systems in the southern region of Durban, South Africa. By examining the solar energy potential and using NASA data, the research employs the Hybrid Optimization Model for Electric Renewables (HOMER) program to evaluate the economic viability of the proposed hybrid power system. A comparative analysis of photovoltaic (PV) systems, grid-only configurations, and grid-integrated PV with wind options is performed to determine the financial feasibility of each alternative. The simulation results suggest that the proposed grid-connected hybrid system, which combines photovoltaic energy and wind energy, represents the most effective and economically advantageous solution for the designated area.

3.2 Introduction

South Africa's uncertain electrical supply poses a national dilemma. The economy is crippled by a lack of energy. Due to insufficient generating capacity, current power plants, which are huffing with age and wear, are being run into the ground and failing increasingly frequently. The capacity to generate power at any time has drastically decreased. The construction of the new power plants is years behind schedule. Every month, power outages and blackouts threaten the country. This necessitates the need to attempt to demonstrate what went wrong and how it may be solved. Local air pollution is directly related to energy supply alternatives, with coal and oil products contributing significantly to urban and rural air pollution and acid rain. To address environmental concerns and minimize pollutant emissions into the atmosphere, several governments, including South Africa, have implemented regulations and incentives to develop renewable energy capacity. Furthermore, due to a lack of sufficient electricity production, South African customers have suffered ever-increasing electricity prices and system instability since 2007. As a result, the use of hybrid renewable energy systems (HRESs) has grown in popularity [85]. Renewable energy sources are characterized by their variability and intermittency. To mitigate the challenges posed by these intermittent characteristics, specific design considerations must be implemented, which can lead to an increase in the overall costs associated with renewable energy installations. One effective strategy to enhance the reliability of renewable energy systems is the incorporation of storage solutions and/or backup power sources, ensuring a consistent power supply to meet demand. Furthermore, renewable energy systems often generate surplus energy, which refers to the excess energy produced that remains unused by loads in off-grid and grid-connected configurations[86]. The challenge is addressed through the incorporation of Energy Storage Systems (ESS), which capture surplus energy and subsequently meet power requirements during periods of high demand.[87]

3.2.1 Study Objective

Multiple research studies concerning high-penetration renewable energy systems have indicated that energy production may not reliably satisfy demand. These studies advocate for the absence of excess energy at any time, suggesting that energy storage should be used to transfer surplus energy to periods of deficit. In addition, other research has indicated that incorporating energy efficiency measures could reduce the required capacity for energy storage, thereby shortening the payback period. This approach enhances the cost effectiveness of the system, as energy storage systems would not incur losses. The allowed level of excess electricity in standalone hybrid renewable energy systems is below 5% [11], while the literature suggests that exceeding a surplus 10% indicates a decrease in energy efficiency in renewable systems. Heide et al (2010). identified an optimal configuration for addressing the seasonal variability of renewable energy sources, such as solar and wind, in Europe and quantified the necessary energy storage to balance energy efficiency. Similarly, Budischak et al. conducted simulations to determine the ideal combination of renewable energy sources and energy storage systems, demonstrating that the integration of energy efficiency can reduce the cost of energy. This study aims to investigate the significance of energy efficiency in a hybrid renewable energy system located in the South Coast area of Durban, KwaZulu-Natal. The main goal is to analyze the present condition of energy efficiency in these systems, focusing specifically on the effects of surplus energy production. Additionally, the study intends to assess how energy efficiency impacts total system expenses and operational effectiveness, while also identifying practical approaches to improve cost efficiency and system dependability.

3.3 Methodology

The study considered a community with a daily energy demand of 291.06 kWh/day, with a morning and evening peak demand profile. The initial sizing used in HOMER Pro was: 60 kW PV array, two 10 kW wind turbines, and 500 kWh battery storage. Meteorological data for Durban South Coast (average solar irradiance 5.85 kWh/m^2/day, wind speed 6.14 m/s) were obtained from the NREL Global Solar Atlas and Wind Atlas.

3.3.1 Aim

This section focuses on examining the prevalence and effects of surplus energy in a grid-connected Hybrid Renewable Energy System (HRES). It seeks to clarify the implications of excess energy on the economic aspects of the system, especially regarding energy costs, and to suggest suitable corrective measures to address related inefficiencies. The evaluation is performed using the HOMER energy optimization software, enabling an in-depth analysis of the connection between surplus energy and system efficiency in a scenario characterized by high renewable energy integration.

3.3.2 Meteorological Data

The techno-economic assessment of the proposed energy system for the residential sector was conducted concentrating on the South Coast of Durban in the KwaZulu Natal province of South Africa. 3.1. This region was chosen for several compelling reasons, mainly due to its current deficiency in renewable energy resources. Despite the urgent need for sustainable energy solutions, the South Coast of Durban has not yet fully realized its potential renewable energy opportunities. The assessment aimed to identify and evaluate the various renewable energy technologies that could be developed or established in this area, to facilitate access to efficient and cost-effective power for local residents. The analysis took into account the unique geographical and climatic conditions of the region, which may support the implementation of solar, wind, or other renewable energy systems. The figure in 3.2 shows the configuration of the network under investigation.

Furthermore, the study considered the implications of the surplus energy generated by the proposed system. In a context where energy demand is often unpredictable, understanding how to manage excess energy production is crucial. This includes exploring options for energy storage, grid integration, and potential avenues to sell surplus energy back to the grid or to neighboring communities. By addressing these factors, the assessment not only highlights the feasibility of implementing renewable energy solutions on the South Coast of Durban, but also underscores the broader economic and environmental benefits that could arise from such initiatives. Ultimately, the findings are designed to inform stakeholders, including policymakers, investors, and local communities, about the potential to transform the energy landscape in this region, paving the way for a more sustainable and resilient energy future.

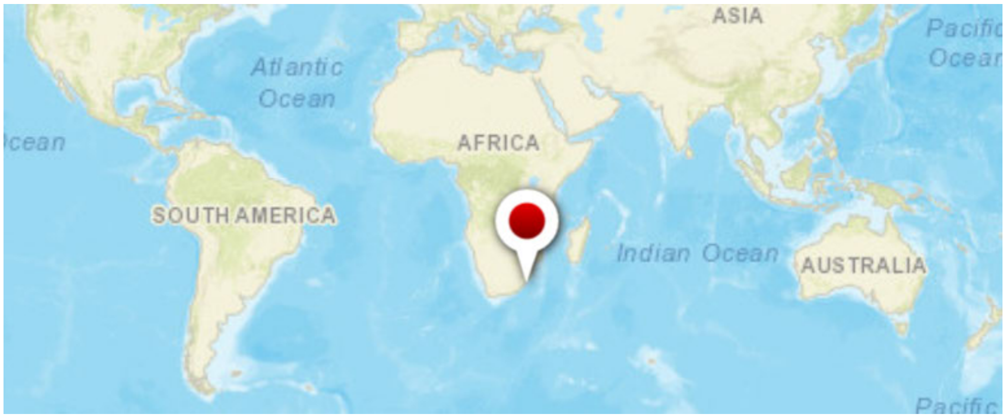

Figure 3.1 Network Configuration Map under Investigation

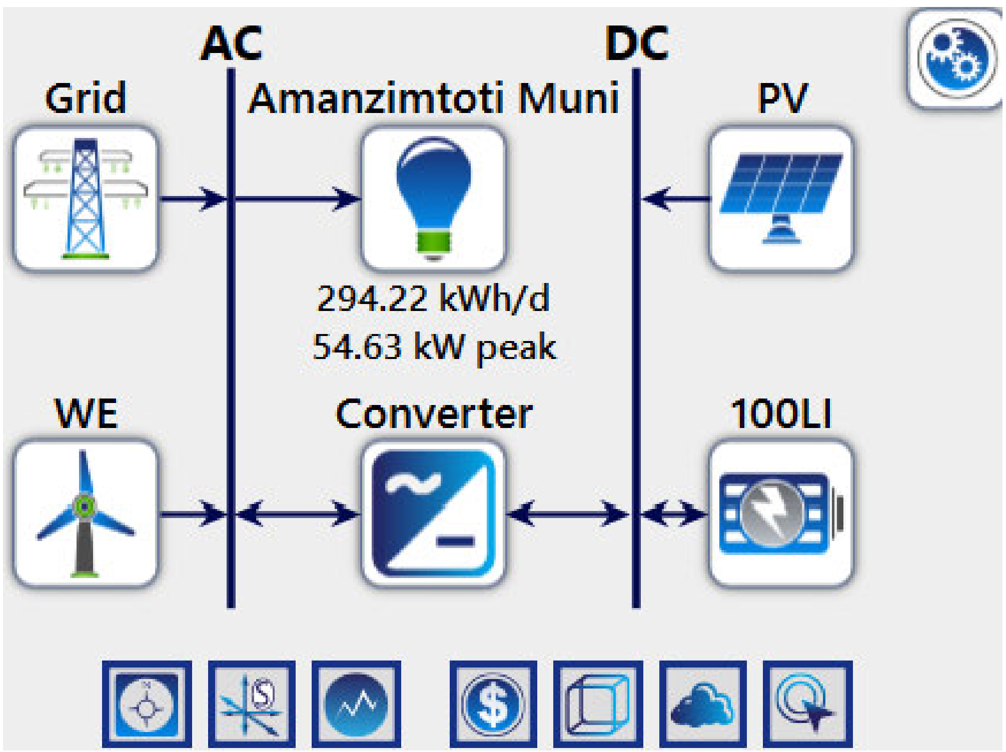

Figure 3.2 Network Configuration under Investigation

3.3.3 Solar PV Modeling

The power generated by the photovoltaic panels under specified conditions was calculated with HOMER Pro. using equation (3.1). While power output from PV panels varies throughout the day due to changing solar irradiance, the efficiency of PV modules is largely constant except for slight variations caused by temperature, dust, and aging effects. Higher temperatures typically decrease module efficiency by about 0.3–0.5% per °C above STC. [88]. Furthermore, variations in solar radiation are considered when calculating efficiency. The efficiency of the photovoltaic system varies throughout the day, depending on solar radiation and weather patterns. The primary characteristics of a PV panel in any hybrid system include efficiency based on output power generation, cost, secure layout, maintaining power production efficiency over long periods of time (the selected project lifespan), and the ability to generate power in low solar radiation scenarios.

The power output of a photovoltaic (PV) system can be expressed as (3.1) [89]:

$$P_{PV} = Y_{PV} f_{PV} \left(\frac{\bar{G}_T}{\bar{G}_{(T,STC)}} \right) \tag{3.1}$$

Where:

- P_{PV} is the power output of the PV system.
- Y_{PV} is the yield of the PV system.
- f_{PV} is the performance factor of the PV system.
- $\bar{G}_T$ is the actual solar irradiance.
- $\bar{G}_{(T,STC)}$ is the standard test condition (STC) solar irradiance.

3.3.4 Battery Modeling

Energy storage plays a crucial role in hybrid systems, significantly enhancing their reliability. In instances where energy is limited, the battery serves to capture surplus electricity generated by alternative power sources for later use. The process involves converting an alternating current (AC) to direct current (DC) for storage in the battery, which is subsequently converted back to AC to fulfill the demands of electricity. The relationship between overall system efficiency and battery round-trip efficiency is expressed in Equation (3.2). Round-trip efficiency is defined as the ratio of useful energy output to useful energy input. For this analysis, a standard lithium-ion battery with a capacity of 100 kWh and a voltage of 600 V is chosen as a representative example, reflecting typical configurations used in commercial-scale systems. This battery, which incurs a capital and replacement cost of ZAR 70,000.00, has a lifespan of 10 years and a specified throughput capacity. Sensitivity analysis was performed assuming a 10-year and 15-year lifetime, revealing that economic feasibility slightly decreases at 10 years but does not affect optimal system sizing

In this initial model, the battery is not optimized for load-specific requirements but is utilized to assess the impact of round-trip efficiency on the overall performance of the system. The energy stored in the battery is released during grid outages to fulfill load requirements. The round-trip efficiency of both the system and the battery is articulated in Equation [90] [91].

$$\eta_{\text{overall}} = \eta_{\text{inv}} \cdot \eta_{\text{rect}} \cdot \eta_{\text{rt}} \tag{3.2}$$

Where:

- η_{overall} is the overall efficiency of the system.

- η_{inv} is the efficiency of the inverter.
- η_{rect} is the efficiency of the rectifier.
- η_{rt} is the efficiency of the other components (e.g., transformers, etc.).

3.3.5 Wind Energy

During the past 10 years, wind energy has experienced remarkable growth, establishing itself as the second largest renewable energy source for electricity generation globally. Wind turbines (WTs) play a crucial role in harnessing this energy, using a specialized wind energy conversion system to transform the wind into electrical power. The efficiency of these turbines in the generation of electricity is influenced by several factors, including hub height, component costs, service life, and cut-in wind speed. Consequently, the power output of different wind generators varies considerably and depends on the wind speed at the hub height. HOMER Pro calculates the wind speed at the hub height of the wind turbine at each time interval, employing the logarithmic law as outlined in (3.3), [92]. The amount of power produced by a wind turbine is fundamentally linked to the characteristics of wind speed at the installation site and the specifications of the turbine. Figure 3.3 shows the power production from integrated renewable energy sources with the grid. The scaled average wind speed of the selected area is 6.14 m/s with a hub height of 17 m and expected lifespan of 10 years [92].

$$U_{\text{hub}} = U_{\text{anem}} \cdot \frac{\ln\left(\frac{Z_{\text{hub}}}{Z_0}\right)}{\ln\left(\frac{Z_{\text{anem}}}{Z_0}\right)} \tag{3.3}$$

Where:

- U_{hub} is the wind speed at hub height.
- U_{anem} is the wind speed measured by the anemometer.
- Z_{hub} is the height of the hub.
- Z_{anem} is the height of the anemometer.
- Z_0 is the roughness length.

The power output of the wind turbine can be expressed as (3.4) [93]:

$$P_{WT} = \begin{cases} P_r^{WT} \cdot \frac{(V-V_{\text{in}})^3}{(V_r-V_{\text{in}})^3} & \text{for } V_{\text{cin}} \leq V \leq V_r \\ P_r^{WT} & \text{for } V_r \leq V \leq V_{\text{cin}} \\ 0 & \text{for } V \geq V_{\text{out}} \leq V_{\text{cin}} \end{cases} \tag{3.4}$$

Where:

- P_{WT} is the power output of the wind turbine.
- P_r^{WT} is the rated power of the wind turbine.
- V is the wind speed.
- V_{in} is the cut-in wind speed.
- V_r is the rated wind speed.
- V_{cin} is the cut-out wind speed.
- V_{out} is the cut-out wind speed.

3.3.6 Grid Modeling

The grid serves as a crucial element of the electrical power system, acting as the primary source of electricity for a designated region. In instances where renewable energy sources are unable to generate electricity during the night, consumers have the option of purchasing electricity from the grid. In contrast, during daylight hours, when renewable energy sources produce surplus electricity, this excess can be sold back to the grid, with financial transactions managed through net metering. These processes are technically termed sell-back rates, demand rates, and grid pricing. For example, when solar photovoltaic systems produce more electricity than is consumed during the day, the surplus can be returned to the grid.

The South African utility used for this study has a power price rate of R2.19 per kilowatt hour (kWh). In situations where renewable energy generation is insufficient and battery storage is depleted, electricity must be acquired from the grid, provided that it is accessible.The net electrical energy exchanged with the grid is calculated as the difference between the energy purchased and the energy sold to the grid, as illustrated in the accompanying equation shown in equation (3.5) [94]. Equation (3.6) shows the generalized energy calculation.

The energy charge is defined as:

$$\text{Energy Charge} = (\text{Energy Purchased from Grid} \times \text{Grid Price}) - (\text{Energy Sold to Grid} \times \text{Sellback Price}) \tag{3.5}$$

The generalized energy calculation:

$$GEC = \sum_{p}^{r} \sum_{q}^{12} \begin{cases} E_{j.\mathrm{net}(G),p,q}^{g(ecG,p)} & \text{if } E_{j.\mathrm{net}(G),p,q} \geq 0 \\ E_{j.\mathrm{net}(G),p,q}^{g(ecS,p)} & \text{if } E_{j.\mathrm{net}(G),p,q} < 0 \end{cases} \tag{3.6}$$

Where:

- GEC is the generalized energy calculation.
- $E_{j.\mathrm{net}(G),p,q}$ is the net energy for the given parameters.
- $g(ecG,p)$ and $g(ecS,p)$ are functions that depend on the energy conditions.

The Case Study Context

Durban South Coast was selected due to its renewable energy potential and challenges with grid stability. This area has moderate solar potential and viable wind resources, making it an ideal candidate for a hybrid system demonstration. The demand profile and resource data are summarized in Table 3.1

Table 3.1 Summary of Case Study Parameters for Durban South Coast

Parameter	**Value**
Average Daily Demand	291.06 kWh/day
Initial PV Size (HOMER input)	60 kW
Wind System Size	20 kW (2 $\times$ 10 kW turbines)
Battery Storage Capacity	500 kWh
Average Solar Irradiation	5.85 kWh/m^2/day
Average Wind Speed	6.14 m/s
Grid Tariff	R2.19/kWh (purchase), R0.85/kWh (sellback)

3.4 Results

The grid provides a sellback rate of R0.85 per kWh, which is considerably lower than the purchase price of ZAR 2.19 per kWh. This disparity in pricing underscores a significant inefficiency: the financial return from selling surplus energy to the grid is less advantageous than the cost of purchasing energy from it. As a result, consumers are discouraged from returning excess energy to the grid due to inadequate compensation. This situation also indicates an issue of surplus energy, where excess energy is sold back inefficiently or wasted. Figure **??** shows the excess energy produced from renewable energy sources.

Table 3.2 shows the grid purchases and the grid sales, which underscores several significant concerns:

- **Dependence on the Grid and Financial Losses:** The cost of grid electricity stands at 2.19 ZAR/kWh, while surplus energy is sold at a mere 0.85 ZAR/kWh. This disparity in pricing creates a buy-sell-low scenario that restricts financial gains, rendering the sale of excess

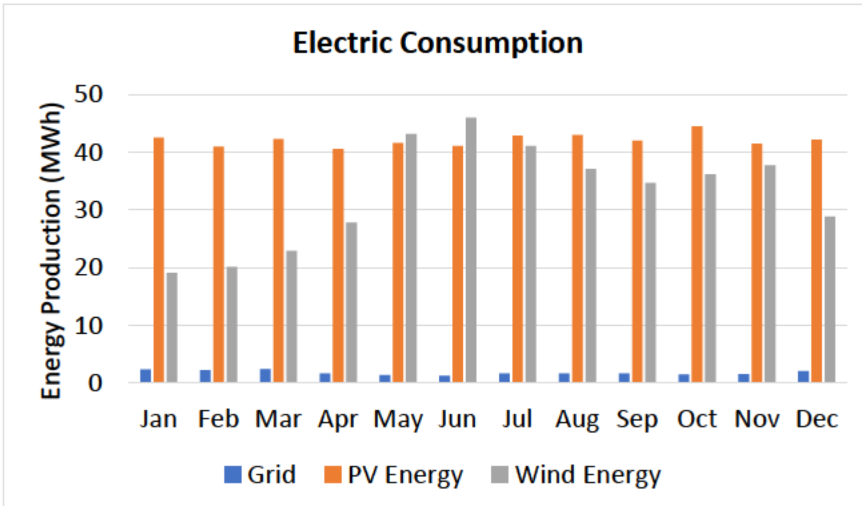

Figure 3.3 Monthly Electricity production from Durban South Region

energy to the grid economically unfeasible.

- **Surplus Energy and Storage Deficiencies:** A notable 11% (16,000 kWh/year) of the energy produced is wasted due to inadequate storage solutions and inefficient distribution. Although the system is overcapacity, the lack of sufficient battery storage results in a significant portion of the surplus energy going unused.

- **Effects on LCOE and ROI:** Energy underuse contributes to an increased Levelized Cost of Energy (LCOE)=0.9889 and diminishes the return on investment (ROI), as surplus energy is not effectively capitalized.

- **Continued High Grid Purchases:** In spite of substantial renewable energy generation, the system still relies on grid purchases that total 12,539 kWh (9%). This dependency undermines the objectives of energy independence, probably due to inadequate storage or peak demand exceeding available supply.

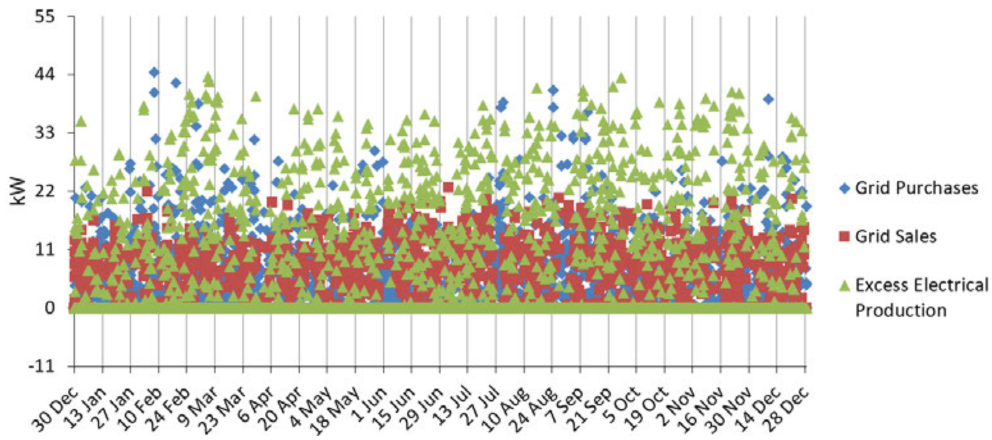

Figure 3.4 Excess Energy Production from Renewables, Grid Sales, and Purchases (kW) Over Time (Monthly.

Table 3.2 Monthly Energy Data

Month	Energy Purchased (kWh)	Energy Sold (kWh)	Net Energy Purchased (kWh)	Energy Charge	Total
Jan	1,014	695	319	R2,187	R2,187
Feb	1,283	712	572	R2,774	R2,774
Mar	1,736	695	1,041	R3,767	R3,767
Apr	845	627	218	R1,820	R1,820
May	751	940	-190	R1,597	R1,597
Jun	711	1,074	-363	R1,503	R1,503
Jul	821	1,181	-360	R1,739	R1,739
Aug	1,335	834	501	R2,883	R2,883
Sept	1,208	946	262	R2,598	R2,598
Oct	704	954	-250	R1,494	R1,494
Nov	963	804	159	R2,069	R2,069
Dec	1,168	734	434	R2,522	R2,522
Annual	**12,539**	**10,196**	**2,344**	**R26,951**	**R26,951**

3.5 Conclusion

The analysis of optimized energy data from the HOMER Pro software provides several important insights on the efficiency and sustainability of the existing microgrid configuration.

- **Energy Dependency:**

  The total annual net energy purchased amounts to 2,344 kWh, even though the system generates a substantial surplus of 10,196 kWh, which is sold back to the grid. Specifically, in months like May, June, July, and October, net purchases are negative, signifying that the microgrid often generates more energy than it utilizes. This phenomenon is mainly attributed to the seasonal fluctuations in renewable energy production, where increased solar or wind output during these months surpasses the load demand, resulting in excess energy being produced. However, during other months (January, February, and March), there is a significant dependence on the electricity of the grid, exposing the system to fluctuating prices on the grid (R2.19 / kWh) and supply uncertainties.

- **Financial Implications:**

  The data indicates a significant financial inefficiency. Although the system generates a considerable surplus, the low sell-back rate (R0.85/kWh) leads to under-utilization of energy resources. The annual energy expenditure of R26,951 illustrates the disparity between high costs for grid purchases and there latively lower revenue generated from energy sellback 3.2, resulting in a financial strain that could be alleviated by improving local energy sharing among consumers at a more advantageous rate. Adopting an optimized Game Theory model could diminish reliance on the grid and enhance consumer savings by facilitating Nash equilibrium for energy sharing.

- **Environmental Impact:**

  Despite the generation of excess renewable energy, the ongoing net dependence of 2,344 kWh on the grid still links the community to fossil fuel-based electricity, sourced primarily from coal in South Africa. By improving the redistribution of surplus energy through long-short-term memory (XGBoost) forecasting to anticipate demand and genetic algorithm-particle swarm optimization (GA-PSO) to size improvements in generation and storage capacity, the microgrid can decrease carbon-intensive grid imports. This approach would promote environmental sustainability and contribute to reductions in carbon emissions.

The GA-PSO method was employed to optimize system sizing, facilitating cost-effective generation and storage while maintaining a balanced energy output for consumers.

In order to achieve the primary objective of this research, a range of advanced techniques will be employed in the subsequent sections. The application of XGBoost-based load forecasting facilitates precise predictions of energy demand, thereby aiding in proactive energy management and ensuring that energy generation and storage are in accordance with user needs. Additionally, Linear Programming (LP) serves as a foundational model, demonstrating a traditional energy distribution system reliant on the grid, while Game Theory is incorporated to develop a dynamic and adaptive energy-sharing framework that encourages fair and economical redistribution of excess energy among households.

Chapter 4

Energy Consumption Forecast by XGBoost-XGBoost Hybrid Model

4.1 Introduction

The forecast of load is an essential aspect of the management and operation of power systems, serving as a cornerstone for effective utility planning and decision making. By accurately predicting future electricity demand across various time frames ranging from short-term (minutes to hours) to long-term (months to years), utility companies can strategically align their resources and infrastructure to meet consumer needs. This predictive capability allows informed decisions regarding infrastructure development, such as the construction of new power plants or the upgrading of existing facilities, as well as the optimization of energy generation and distribution strategies. The significance of accurate load forecasting cannot be overstated. It is vital to ensure reliable electricity supply, which is crucial for both residential and commercial consumers. In addition, precise forecasting helps reduce operational costs by minimizing the risk of overproduction or underproduction of electricity, which can lead to financial losses. In addition, effective load forecasting improves the overall efficiency of power systems by enabling utilities to better manage their resources, reduce waste, and improve service reliability.

This section aims to provide a comprehensive overview of load forecasting, including its various types such as short-, medium-, and long-term forecasting, and the primary methods and techniques employed in the field, including statistical methods, machine learning algorithms, and hybrid approaches. The process of forecasting the load is fundamental to effective management and operation of power systems. It involves estimating future electricity demand over different time intervals, which is critical for utility providers to make strategic decisions related to infrastructure expansion, energy production, and distribution logistics. The ability to predict load accurately allows utilities to allocate resources efficiently, ensuring that they can meet consumer demand without incurring unnecessary costs. Reliable load forecasting is essential to maintain stable electricity supply, as it helps prevent outages and ensures that the grid runs smoothly. In addition, it plays an important role in lowering operational costs by optimizing generation schedules and reducing the need for costly peak power plants that are only used during periods of high demand. By improving the overall performance of power systems, load forecasting contributes to a more sustainable and resilient energy landscape.

In this section, a comprehensive examination of various machine learning techniques for forecasting is undertaken. The models applied in this analysis include ARIMA, Random Forest, XGBoost, and XGBoost, all of which were specifically tailored to make predictions related to energy consumption. The assessment of these models was carried out using the Mean Absolute Error (MAE) as the evaluation criterion, which serves as a standard measure for quantifying the accuracy of the predictions. The findings from this assessment revealed that XGBoost consistently outperformed all other models in terms of predictive accuracy. As a direct consequence of these results, XGBoost was ultimately chosen as the most suitable model for the load forecasting task, given its ability to capture complex patterns in the data and provide reliable forecasts.

4.2 Data Analysis

The dataset pertaining to the Durban Southern Coast encompasses 300 households. This data set has been classified into ten clusters based on households exhibiting the most comparable load demand patterns [95].

The energy consumption for each household is recorded in kilowatt hours (kWh) at 30-minute intervals, offering an in-depth analysis of the trends in energy use. Data collection for this data set was carried out over a specific time frame, which spanned from midnight on July 1, 2012, to August 30, 2013. This period allows for a comprehensive examination of the trends and variations in energy consumption among selected households. The focus on such a timeframe is crucial for understanding seasonal and temporal influences on energy usage.

Retained households serve as a representative sample for categorization analysis, enabling researchers to draw meaningful conclusions about energy consumption behaviors. By focusing on this subset, the analysis can yield insights that may be applicable to broader contexts within the region. To ensure data integrity, all entries were verified to contain valid timestamps, thus preparing the DataFrame for subsequent analytical or visual tasks. The removal of rows containing NaT values was carried out to enhance the cleanliness and reliability of the dataset, marking a crucial phase in the data preprocessing process.

Statistical Analysis

Statistical analysis is used in energy consumption data to gain a deeper understanding of the data and to make informed decisions. Statistical analysis plays a crucial role in the examination of energy consumption data, enabling a deeper understanding of information and facilitating well-informed decision making. By applying (ARIMA, Random Forest, LSTM, XGBoost), stakeholders can gain insight into energy usage patterns, identify inefficiencies, and develop optimization strategies. The applications of statistical analysis in this context can significantly improve the understanding of the patterns and trends of energy usage, ultimately leading to more sustainable energy practices.

One key application is descriptive analysis, which employs statistical measures to effectively summarize the data. This includes calculating essential metrics such as the mean, median and standard deviation of energy consumption, providing a clear overview of the central tendencies and variability of the data. For example, the mean offers a simple average of energy usage, while the median provides insight into the midpoint of the data, helping to mitigate the influence of outliers. The standard deviation, on the other hand, quantifies the extent of variation or dispersion in energy consumption, allowing analysts to understand how consistent or erratic the use of energy over time. Such descriptive statistics serve as foundational tools for further analysis and interpretation, enabling stakeholders to identify anomalies and establish benchmarks for energy performance.

Another important application is time-series analysis, which focuses on evaluating energy consumption data across different time intervals to uncover trends and recurring patterns. By analyzing data collected over days, months or years, analysts can identify seasonal variations in energy usage, such as increased consumption during summer months due to air conditioning or winter months due to heating needs. Time-series analysis also allows for the detection of long-term trends, such as a gradual increase in energy efficiency due to technological advancements or shifts in consumer behavior. This understanding is vital for energy providers and policymakers, as it informs demand forecasting and resource allocation. Statistical metrics can be calculated as in (4.1) and (4.2), where the total number of values is calculated in (4.3) and (4.4) :

$$\text{Count} = \text{Number of non-null values in the column} \tag{4.1}$$

$$[\text{Mean} = \frac{\sum_{i=1}^{n} x_i}{n}] \tag{4.2}$$

Where

$$(x_i)$$

is each value in the column and (n) is the total number of values.

$$[Q_2 = \text{Median} = \text{Value at the 50th percentile}] \tag{4.3}$$

$$[\text{Max} = \text{Largest value in the column}] \tag{4.4}$$

Additionally, regression analysis is utilized to develop statistical models that can forecast energy consumption based on various influencing factors, including weather conditions, time of day, and historical usage trends. For example, multiple regression models can incorporate variables such as temperature, humidity, and occupancy rates to predict energy demand with greater precision. Organizations can make data-driven decisions regarding energy supply, pricing strategies, and infrastructure investments by understanding the relationships between these variables and energy consumption. These analytical methods collectively contribute to a more comprehensive understanding of the dynamics of energy consumption, enabling stakeholders to implement targeted interventions that promote energy efficiency and sustainability.

In addition, advanced statistical techniques, such as cluster analysis, can be employed to segment energy consumers into distinct groups based on their usage patterns. This segmentation allows for tailored energy management strategies, such as targeted marketing campaigns for energy-saving programs or customized pricing plans that incentivize lower consumption during peak demand periods.

Although the dataset comprised 300 households, they were clustered into ten representative groups based on load demand similarity.Table 4.1 summarizes these ten clusters (labeled Cons 1–10) rather than all 300 households including statistical measures with essential metrics such as; the mean, median, and standard deviation for each consumer where: count is the number of non-null values in each column, mean is the value of each column, std is the standard deviation of each column, min is the minimum value in each column, 25%, 50% and 75% represents the 25th percentile (Q1) of each column, the 50th percentile (Q2 or median) of each column and the 75th percentile (Q3) of each column respectively.

Table 4.1 Energy Consumption Statistical Analysis

Statistical Metric	Cons 1	Cons 2	Cons 3	Cons 4	Cons 5	Cons 6	Cons 7	Cons 8	Cons 9	Cons 10
Count	17473	17473	17473	17473	17473	17473	17473	17473	17473	17473
Mean	0.439490	0.533408	0.253917	0.428771	0.513784	0.701526	0.281534	0.698481	0.24314	0.280547
std	0	0	0.253917	0.428771	0.513784	0.691453	0.416834	0.723912	0.49409	0.492572
min	0.121	0.102	0	0	0	0	0	0	0	0
25%	0.121	0.102	0.053	0.105	0.095	0.295	0.094	0.157	0.012	0.05
50%	0.656	0.297	0.082	0.262	0.29	0.458	0.14	0.461	0.05	0.113
75%	0.656	0.297	0.082	0.262	0.29	0.849	0.235	0.932	0.158	0.218
max	21.415	9.47	4.564	8.69	33.42	30.453	7.782	8.728	7.391	8.823

Load Factors

In the context of power systems, particularly with the integration of renewable energy sources, the load factor plays a crucial role in calculations. Understanding the load factor is crucial, as it provides information on the electricity consumption patterns in residential settings. This metric is essential for several reasons, primarily because of its ability to assess the efficiency of the system. By evaluating this metric, stakeholders including utility companies, policymakers, and consumers can identify areas for improvement and optimize energy usage. For instance, a low load factor may indicate that households are using electricity in a highly variable manner, with significant peaks and troughs in demand. This variability can lead to inefficiencies in energy production and distribution, as utilities must maintain capacity to meet peak demands, often resulting in higher operational costs.

Moreover, an accurate calculation of the load factor can lead to better resource management and planning. Utilities can use load factor data to forecast demand more accurately, allowing them to adjust their energy generation strategies accordingly. This proactive approach can help integrate renewable energy sources, as understanding consumption patterns can facilitate the alignment of energy supply with demand. Furthermore, by identifying trends in electricity usage, utilities can implement demand-side management programs that encourage consumers to shift their usage to off-peak times, thereby flattening the load curve and enhancing overall system efficiency.This, in turn, can contribute to more sustainable energy practices and potentially lower costs for consumers. When energy consumption is optimized, it reduces the need for additional power plants and infrastructure, which can be costly and environmentally detrimental. Furthermore, consumers may benefit from lower electricity bills as utilities pass on savings achieved through improved efficiency and reduced peak demand.In summary, the load factor is not merely a technical metric; it is a critical tool to improve the efficiency of electricity consumption in households.

By understanding and leveraging this information, stakeholders can foster a more sustainable energy future, characterized by lower costs, improved resource management, and lower environmental impact. The load factor is defined as the ratio of the actual output of a power plant to its potential output if it were to operate at full capacity for a specific period. In this study, following the assessment of Total Energy Consumption, the Average Load and Maximum Load from the households were calculated to derive the load factor. The analysis involved a systematic approach to quantify the patterns of energy use within households, which facilitated the calculation of both the average load and the maximum load. These metrics are essential to understand the overall energy demand and efficiency of the households in question.Ultimately, the load factor was determined utilizing the calculated values of the average load and the maximum load, providing information on the behavior of energy consumption and potential areas for improvement in energy management. In equations (4.5) and (4.6) the average load and the load factor are calculated respectively.

$$\text{Average Load (kW)} = \frac{\text{Total Energy Consumed (kWh)}}{\text{Total Time (hours)}} \tag{4.5}$$

$$\text{Load Factor} = \frac{\text{Average Load}}{\text{Peak Load}} \tag{4.6}$$

Where Average load is the total energy consumed during the period divided by the total time (in hours), and Peak load is the maximum load recorded during that period.

Diversity Factor

In this study, the calculation of the diversity factor was a critical component to understanding the overall patterns of energy consumption within the residential area under investigation. To begin with, we focus on determining the maximum individual demands for each household. This process involved analyzing the peak load for each residence, which is defined as the highest level of electricity consumption recorded over a specific period. Identifying these peak loads allowed us to gain insight into the behavior of energy usage of each household, taking into account factors such as the number of occupants, the types of appliances used, and the time of day when energy consumption was highest. Furthermore, the analysis was extended to assess the maximum demand of the system as a whole. This metric reflects the highest load recorded at any given time in all households within the study area.

By evaluating the maximum demand for the system, the cumulative impact of individual household demands on the overall energy infrastructure is easily understood. This assessment is crucial for utility providers and energy planners, as it helps to determine the capacity requirements of the electrical grid and ensure that there is sufficient supply to meet peak demands. The relationship between individual maximum demands and the overall demand of the system is encapsulated in the diversity factor, which quantifies the extent to which the peak demands of different households coincide. A higher diversity factor indicates that not all households are consuming their maximum load simultaneously, which can lead to more efficient energy distribution and reduced infrastructure costs. By analyzing these factors in detail, this study aims to provide valuable information on energy consumption patterns, inform future energy planning, and contribute to the development of more sustainable energy systems. The diversity factor formula used in the study is provided in equation (4.7) as:

$$\text{Diversity Factor} = \frac{\text{Sum of Individual Peak Loads}}{\text{Maximum System Load}} \tag{4.7}$$

The diversity factor was then applied to estimate the actual peak demand that the system needed to accommodate. This factor helps to estimate a more realistic peak demand for the entire system. The total and estimated peak demand considering the diversity factor was then calculated as shown in (4.8).

$$P_{\text{total}} = \sum_{i=1}^{n} P_i \tag{4.8}$$

where:

- P_{total} is the total peak demand.
- P_i is the peak demand of the i-th consumer.
- n is the total number of consumers.

4.3 Methodology

The initial phase of the methodology involved the systematic collection and preparation of data, which served as the basis for subsequent analyzes. This phase included gathering relevant datasets, cleaning the data to remove inconsistencies or errors, and organizing it in a manner conducive to analysis. Following this preparatory work, the second phase included a thorough data exploration. This exploration aimed to uncover patterns, trends, and anomalies within the dataset, providing valuable insights that would inform the modeling process. The third phase focused on reprocessing the data specifically for machine learning applications. This step involved transforming the data into a suitable format, normalizing the values, handling missing data, and creating features that would enhance the predictive power of the models.

In the fourth phase, a variety of machine learning algorithms were used to forecast short-term electrical loads. The algorithms used included long-short-term memory (XGBoost) networks, auto-regressive Integrated Moving Average (ARIMA) models, Random Forest, and XGBoost. Each of these algorithms has unique strengths and characteristics that make them suitable for different aspects of the forecasting challenge. The performance of these algorithms was rigorously evaluated using Mean Absolute Error (MAE) metrics, which provided a quantitative measure of the accuracy of the prediction. This evaluation process was critical to identifying the most effective approach for the specific forecasting task at hand. Ultimately, the best model for electric load forecasting was selected on the basis of comprehensive optimization and tuning processes applied to the various models. These processes ensured that the chosen model was not only accurate, but also robust and reliable for practical applications. A visual summary of the methodology is presented in figure 4.1, illustrating the sequential steps taken throughout the chapter.

4.3.1 Data Collection and Description

This section provides a comprehensive overview of the methodology used in this research work, detailing each step taken to ensure a robust and effective analysis of data on electricity consumption. The process began with the meticulous collection of metered data from the local electricity municipality, which serves as the primary source of information for this study. This data was subsequently organized and stored in a CSV file format (Comma-Separated Values), facilitating easy access and manipulation for further analysis.

Once the data was securely stored, the next phase involved Data Exploration. This critical step aimed to uncover insights into the inherent characteristics of the data, including its distribution, trends, and any potential anomalies. (ARIMA, Random Forest, LSTM, XGBoost) and visualization tools were employed during this phase to assess the quality of the data, identify missing values, and understand the relationships between different variables. This exploratory analysis provided a solid foundation for the subsequent steps in the methodology.

Following the exploration, the focus shifted to Data Processing. This phase was essential for preparing the data for analysis and involved several key activities. Normalization techniques were applied to ensure that the data was on a consistent scale, which is particularly important for machine learning algorithms that are sensitive to the magnitude of input features. In addition, feature selection was performed taking into account various factors such as the specific day, day of the week, whether it was a weekday or weekend, and the month of the year. This selection process aimed to identify the most relevant features that would contribute to the predictive power of the model.

After processing and preparing the data, the next step was training the machine learning models. Various algorithms were used to develop a predictive model capable of forecasting future electricity load consumption, utilizing the historical data collected. This phase included selecting appropriate machine learning techniques, such as regression analysis, decision trees, or neural networks, depending on the nature of the data and the specific forecasting requirements.

Once the models were trained, a critical evaluation of their accuracy was performed. This evaluation process involved comparing the model's predictions against actual historical data to determine its performance metrics, such as Mean Absolute Error (MAE) or Root Mean Squared Error (RMSE). If the model met the predetermined accuracy thresholds, it was deemed suitable for application in forecasting future load consumption. Conversely, if the model performance was found to be lacking, a retraining process was initiated, which involved revisiting the data processing and feature selection steps to enhance the model's predictive capabilities.

Finally, after achieving a satisfactory level of precision, the trained model was used to forecast load consumption. This culminated in the generation of forecasts that provided valuable insights.

4.4 Energy Consumption Forecasting

In this section, a comparative analysis was performed using these four methodologies - ARIMA, XGBoost, Random Forest, and XGBoost - to evaluate their effectiveness in forecasting energy consumption. The results of the analysis indicated that the XGBoost demonstrated a superior performance among the models tested. This can be attributed to its capacity to learn intricate patterns and dependencies in the time-series data, which is particularly relevant in the context of energy consumption that often exhibits nonlinear trends and seasonal variations.

Overall, the findings underscore the importance of selecting appropriate forecasting methodologies based on the specific characteristics of the data and the forecasting objectives. The integration of advanced techniques such as XGBoost into energy consumption forecasting not only enhances predictive accuracy but also provides valuable insights into the underlying dynamics of energy consumption patterns.

Load Forecasting models

Given that stationarity is a crucial requirement for the ARIMA model, the first actions taken included differencing and implementing power transformations to eradicate trends. Following this, the parameters (p, d, q) were established to facilitate the development of the ARIMA model. In this framework, p represents the count of preceding periods in the time series that are taken into account, d represents the differencing order essential to achieve stationarity in the time series, and q relates to the number of lags associated with the error term. During our initial exploration, we examined various combinations of these parameters, relying on insights from the autocorrelation function (ACF) and the partial autocorrelation function (PACF). Eventually, a different parameterization strategy was selected, based on the analysis of Akaike Information Criterion (AIC) values, to minimize error metrics. During advances in statistical-based methodologies, the machine learning technique known as random forest was also explored, developed, and used for time series forecasting. This method can be characterized as an ensemble approach that consolidates the predictions from a multitude of distinct decision trees. Random forest requires minimal hyperparameter tuning; however, its implementation necessitates the construction of a series of decision trees derived from a limited set of features. During data analysis, it became evident that there were insufficient features to effectively utilize the random forest algorithm, which requires a larger dataset. Consequently, the variables in the initial column, which contained the timestamps of the observations, were divided into four distinct variables (day, month, weekday, and day of the week).

Additional features, such as total electricity consumption for the subsequent day, were also incorporated. Following this, one of the widely recommended practices in machine learning, one-hot encoding, was applied to transform nonnumerical values into numerical representations, resulting in a balanced matrix. The final stage of data preparation involved the establishment of training and testing datasets by determining a fixed split point. Similarly, the XGBoost model was developed. This machine learning algorithm, similar to the random forest, is based on decision trees but employs a gradient boosting technique for regression and classification tasks. Initially, training and testing datasets were defined and the same data transformation processes used in random forests to generate additional characteristics were adopted. The model was subsequently created and trained for deployment in load forecasting. The last category pertains to deep learning methods, and XGBoost is selected among various options for forecasting. The fitting process involved three transformations: rendering the time series data stationary, converting it into a supervised learning format, and applying data scaling. Finally, the data were divided into 60% for training and 20% for validation and 20% for testing, and lastly, the most suitable parameterization for the model implementation was chosen, incorporating 50 epochs, one dense unit and a batch size of 32.

4.4.1 Data Understanding and Exploration

1. **Data Understanding and Exploration:** Clearly define the problem and objectives (e.g., energy consumption prediction). Perform exploratory data analysis (EDA) to understand the dataset, identify missing values, and detect outliers. Visualize trends, seasonality, and correlations in the data.

2. **Data Preprocessing:** Handle missing values and outliers systematically (e.g., imputation, removal). Encode categorical variables using appropriate techniques (e.g., one-hot encoding, label encoding). Normalize or standardize numerical features to ensure uniform scaling. Split the data set into training, validation and test sets with a clear rationale for the split ratio.

3. **Feature Engineering:** Create time-based features (e.g., hour, day, month) if applicable. Perform feature selection to retain only relevant features. Consider lag features or rolling

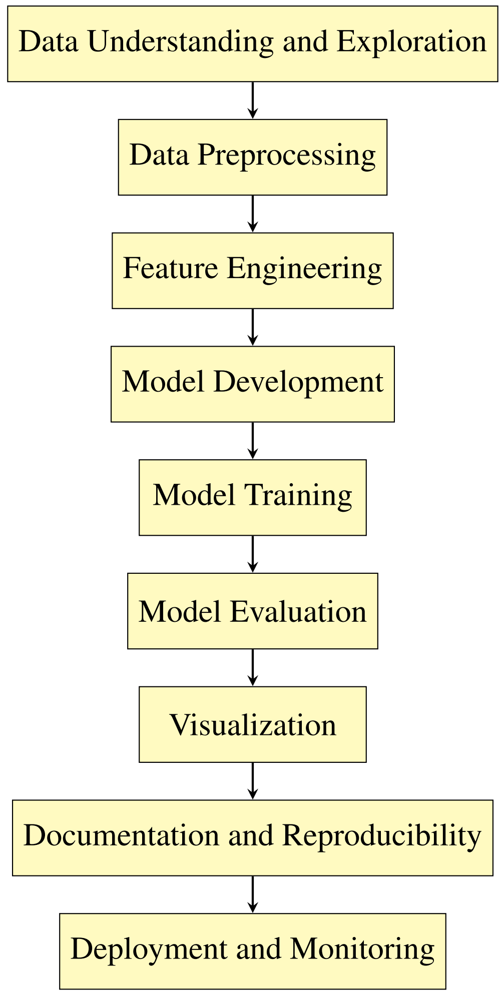

Figure 4.1 Framework for Prediction Modeling.

statistics for time series data.

4. **Model Development:** Choose the appropriate model architecture (e.g., XGBoost for time-series data). Define hyperparameters and use grid search or random search for optimization. Incorporate regularization techniques to prevent overfitting.

5. **Model Training:** Use early stopping to monitor validation loss and prevent overfitting. Train the model on the training set and validate on the validation set. Save the best model weights during training.

6. **Model Evaluation:** Evaluate the model on the test set using multiple metrics such as MAE, RMSE and R^2. Analyze residuals to identify patterns or biases in predictions.

7. **Visualization:** Plot actual vs. predicted values to visually assess model performance. Create additional plots (e.g., residuals histogram, time-series plots) for deeper insights.

8. **Documentation and Reproducibility:** Document all steps, including data preprocessing, feature engineering, and model training. Save the final model and preprocessing pipeline for future use. Ensure that the code is modular.

Comparison of Models and Performance Evaluation

Compared to the aforementioned models, a series of innovative methodologies were developed and implemented specifically for the purpose of load forecasting, utilizing the same time series data that were previously mentioned. These methodologies were designed to improve the accuracy and reliability of load predictions. To evaluate the effectiveness of these methods, a comprehensive assessment was performed, focusing on mean absolute error (MAE) as a primary performance indicator. This metric was chosen because of its ability to provide a clear and interpretable measure of forecast accuracy, allowing for a straightforward comparison between the different methodologies employed.

The Mean Absolute Error (MAE) for the long-short-term memory (XGBoost) model was the lowest among all models evaluated, indicating its superior performance. This metric, which measures the average magnitude of errors in a set of predictions, highlights the XGBoost's ability to provide more accurate forecasts compared to other models in the evaluation. The consistently lower MAE suggests that the XGBoost model is better equipped to capture the underlying patterns and trends in the energy data, making it a reliable choice for forecasting future energy demands.

Consequently, due to its exceptional performance and reliability, the XGBoost model was selected as the preferred model for energy forecasting. Equation (4.9) shows the error metric used for the evaluation of models:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{4.9}$$

Where:

- n is the total number of observations.
- y_i is the actual value for the i-th observation.
- $\hat{y}_i$ is the predicted value for the i-th observation.

4.5 Long-Short-Term-Memory Model

This section outlines the methodology for forecasting aggregated community loads using smart meter data in conjunction with long-short-term memory (XGBoost) models. The XGBoost model was specifically designed to predict energy consumption. The process involves importing a dataset, preprocessing it, splitting it into training and validation subsets, constructing and training the XGBoost model, evaluating its performance, and visualizing the results. Keras was utilized as the deep learning environment, with TensorFlow serving as the backend for Python. The XGBoost network layer in Keras was responsible for maintaining states across batches of data, where each batch consisted of a predetermined number of rows from the training dataset, thereby establishing the frequency of updates to the network's weights. In addition, pandas were used for data manipulation and analysis, while NumPy facilitated numerical operations and array manipulation.

Data Preprocessing and Exploration

In this phase, a comprehensive examination of the data types assigned to each column in the data set is performed. This step is crucial as it ensures that the data types are appropriate for the intended machine learning processes. For instance, numerical features should be of types such as integers or floats, while categorical features should be represented as strings or categorical types. Any discrepancies or misclassifications in data types can lead to errors during model training and evaluation, making this validation a critical step in the data preparation process.

1. Label Encoding

    - In the preprocessing phase, categorical variables such as Consumer and their categorical electrical appliances (load) of consumers were transformed into numerical formats using the 'LabelEncoder' from the 'sklearn.preprocessing module '. This transformation is crucial because machine learning algorithms typically require numerical input. By encoding these categorical variables, we ensure that the model can interpret and utilize them effectively during training and prediction.

2. Feature and Target Selection

    - The dataset was structured such that the features (denoted as X) include the columns for Consumer, load, and Energy Consumption (kWh). These features are essential as they provide the necessary information for the model to learn patterns and relationships. The target variable (denoted as y) is represented by the total_energy_consumption_kWh column, which is the outcome we aim to predict. This clear distinction between features and target is fundamental for supervised learning tasks.

3. Train-Validation-Test Split

- To evaluate the model's performance effectively, the dataset is divided into three distinct subsets:
  - **Training Data (60%)**: This portion ($X_{\text{train}}, y_{\text{train}}$) is used to train the model, allowing it to learn from the data.
  - **Validation Data (20%)**: This subset ($X_{\text{val}}, y_{\text{val}}$) is utilized to tune the model's hyperparameters and make decisions about model architecture without overfitting to the training data.
  - **Testing Data (20%)**: Finally, the test set ($X_{\text{test}}, y_{\text{test}}$) is reserved for evaluating the model's performance on unseen data, providing an unbiased assessment of its predictive capabilities.

4. Feature Scaling

   - To ensure that all features contribute equally to the model's learning process, feature scaling is applied. Specifically, ‘StandardScaler‘ is employed to standardize the features, transforming them to have a mean of 0 and a standard deviation of 1. This standardization is vital, especially when features are on different scales, as it prevents any single feature from dominating the learning process due to its larger range of values. By scaling the features, we enhance the model's convergence speed and overall performance.

5. Reshaping Input Data for XGBoost

   - After scaling, the input data is reshaped into a 3D format suitable for Long Short-Term Memory (XGBoost) networks. The required structure is organized as (samples, time steps, features).

XGBoost Model Development

1. Model Architecture

    - The architecture of the model is built around a single Long Short-Term Memory (XGBoost) layer, which is a type of recurrent neural network (RNN) well-suited for sequence prediction tasks. This XGBoost layer consists of 20 units, allowing it to capture temporal dependencies in the data effectively. The activation function used in this layer is the Rectified Linear Unit (ReLU), which helps in introducing non-linearity into the model and mitigating the vanishing gradient problem often encountered in deep networks.

2. Dense Output Layer

    - At the output stage, the model incorporates a single neuron. This neuron is specifically tasked with predicting energy consumption, which is a continuous variable. The choice of a single output neuron is appropriate for regression tasks, where the goal is to output a single value rather than a class label.

3. Compilation

    - For the compilation of the model, the Adam optimizer is employed. Adam is a popular optimization algorithm that combines the advantages of two other extensions of stochastic gradient descent, namely AdaGrad and RMSProp. It is particularly effective for training deep learning models. The loss function used is Mean Squared Error (MSE), which quantifies the average squared difference between the predicted values and the actual values. MSE is a standard choice for regression problems, as it provides a clear measure of prediction accuracy.

4. Training the Model

    - Early Stopping

        - To prevent overfitting during the training process, an early stopping mechanism is implemented. This technique monitors the validation loss, and if there is no improvement over 5 consecutive epochs, the training is halted. At this point, the model reverts to the weights that yielded the best validation performance, ensuring that the model retains its ability to generalize well to unseen data.

    - Epochs and Batch Size

        - The model is trained for a maximum of 50 epochs, which is the number of complete passes through the training dataset. The batch size is set to 32, meaning that the model updates its weights after processing 32 samples at a time. This batch size strikes a balance between computational efficiency and the stability of the gradient estimates.

5. Evaluation

    - The effectiveness of the model is evaluated using a separate test dataset. The primary metric for assessment is the test loss, calculated using Mean Squared Error (MSE). This metric provides insight into how well the model performs on unseen data, indicating its predictive accuracy. Additionally, predictions are generated for the validation dataset, allowing for a comparison between predicted and actual values.

4.6 Problems and Limitations of XGBoost Model

Long Short-Term Memory (XGBoost) networks have demonstrated significant effectiveness in identifying temporal dependencies within time series data. Nevertheless, independent XGBoost models frequently encounter several challenges, including:

- **Overfitting:** XGBoosts may memorize irrelevant noise, particularly in smaller datasets, which results in inadequate generalization.
- **Slow Convergence:** The training process for deep XGBoost networks can be both computationally demanding and time-consuming.
- **Limited Linear Trend Capture:** While XGBoosts are proficient at detecting nonlinear patterns, they may struggle to model linear trends accurately.

These challenges highlight the need for a more robust hybrid model that integrates the advantages of various architectures.

4.7 Hybrid XGBoost-XGBoost Model Overview

This section introduces a hybrid model that fuses XGBoost and XGBoost to increase forecast accuracy. XGBoost is proficient in extracting complex temporal relationships and encoding patterns found in time series data. Meanwhile, XGBoost enhances XGBoost's capabilities by capturing linear trends and addressing residual errors. This hybrid architecture aims to boost performance through the integration of deep learning and gradient boosting techniques.

4.7.1 XGBoost Model

The XGBoost model is designed to capture the time dependence in the data. The architecture consists of the following components:

- **XGBoost Units:** 64 XGBoost units are used to learn the sequential patterns in the data.
- **Dropout Layer:** A dropout layer with a rate of 0.2 is implemented to prevent overfitting during training.
- **Dense Layers:** The model includes dense layers for regression output, allowing for continuous value predictions.
- **Early Stopping:** Early stopping is used to avoid unnecessary training epochs, ensuring efficient training and preventing overfitting.

4.7.2 XGBoost Model

The XGBoost model is trained using XGBoost embeddings combined with original features. This integration allows the model to take advantage of both sequential and static features, enhancing prediction accuracy.

The hyperparameter for the XGBoost model are as follows:

- **n_estimators:** 100
- **learning_rate:** 0.05
- **max_depth:** 5

4.7.3 Evaluation of XGBoost-XGBoost Model

In evaluating the effectiveness of the hybrid model, three primary metrics were utilized:

- **Mean Absolute Error (MAE):** This metric quantifies the average size of the discrepancies between predicted and actual values, yielding an understandable error measurement in kilowatt-hours (kWh).
- **Root Mean Square Error (RMSE):** This metric places greater emphasis on larger errors compared to MAE, thereby providing a clearer understanding of the model's performance in the presence of significant deviations.
- **R^2 Score (Coefficient of Determination):** This score reflects the model's ability to account for variance in the target data, with values approaching 1 indicating highly accurate predictions.

4.8 Results

This section delves into the critical data preparation steps required for implementing a machine learning approach to load forecasting. It begins with a comprehensive overview of data preprocessing techniques, which are essential for ensuring the quality and reliability of the data used in modeling. Following this, the section explores the intricate patterns of household electricity consumption, highlighting how these patterns can vary based on factors such as time of day, seasonality, and individual household behaviors. Additionally, the analysis of load and diversity factors is discussed, providing insights into how different households contribute to overall electricity demand and how this can impact forecasting accuracy. The section culminates in a comparative analysis of various machine learning models, evaluating their performance in the context of load forecasting. Special emphasis is placed on the outcomes associated with the selected model, specifically the Long Short-Term Memory (XGBoost) network, which is known for its effectiveness in capturing temporal dependencies in time series data.

4.8.1 Electrical Household Behavior Analysis

The prepared dataframe includes a timestamp that denotes the date and time of the recorded energy consumption data, as well as a column indicating the number of households serving as consumers. These columns capture the energy consumption values (in kWh) for ten different consumers. Each column is linked to a specific consumer, and each row represents a specific time interval, occurring every 30 minutes, which contains the energy consumption values for each consumer at that designated time.

4.8.2 Comparison and performance Evaluation of Models

Forecasting was conducted using various models as follows: The consumer energy consumption data is organized within a dictionary (consumer data). This dataset encompasses features such as timestamp, Consumer, Power in watts, quantity, Backup time (in hours), and the target variable, Energy Consumption (in kWh). One-hot encoding was applied to transform the categorical Consumer column into numerical binary columns for each consumer, which are utilized as features. The feature matrix X comprises all pertinent columns, including Power in watts, quantity, Backup time (in hours), and the one-hot encoded consumer columns. The target variable y corresponds to the column of energy consumption (in kWh).The implementation of different models was carried out to identify the most effective. The models evaluated included Random Forest, XGBoost, ARIMA, and XGBoost (figure 4.2).

The results indicate the performance summary for different forecasting models, including the completion of training for the XGBoost model as follows: The Mean Absolute Error (MAE) for the Random Forest model is 12.89, indicating satisfactory performance; however, it exhibits a higher error rate in comparison to the XGBoost model. The XGBoost model has an MAE of 12.86, which slightly surpasses that of Random Forest, demonstrating effective handling of feature interactions. In contrast, the ARIMA model shows an MAE of 13.46, performing less effectively than the machine learning models, as it is primarily designed for time-series forecasting and may struggle with feature interactions. The XGBoost model, with an impressive MAE of 0.20, significantly outperforms all other models, highlighting its capability to learn intricate temporal patterns and dependencies within the data. Throughout the training process, the loss consistently decreased, indicating effective learning by the model. The final loss values, approximately 0.034, correspond with the low MAE observed during evaluation. The notably low XGBoost MAE of 0.20 implies that the model adeptly captures both temporal dependencies and non-linear patterns within the dataset, although the XGBoost MAE was initially over 10%, which has since been improved.

The XGBoost model exhibits a low Mean Absolute Error (MAE) of 0.20, suggesting superior performance compared to other models, although the error is still relatively high and subject to further hyperparameter tuning. This low MAE can be attributed to the model's sequential nature and its reliance on historical data to predict future outcomes. In contrast, while Random Forest and XGBoost show acceptable performance, they do not match the sequential modeling advantages provided by XGBoost.

In contrast, ARIMA faces challenges because it concentrates exclusively on the time series component and may overlook other independent variables. In figure 4.3, a comprehensive comparison of the forecast outcomes generated by four different models is presented. ARIMA, XGBoost, Random Forest, and XGBoost is presented. Each of these models employs different methodologies and algorithms to predict future values based on historical data, and their performance is evaluated based on various accuracy metrics. Furthermore, the performance of the XGBoost learning approach is particularly noteworthy when examining the evaluation of error metrics. This figure provides a detailed breakdown of the prediction errors associated with each model, highlighting the significant advantage of the XGBoost. The lower prediction errors associated with the XGBoost model not only underscore its effectiveness but also suggest that it may be a more reliable choice for forecasting tasks, especially in scenarios where accuracy is paramount. A lower MAE indicates that the model's predictions are closer to the actual observed values, thereby reflecting higher accuracy. In this case, the XGBoost's MAE of 0.2 suggests that it consistently produces forecasts that are more precise than those generated by the other models under consideration.

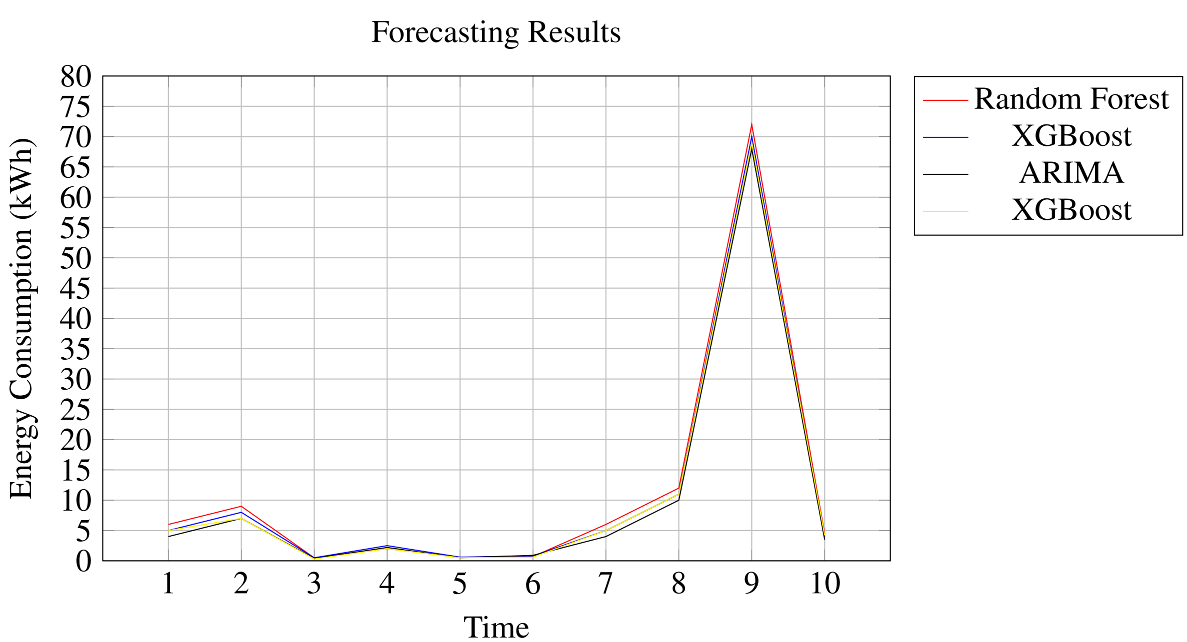

Figure 4.2 Forecasting results from different models

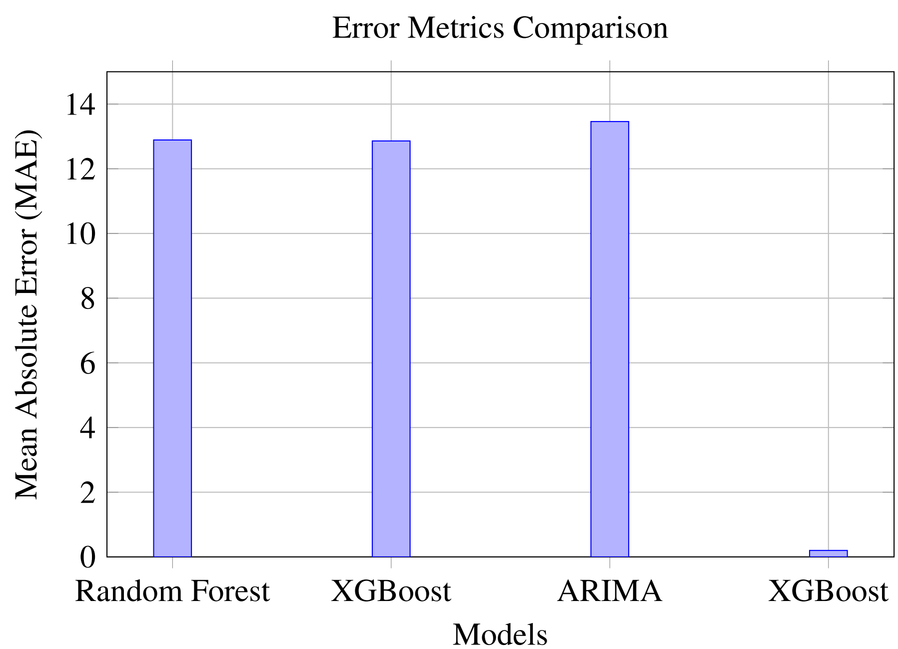

Figure 4.3 Comparison of Mean Absolute Error (MAE) for different models

4.8.3 Hybrid XGBoost-XGBoost Model

The hybrid XGBoost-XGBoost model aims to integrate the sequential pattern recognition abilities of XGBoost networks with the powerful regression features of XGBoost. XGBoost is adept at detecting time-related dependencies and nonlinear dynamics in energy consumption data. At the same time, XGBoost improves the accuracy of the final predictions by learning from the patterns identified by XGBoost, thus reducing errors.

Following the consideration of this model, enhancements were made to improve its predictive capabilities. The scatter plot presented in figure 4.4 illustrates a visual comparison between the actual energy consumption values (on the x-axis) and the predicted energy consumption values (on the y-axis). The dashed red line indicates the ideal scenario in which the predicted values perfectly align with the actual values y equal to x. The proximity of data points to this dashed line indicates the accuracy of the model's predictions. In contrast, points that deviate significantly from the red line reflect larger prediction errors.

The load forecast for both the short-term and medium-term periods is depicted in figure 4.5. This figure provides a visual representation of the anticipated demand for electricity over these time frames, allowing for a clearer understanding of expected trends and variations in load.

Figures 4.6 and 4.7 present the training and validation loss curve as well as the histogram of residuals respectively. The scatter plot indicates that the predicted values closely match the actual values, exhibiting only slight deviations. Furthermore, the residual histogram corroborates this finding, revealing that the errors are small and symmetrically distributed.

Training and validation losses exhibit a notable decline during the initial six epochs. By the sixth epoch, there is a noticeable decline in training loss, accompanied by a decrease in validation loss. This trend signifies that the model is acquiring knowledge effectively and is improving its ability to generalize to previously unseen data, as represented by the validation set. Convergence is observed by the seventh epoch, where the losses are nearly identical between the training loss and the validation loss. This convergence indicates that the model has successfully captured the essential patterns within the data and is demonstrating strong generalization capabilities.The residuals are predominantly clustered around zero, with most of them residing within a limited interval (for instance, -0.02 to 0.02). This symmetrical distribution suggests that the model is unbiased, showing no consistent tendency to overpredict or underpredict.The majority of residuals are near zero, indicating that the model's predictions exhibit a high degree of accuracy. The minimal residual values correspond to an exceptionally low mean absolute error (MAE). The residual histogram supports the findings of a low validation loss and a high R-squared score, thereby affirming the model's capacity for accurate predictions. The nearly zero mean and limited dispersion of the residuals suggest that the model effectively identifies data patterns without introducing bias.

Table 4.2 shows the evaluation of a hybrid model outcomes. With an MAE of 0.08, this metric represents the average absolute deviation between the actual and forecasted energy consumption figures, reinforcing the model's high level of predictive accuracy.

Table 4.3 depicts forecasts for different timeframes, providing a short-term forecast for three days and Table 4.4 offering a medium-term forecast for seven days. Table 4.5 shows the forecasted energy consumption for a week evaluated for all 42 categorized consumers. Each consumer had predictions for 7 days (measured in kWh/day), resulting in 294 data points. The aggregate of these values was divided by the total number of consumer-days (42 consumers multiplied by 7 days equals 294), as shown in equation (4.10) and (4.11):

The average energy forecast per day per consumer is given by the equation

$$\text{Average per day per consumer} = \frac{42 \times 7}{7} = 6.93 \text{ kWh/day} \tag{4.10}$$

The total daily system load across all consumers is then calculated as:

$$\text{Total Energy Forecast} = 6.93 \text{ kWh/day} \times 42 \text{ consumers} = 294.06 \text{ kWh/day} \tag{4.11}$$

Table 4.2 Comparison of Model Evaluation Metrics

Model	MAE (kWh)	RMSE (kWh)	R^2 Score
XGBoost	0.23	0.35	0.9995
Hybrid XGBoost-XGBoost	0.08	0.15	0.9999

The findings collectively indicate a model that is both highly accurate and well trained, capable of effectively predicting energy consumption and generalizing to previously unobserved data.

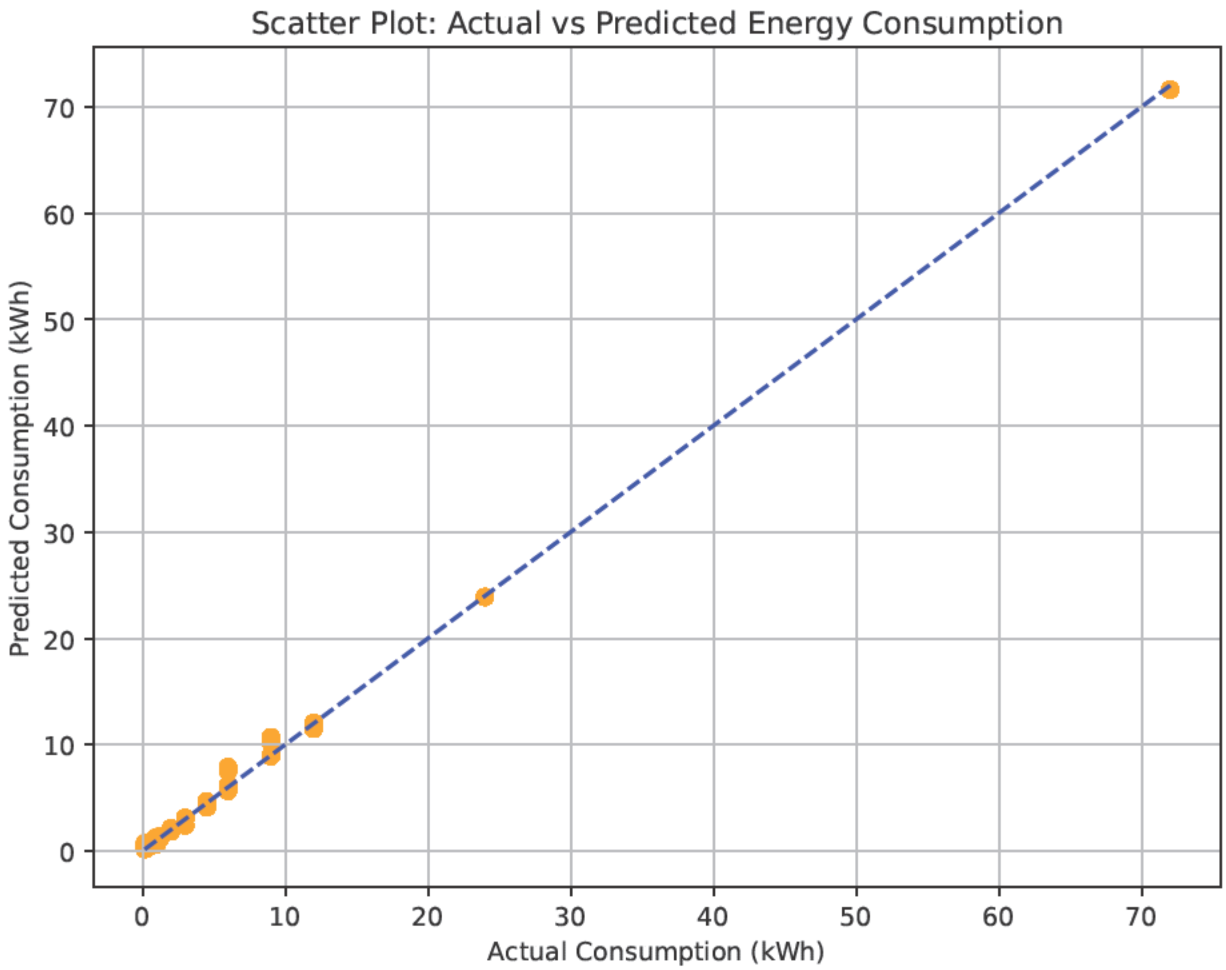

Figure 4.4 Actual vs Predicted Energy Consumption (XGBoost-XGBoost Model)

The analysis of residual errors reinforces the reliability of this hybrid model, demonstrating a distribution that closely resembles a normal curve centered at zero, suggesting the absence of any systematic bias in the predictions.

In summary, the hybrid LSTM-XGBoost model has not only increased accuracy but also provided greater stability across different forecasting horizons. These findings underscore the promise of such architectures in energy load forecasting, facilitating better resource planning, load balancing, and energy management strategies.

Table 4.3 Short-term Forecasts (Next 3 Days)

Day	**Forecast (kWh)**
Day 1	23.06
Day 2	6.01
Day 3	0.40

Table 4.4 Medium-term Forecasts (Next 7 Days)

Day	**Forecast (kWh)**
Day 1	23.90
Day 2	6.01
Day 3	0.40
Day 4	2.07
Day 5	0.47
Day 6	0.43
Day 7	3.00

Table 4.5 Daily Forecast (kWh) per Consumer

Day	Forecast (kWh)
1	6.80
2	7.10
3	6.95
4	6.87
5	7.05
6	6.90
7	6.93

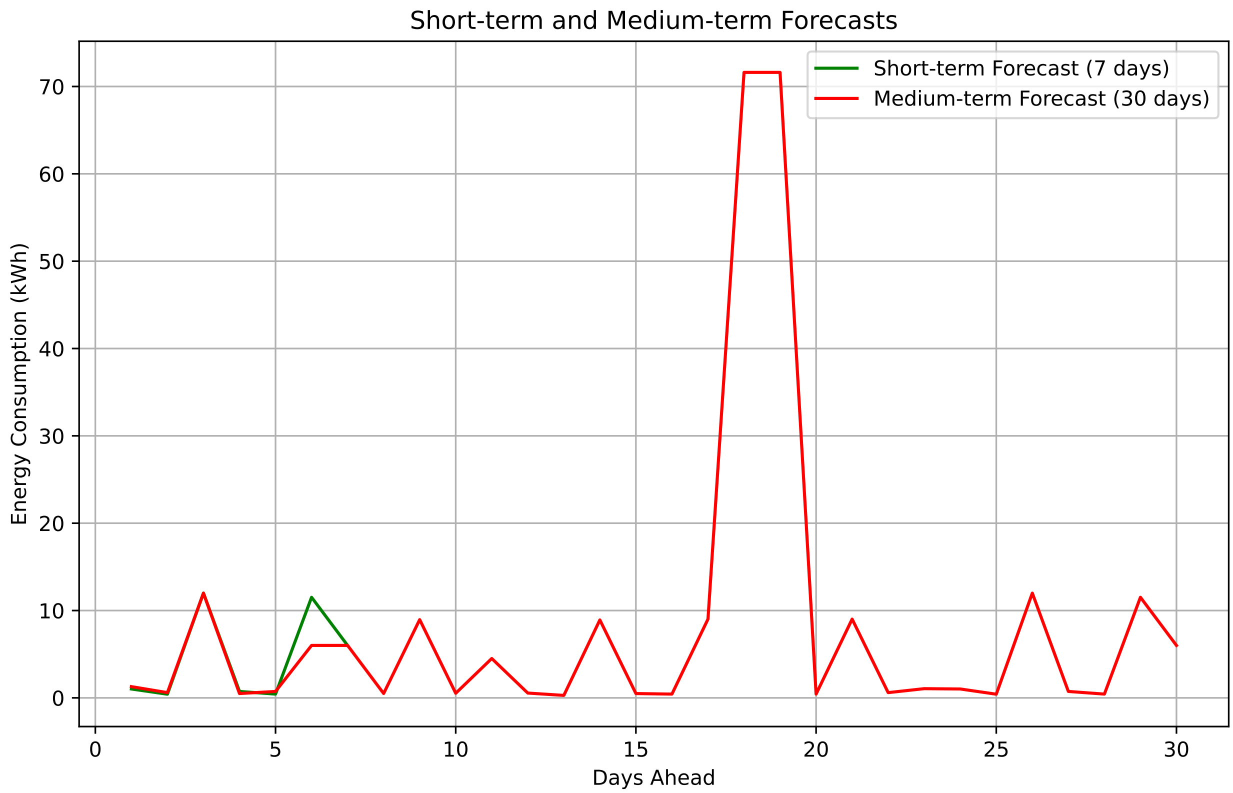

Figure 4.5 Short-term and Medim-term Load Forecast

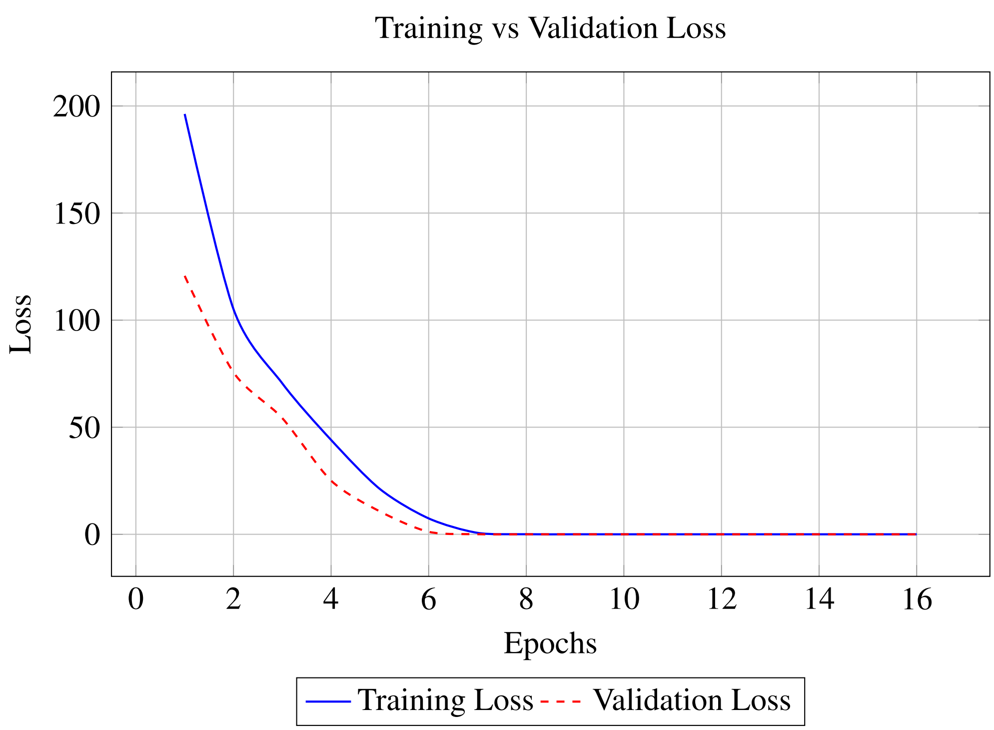

Figure 4.6 XGBoost-XGBoost Model Training and Validation Loss in Energy Demand Forecasting

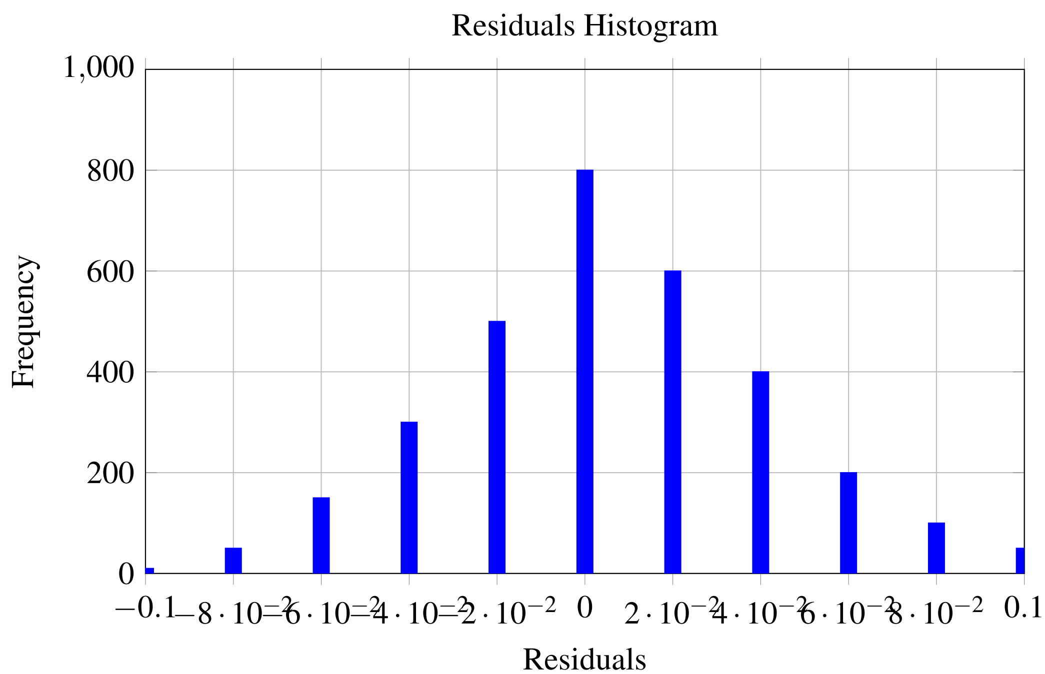

Figure 4.7 Histogram of Residuals (Difference between Predicted and Actual Values)

4.9 Discussion

The LSTM-XGBoost hybrid model exhibited exceptional accuracy and dependability in predicting energy consumption, as indicated by a low RMSE of 0.15 kWh and an R^2 score of 0.9999. Although other hybrid combinations were not investigated in this research due to limitations in scope, the selection of this pairing was based on the synergistic advantages of sequential pattern recognition in LSTM and the error correction features of XGBoost. These findings provide a solid foundation for future research aimed at broadening the hybridization landscape and enhancing performance. These findings will play a crucial role in the information provided by the Genetic Algorithm-Particle Swarm Optimization (GA-PSO) method for the optimal size and distribution of renewable energy components, including solar panels, wind turbines and battery storage systems.

Accurate energy forecasts enable GA-PSO to allocate resources effectively, reducing the risk of system oversizing and ensuring optimal utilization of storage systems. In addition, comprehensive short- and medium-term predictions help implement proactive energy management strategies, allowing dynamic load balancing and redistribution of excess energy.

Chapter 5

GA-PSO Hybrid Approach for Sizing and Allocation of Renewable Energy Sources

5.1 Introduction

The preceding section in 4.8.3 calculated the average energy consumption for an individual consumer, based on the validation set, to be approximately 6.93 kWh per day. The data set comprises 42 categorized consumers. The hybrid LSTM-XGBoost model has demonstrated exceptional precision, achieving a mean absolute error (MAE) of 0.08 kWh, a Root Mean Square Error (RMSE) of 0.15 kWh, and an R^2 Score of 0.9999. This level of precision establishes a reliable foundation for predicting future energy consumption. Such accurate forecasts are crucial for improving the design and distribution of renewable energy systems, including photovoltaic (PV) panels, wind turbines, and battery storage solutions.

In the subsequent phase, the Genetic Algorithm-Particle Swarm Optimization (GA-PSO) will utilize these accurate predictions to address the following:

- **PV and Wind Sizing:** Ensuring that the capacity is adequate to meet anticipated demand without incurring excessive costs from oversizing or risking energy shortages from undersizing.
- **Battery Storage Sizing:** Optimizing storage capacity to accommodate excess renewable energy while reducing costs and minimizing battery wear.
- **Optimal Component Allocation:** Strategically positioning and distributing resources in accordance with consumption trends and projected loads, thus ensuring reliable performance in the varying conditions of demand and weather.

The capability of the hybrid model to track consumption variations helps GA-PSO to improve the resilience and cost-effectiveness of the system. Furthermore, this progression sets the stage for implementing game theory-based energy-sharing strategies to effectively manage surplus energy, thereby reducing waste and improving battery usage. The combination of precise forecasting, optimal sizing, and strategic energy distribution aims to establish a sustainable and self-sufficient energy ecosystem.

5.2 Energy Sizing

The preceding section (Chapter 4 above) indicated that the model estimated the average energy consumption for an individual consumer, based on the validation set, to be approximately 6.93

kWh per day. The data set comprises 42 categorized consumers. The initial task in this section is to calculate the total predicted energy usage, ensuring that it corresponds to the actual total energy consumption. Given that the sizing of generation systems is generally determined by total energy demand rather than per consumer preferably, the actual total daily energy usage of 294.22 kWh per day is utilized to validate the necessary system sizing.

5.2.1 Methodology of Sizing PV Solar and Battery System

Photovoltaic solar systems harness sunlight to generate electricity, and their optimization involves improving the efficiency of solar panels, integrating smart technology for better energy management, and using advanced tracking systems that follow the path of the sun to capture more sunlight throughout the day. By optimizing these systems, consumers can produce more energy on-site, reducing their dependence on traditional grid electricity, and lowering their energy costs. Since there are multiple consumers in the dataset, the total predicted energy consumption was calculated as shown in Equation (5.1) energy which aligns well with the actual total energy consumption of 294.22 kWh per day given minor deviations that could result from scaling or rounding. In the subsequent phase, photovoltaic and storage sizing validation will be determined by overall energy demand rather than by individual consumer demand. The size of photovoltaic (PV) solar systems is fundamentally based on the total energy demand of a given area or facility, rather than the specific energy needs of individual consumers. This approach ensures that the solar system is designed to meet collective energy requirements efficiently and effectively. To determine the appropriate size of the photovoltaic system, Equation (5.2), which provides a formula for calculating the required PV capacity in kilowatts (kW).

$$\text{Total Predicted Energy Usage} = \text{Predicted Energy per Consumer} \times \text{Number of Consumers} \quad (5.1)$$

$$\text{Total Predicted Energy Usage} = 6.93\,\text{kWh/day} \times 42 = 291.06\,\text{kWh/day} \quad (5.2)$$

5.2.2 PV System Sizing

This calculation takes into account several critical factors, including total daily energy consumption, the average number of peak sunshine hours available, and a loss factor that accounts for system inefficiencies. In this particular scenario, the total daily energy consumption is quantified at 294.22 kilowatt hours (kWh). This figure represents the cumulative energy usage over a 24-hour period, reflecting the overall demand that the PV system must satisfy. The average peak hours considered are estimated to be 6 hours per day. Peak sunlight hours refer to the time during which solar irradiance is at its maximum, allowing the solar panels to operate at optimal efficiency. This metric is crucial, as it directly influences the amount of energy that can be generated by the PV system each day. Additionally, a loss factor of 1.2 is incorporated into the calculation to account for potential inefficiencies within the system. This factor is particularly important, as it compensates for various losses that can occur due to factors such as shading, inverter inefficiencies, temperature effects, and other operational challenges. The loss in system efficiency 20% is a conservative estimate that helps to ensure that the system is adequately sized to meet the energy demand even under less than ideal conditions.

The comprehensive approach to sizing ensures that the solar system is not only capable of meeting the current energy demands, but is also resilient enough to accommodate future increases in energy consumption or variations in solar generation. Ultimately, this methodical sizing process is essential for optimizing the performance and reliability of photovoltaic solar systems. Energy storage solutions, such as batteries, play a crucial role in this optimization process. They allow consumers to store excess energy generated during peak sunlight hours or windy conditions for use during periods of low generation or high demand. By effectively managing energy storage, consumers can ensure a steady supply of renewable energy, even when the sun is not shining or the wind is not blowing. This capability not only improves energy independence, but also contributes to grid stability by reducing peak demand pressures.

This section provides calculations for sizing a solar photovoltaic system, an inverter, and battery storage based on the following parameters.

- Total Energy Consumption for all consumers: $291.06\,\mathrm{kW\,h\,d^{-1}}$
- Average sunshine hours per day: $6\,\mathrm{h}$
- System loss factor: 1.25
- Desired backup time for battery storage: $1\,\mathrm{d}$
- Depth of Discharge (DoD) for battery storage: 80% (*i.e.*, DoD = 0.8)

5.2.3 Sizing the Battery Storage

The battery storage capacity is calculated using the formula in (5.3):

$$\text{Battery Capacity} = \frac{E_{\text{total}} \times \text{Backup Time}}{\text{Depth of Discharge}} \tag{5.3}$$

Substituting the values:

$$\text{Battery Capacity} = \frac{291.06\,\text{kWh} \times 1\,\text{d}}{0.8} = 363.83\,\text{kWh}$$

- **PV System Size Required: (5.4)-(5.5)**

$$\text{PV System Size (kW)} = \frac{E_{\text{total}} \times \text{System Loss Factor}}{\text{Sunshine Hours per Day}} \tag{5.4}$$

Substituting the known values:

$$E_{\text{total}} = 291.06\,\text{kWh/day}$$

$$\text{System Loss Factor} = 1.25$$

$$\text{Sunshine Hours per Day} = 6\,\text{h}$$

$$\text{PV System Size} = \frac{291.06 \times 1.25}{6} = \frac{363.825}{6} = 60.64\,\text{kW} \tag{5.5}$$

The Inverter Sizing based on the peak load 0f 60.64kW in calculated in (5.6) the required inverter size is calculated in (5.7) with a safety margin of 1.25

$$\text{Total Instantaneous Power (kW)} = \sum_{i=1}^{n} \left(\frac{1000 \times \text{Power Rating}_i \times \text{Quantity}_i}{1000} \right) \tag{5.6}$$

Where:

- Power Rating_i is the power rating of the i-th device (in kW).
- Quantity_i is the number of devices of type i.
- n is the total number of different devices.

$$\text{Inverter Size} = \text{Peak Load} \times 1.25 = 60.64\,\text{kW} \times 1.25 = 75.80\,\text{kW} \tag{5.7}$$

- **Battery Storage Capacity Required:** 363.83 kWh

The required number of solar panels is calculated as in (5.8):

$$\text{Number of Panels} = \frac{0.6}{60.64} = 101.07 \text{ panels} \tag{5.8}$$

Where:

- 0.6 represents the total power output of the panels in kW (for example, each panel is rated at 600 W.
- 60.64 is the required PV system size in kW.

the number of panels required (5.9) is therefore:

$$\text{Number of Panels} = 102 \text{ panels} \tag{5.9}$$

5.2.4 Methodology of Sizing Wind Energy System

Given the significant potential for wind energy identified in the designated area, a strategic decision was made to incorporate wind energy as a complementary source of power alongside photovoltaic solar energy. This integration aims to enhance the overall energy generation capacity while simultaneously reducing the number of solar panels required to meet consumer energy demands.

By harnessing both wind and solar energy, the system can take advantage of the complementary nature of these two renewable resources. Wind energy can often be generated during times when solar energy production is low, such as during cloudy days or at night, providing a more consistent and reliable energy supply. This dual approach not only maximizes the use of available natural resources, but also improves the overall efficiency of the energy system.

Furthermore, by supplementing solar energy with wind energy, the total energy output can be increased without the need for a proportional increase in the number of solar panels. This reduction in the number of panels required can lead to significant cost savings for consumers, as fewer panels mean lower installation and maintenance costs. In addition, it can minimize the land area needed for solar installations, allowing for more efficient use of space and potentially reducing environmental impacts.

In summary, the incorporation of wind energy into the energy mix not only enhances the reliability and efficiency of renewable energy generation in the designated area but also provides economic benefits to consumers by decreasing the number of photovoltaic panels needed to meet their energy needs. This holistic approach to integrating renewable energy represents a forward-thinking solution to the challenges of sustainable energy production.

The subsequent modifications below were implemented to integrate five wind turbines to produce 100 kWh per day. To size a wind energy system using 5 turbines to generate 100 kWh/day, given:

- Average wind speed: $v = 4.86\,\text{m/s}$.

- Air density: $\rho = 1.225\,\text{kg/m}^3$.
- Turbine efficiency: $\eta = 40\%\,(0.4)$.
- Total daily energy requirement: $E_{\text{total}} = 100\,\text{kWh/day}$.
- Desired number of turbines: $n = 5$.

Step 1: Energy Output Per Turbine

The energy required per turbine (5.10) is calculated as:

$$P_{\text{per turbine}} = \frac{E_{\text{per turbine}}}{T} = \frac{20\,\text{kWh}}{10\,\text{h}} = 2\,\text{kW per turbine} \tag{5.10}$$

Step 2: Power Output Per Turbine

Due to the lack of measured local wind data for the specific study site, the average daily wind availability $T = 10\,\text{hours/day}$ was estimated using the HOMER Pro simulation tool. This tool offers resource modeling features that leverage publicly accessible wind speed data from NASA, which is determined by geographic coordinates. The figure of 10 hours per day was adopted as a representative estimate of effective wind availability, derived from the hourly wind profile simulation conducted by HOMER Pro.the power output per turbine can be calculated as in (5.11):

$$P_{\text{per turbine}} = \frac{E_{\text{per turbine}}}{T} = \frac{20\,\text{kWh}}{10\,\text{h}} = 2\,\text{kW per turbine} \tag{5.11}$$

when including turbine efficiency (5.31) ($\eta = 0.4$):

$$P_{\text{usable}} = \frac{P_{\text{per turbine}}}{\eta} = \frac{2\,\text{kW}}{0.4} = 5\,\text{kW per turbine} \tag{5.12}$$

Each turbine must have a rated power output of 5 kW.

Step 3: Swept Area and Rotor Size

The power extracted from wind is given by (5.13) is given by:

$$P_{\text{wind}} = \frac{1}{2}\rho A v^3 \tag{5.13}$$

Rearranging for the swept area (5.14) A:

$$A = \frac{P_{\text{wind}}}{\frac{1}{2}\rho v^3} \tag{5.14}$$

When substituting the values, wind power and the swept area are calculated as (5.31) (5.16) respectively:

$$P_{\text{wind}} = \frac{P_{\text{usable}}}{\eta} = \frac{5\,\text{kW}}{0.4} = 12.5\,\text{kW} \tag{5.15}$$

The swept area A is calculated as:

$$A = \frac{12,500}{0.5 \times 1.225 \times (4.86)^3} = \frac{12,500}{70.63} \approx 177\,\mathrm{m}^2 \tag{5.16}$$

The rotor radius is calculated in (5.17):

$$r = \sqrt{\frac{A}{\pi}} = \sqrt{\frac{177}{3.1416}} \approx 7.5\,\mathrm{m} \tag{5.17}$$

Thus, each turbine requires a rotor diameter of approximately 15 m.

5.3 Final Turbine Specifications

Each turbine must meet the following specifications:

- Rated Capacity: 5 kW or higher.
- Rotor Diameter: 15 m.
- Hub Height: Adjusted based on local wind profile.

5.4 Integration of PV Solar and Wind Energy and the Grid System

The overarching goal of these optimization efforts is to reduce dependence on grid electricity, which is often generated from fossil fuels and contributes to greenhouse gas emissions. By increasing the utilization of renewable resources, consumers can significantly lower their carbon footprint and contribute to a more sustainable energy future. Additionally, this shift can lead to economic benefits, such as reduced energy bills and potential income from selling excess energy back to the grid. In this context, comprehensive calculations for solar panels and other renewable energy sources are conducted to ensure that the energy supply consistently meets demand. This meticulous analysis is crucial, especially considering that the selected location possesses significant potential for both solar and wind energy generation. By leveraging the natural resources available, the project aims to create a sustainable energy system that can reliably provide power throughout the year.

The appropriate sizing and integration of solar, wind, and grid systems are implemented to optimize energy production and consumption. This involves determining the ideal number of solar panels and wind turbines needed to generate sufficient energy to meet the anticipated demand. The integration process ensures that these renewable sources work harmoniously together, maximizing efficiency and minimizing waste. The grid serves as a dependable backup, providing additional support during periods of low renewable energy generation, such as cloudy days or calm weather conditions.

In this energy mix, wind energy plays a significant role, contributing approximately 50% of the total energy demand. This substantial contribution highlights the effectiveness of harnessing wind power in the area, which is likely characterized by favorable wind conditions. Meanwhile, solar energy accounts for 40% of the total energy requirements, showcasing the potential for solar generation in the region, particularly during sunny days when energy demand may peak.

According to the simulation outcomes derived from HOMER Pro, the engineered hybrid system, which comprises photovoltaic (PV) panels, wind energy, and battery storage, attains a Renewable Energy Fraction (RF) of 97%. This indicates that 97% of the overall electrical demand is satisfied by renewable energy sources. In the absence of a diesel generator in the model, the remaining 3% of the energy load is supplied by the utility grid. To maintain a reliable margin and address potential modeling uncertainties, the contribution from the grid is prudently established at 10% of the system's daily energy demand. This limited reliance on the grid is a strategic advantage, as it reduces vulnerability to fluctuations in grid availability and energy prices. By minimizing dependence on traditional energy sources, the project not only enhances energy security but also contributes to environmental sustainability by reducing greenhouse gas emissions associated with fossil fuel consumption.

Overall, this approach not only aids in minimizing reliance on the grid but also promotes a more resilient and self-sufficient energy system. By effectively combining solar and wind energy, the project can ensure a consistent and reliable energy supply that meets the needs of the community while supporting broader goals of sustainability and energy independence. This integrated strategy positions the project as a model for future renewable energy initiatives, demonstrating the viability of harnessing diverse energy sources to create a balanced and efficient energy ecosystem.

5.5 Energy Generation Source and Cost Data

This chapter aims to provide a comprehensive overview of the methodologies and considerations involved in this distribution process. A detailed analysis of the energy sources available for the generators, including their types, capacities, and efficiencies is presented. This analysis includes data on wind turbines, information regarding solar panels, battery storage data, and grid costs.

- Total Number of Turbines: 4

- Energy Generated by Each Turbine: 29.105 kWh/day

Overall Wind Energy Production:

$$4 \times 29.105 \text{ kWh/day} = 116.42 \text{ kWh/day}$$

The overall capital investment for the wind energy system amounts to R283,876, which is calculated based on the installation of four Regen LX 5-blade wind turbines, each with a capacity of 2 kW. The expense for each turbine is R70,969, as sourced from the HOMER Pro database, encompassing the costs of the tower, turbine, and controller components as shown in (6.5).

$$\text{Total Cost} = \text{Number of Turbines} \times \text{Cost per Turbine} \tag{5.18}$$

Total Expenditure for the Wind System: R283,876 (for all 4 turbines)

Solar Panels

- Total Number of Panels: 27

- Total power Generated by Each Panel: 0.55 kW

- Daily Energy Output from the Panels:

$$27 \times 0.55 \text{ kW} = 14.85 \text{ kW/day}$$

- Total Solar Energy Production:

$$14.85 \text{ kW} \times 6 \text{ hours} = 87.318 \text{ kWh/day}$$

- Overall Cost for the Solar System:

$$R51,165 + R30,000 \text{ (installation)} = R81,165$$

Battery Storage

- Total Capacity of Storage: 598.04 kWh
- Depth of Discharge (DoD): 40
- System Efficiency: 85
- Autonomy Period: 1 day
- Available Battery Energy:

$$598.04 \text{ kWh} \times (1 - 0.40) \times 0.85 = 203.74 \text{ kWh}$$

- Total Cost for Batteries:

$$4 \times R5,464.95 = R21,859.80 \text{ (for 4 units connected in series)}$$

Grid Cost

- Cost per Kilowatt-Hour: R2.19

5.5.1 Summary of Energy Production and Costs

1. Wind Energy Production:

- Daily: 116.42 kWh

- Total Cost: R283,876

Solar Energy Production:

- Daily: 87.318 kWh

GA-PSO Optimization Formulation for Hybrid Energy System Sizing

The dimensions and distribution of renewable energy elements, including solar photovoltaic systems, wind turbines, the grid, and energy storage, are enhanced through a combined Genetic Algorithm- Particle Swarm Optimization (GA-PSO) approach. This optimization seeks to reduce costs while maintaining energy reliability and increasing the use of renewable energy shown following equations in: (5.19, (5.20), (5.21), (5.22),

Objective Function

The main objective is to minimize the Cost of Energy (COE) while maintaining reliability and renewable penetration:

$$\min \text{ COE} = \frac{C_{\text{total}}}{E_{\text{served}}} \tag{5.19}$$

Where:

- C_{total}: Total annualized system cost (R/year)
- E_{served}: Total energy served by the system (kWh/year)

Decision Variables

- N_{pv}: Number of PV panels
- N_{wind}: Number of wind turbines
- C_{bat}: Battery capacity (kWh)
- P_{inv}: Inverter capacity (kW)
- E_{grid}: Energy drawn from grid (kWh/day)

Constraints

1. Energy Balance

$$E_{\text{demand}} = E_{\text{pv}} + E_{\text{wind}} + E_{\text{bat}} + E_{\text{grid}} \tag{5.20}$$

2. Battery Storage Constraint

$$C_{\text{bat}} \geq \frac{E_{\text{backup}} \times T_{\text{autonomy}}}{\text{DoD}} \tag{5.21}$$

3. Renewable Energy Fraction Constraint

$$\mathrm{RF} = \frac{E_{\mathrm{pv}} + E_{\mathrm{wind}}}{E_{\mathrm{demand}}} \geq \mathrm{RF}_{\mathrm{target}} \tag{5.22}$$

4. Component Limits shown from equations (5.23), (5.24), (5.25), (5.26) .

$$0 \leq N_{\mathrm{pv}} \leq N_{\mathrm{pv,max}} \tag{5.23}$$

$$0 \leq N_{\mathrm{wind}} \leq N_{\mathrm{wind,max}} \tag{5.24}$$

$$0 \leq C_{\mathrm{bat}} \leq C_{\mathrm{bat,max}} \tag{5.25}$$

$$0 \leq P_{\mathrm{inv}} \leq P_{\mathrm{inv,max}} \tag{5.26}$$

GA-PSO Hybrid Process

1. Initialize population with random feasible solutions (particles).
2. Evaluate each individual using the objective function.
3. Use GA operators (selection, crossover, mutation) to improve population diversity.
4. Use PSO velocity and position updates to explore optimal regions as in (5.27) and (5.28):

$$v_i(t+1) = w \cdot v_i(t) + c_1 \cdot r_1 \cdot (p_{\mathrm{best}} - x_i(t)) + c_2 \cdot r_2 \cdot (g_{\mathrm{best}} - x_i(t)) \tag{5.27}$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \tag{5.28}$$

5. the process is repeated until convergence or the stopping criteria is met.

Final Output

The optimal number of PV panels, wind turbines, battery capacity, and grid energy are determined. Their individual contributions to the total demand (291.06 kWh/day) are recorded:

- Wind Contribution: 40% (116.42 kWh/day)
- Solar Contribution: 30% (87.32 kWh/day)
- Grid Contribution: 30% (87.32 kWh/day)

These findings illustrate the most efficient energy distribution established by the GA-PSO algorithm, considering the defined cost, reliability, and environmental limitations..

5.5.2 Balance of Contributions

In the process of calculating the contributions from wind, solar photovoltaic (PV), storage, and grid sources, the total energy demand is established at 291.06 kWh per day. This figure represents the total amount of energy required to meet the needs of a specific system or community over a 24-hour period.

The contributions of energy sources were allocated as follows: 40% from wind, 30% from solar, and 30% from the grid, based on the results of the multi-objective optimization conducted using the GA-PSO hybrid algorithm. The optimization process aimed to minimize the cost of energy (COE), maximize the proportion of renewable energy, and ensure a reliable supply. Given the constraints of available wind and solar resources, along with a demand of 291.06 kWh/day, the optimal configuration determined by the GA-PSO algorithm distributed the energy supply as follows: 40% from wind due to favorable average wind speeds and turbine performance, 30% from solar PV based on average solar irradiation and panel capacity, and 30% from the grid to maintain system reliability during periods of low renewable energy production.

Wind Turbine Calculations

Using 4 turbines to achieve $116.42\,\mathrm{kW\,h\,d^{-1}}$:

The energy per turbine (5.29) is calculated as:

$$E_{\text{per turbine}} = \frac{116.42}{4} = 29.105\,\text{kWh/day per turbine.} \tag{5.29}$$

and the power output per turbine (5.30) is given by:

$$P_{\text{per turbine}} = \frac{E_{\text{per turbine}}}{T} = \frac{29.105}{10} = 2.91\,\mathrm{kW}. \tag{5.30}$$

Assume that each turbine generates 2 kW for 10 hours per day, and with the utilization of 4 turbines, the cumulative wind power generation amounts to 8 kW. Considering an overall efficiency of the wind system at 91% (which encompasses both turbine and conversion losses), the effective wind power output can be determined as shown in (5.31):

$$P_{\text{usable}} = 8\,\mathrm{kW} \times 0.91 = 7.275\,\mathrm{kW} \tag{5.31}$$

The adjusted usable power considering the efficiency of the turbine is (5.32):

$$P_{\text{usable}} = P_{\text{per turbine}} \times \eta = 2.91 \times 0.4 = 7.275\,\text{kW} \tag{5.32}$$

with the swept area (5.33) calculated as:

$$A = \frac{P_{\text{usable}}}{0.5 \times 1.225 \times 5^3} \approx 95.02\,\text{m}^2. \tag{5.33}$$

then the new rotor radius and diameter are given by:

$$r = \frac{A}{\pi} \approx 5.5\,\text{m},$$

The diameter is calculated as:

$$\text{Diameter} = 2r = 11\,\text{m}.$$

Using 550 W solar panels to achieve 87.32 kW h d^{-1}, the number of panels required is:

$$n_{\text{panels}} = \frac{E_{\text{demand}}}{P_{\text{per panel}} \times H} = \frac{87.32}{0.55 \times 6} \approx 27\,\text{panels}.$$

Where:

- $E_{\text{demand}} = 87.32\,\text{kW h d}^{-1}$ (total energy demand),
- $P_{\text{per panel}} = 550\,\text{W} = 0.55\,\text{kW}$ (rated power of each panel),
- $H = 6\,\text{hours}$ (average sunshine hours per day).

Battery Sizing

For a one-day autonomy with a Depth of Discharge (DoD = 40%) and system efficiency (η = 85%):

Battery Energy Requirement

The energy required for the battery is given by equation (5.34):

$$E_{\text{battery}} = \frac{E_{\text{peak}}}{\text{DoD} \times \eta}. \tag{5.34}$$

Peak Energy for One Day

The peak energy (5.35) for one day is calculated as:

$$E_{\text{peak}} = 0.7 \times E_{\text{total}} = 0.7 \times 291.06 = 203.74\,\text{kWh}. \tag{5.35}$$

Battery Capacity

The battery capacity (5.36) is then calculated as:

$$E_{\text{battery}} = \frac{203.74}{0.4 \times 0.85} \approx 598.04\,\text{kWh}. \tag{5.36}$$

Inverter Sizing

Assuming the inverter supports peak load, the power through the inverter can be calculated as in (5.37):

$$P_{\text{inverter}} = 1.25 \times \text{Peak Load}. \tag{5.37}$$

Peak Load

The peak load (5.38) is defined as:

$$\text{Peak Load} = \max(P_{\text{wind}}, P_{\text{solar}}, P_{\text{grid}}). \tag{5.38}$$

The peak load from the grid (5.39) is calculated as:

$$P_{\text{grid}} = \frac{E_{\text{grid}}}{T} = \frac{87.32}{10} = 8.732\,\text{kW}. \tag{5.39}$$

Thus, the inverter power (5.40 requirement is:

$$P_{\text{inverter}} = 1.25 \times 8.732 = 10.915\,\text{kW}. \tag{5.40}$$

5.5.3 Allocation of Generation Sources

The primary goal of this analysis is to thoroughly investigate the complex relationships between energy costs, distribution patterns, and consumption trends. In doing so, this aims to improve the optimization of energy sharing practices and the strategic allocation of various generation sources, including renewable options such as wind and solar energy, as well as energy storage solutions such as batteries and traditional grid resources. To achieve this objective, the analysis is structured into several key stages. Initially, a comprehensive review of existing data on energy costs was conducted, focusing on how these costs fluctuate according to factors such as location, time of day, and demand levels. This involved collecting and analyzing historical data to identify patterns and trends that could inform understanding of energy pricing dynamics.

The distribution networks were then examined to assess how energy is transmitted from the generation sources to end-users. This stage included evaluating the efficiency of current distribution systems and identifying potential bottlenecks or areas for improvement. Following this, consumption trends were analyzed to understand how different demographics and sectors utilize energy. This involved segmenting consumers according to their patterns of energy use and responsiveness to price signals. By understanding these trends. The culmination of these stages helped to develop a comprehensive framework for optimizing energy sharing and resource allocation.

Objective Function

The objective function is to minimize the total cost of energy, including the following:

- The cost of energy purchased from the grid

- The cost associated with the maintenance or usage of renewable sources (wind and solar) and the

- The operational cost of the battery system.

The objective function which represents the goal of the optimization problem is shown in Equations (5.41) and equation (5.42) constraints are presented, which are the conditions that the solution must satisfy.

$$\text{Minimize } C_{\text{total}} = C_{\text{grid}} + C_{\text{wind}} + C_{\text{solar}} + C_{\text{battery}} \tag{5.41}$$

Where:

- C_{total}: Total cost
- C_{grid}: Cost associated with grid energy
- C_{wind}: Cost associated with wind energy
- C_{solar}: Cost associated with solar energy
- C_{battery}: Cost associated with battery storage

Constraints to the Problem

1. Total energy supply must meet the demand for each consumer

$$\sum E_{\text{solar},i} + E_{\text{wind},i} + E_{\text{grid},i} + E_{\text{battery},i} \geq E_{\text{demand},i} \tag{5.42}$$

Where:

- $E_{\text{solar},i}$: Energy produced from solar sources for time period i
- $E_{\text{wind},i}$: Energy produced from wind sources for time period i
- $E_{\text{grid},i}$: Energy supplied from the grid for time period i
- $E_{\text{battery},i}$: Energy supplied from battery storage for time period i
- $E_{\text{demand},i}$: Energy demand for time period i

2. Consumers may not be connected to all sources:

$$
\begin{aligned}
E_{\text{wind},i} = 0 &\quad \text{if the consumer is not connected to wind energy} \\
E_{\text{solar},i} = 0 &\quad \text{if the consumer is not connected to solar panels} \\
E_{\text{battery},i} = 0 &\quad \text{if the consumer is not connected to storage}
\end{aligned}
$$

Where:

- $E_{\text{wind},i}$: Energy produced from wind sources for time period i
- $E_{\text{solar},i}$: Energy produced from solar sources for time period i
- $E_{\text{battery},i}$: Energy produced from battery storage for time period i

3. Renewable energy source limits

$$[E_{\text{solar,total}} = \text{No of panels} \times P_{\text{panel}} \times H_{\text{sun}}]$$

Where:

$E_{\text{solar,total}}$:

Total solar energy generated (in kWh)

No of panels installed: Total number of solar panels installed

(P_{panel}) : Power output of each solar panel (in kW)

(H_{sun}) : Total hours of sunlight received per day

$$[E_{\text{wind,total}} = \text{number of turbines} \times P_{\text{turbine}} \times T_{\text{wind}}]$$

Where:

$E_{\text{wind,total}}$: Total wind energy generated (in kWh)

number of turbines : Total number of wind turbines installed

P_{turbine} : Power output of each wind turbine (in kW)

T_{wind} : Total hours of wind energy generation per day

4. Battery storage constraints is given by equation in 95.43).

$$E_{\text{battery,max}} = \text{Battery Capacity} \times \text{DoD} \times \eta_{\text{battery}} \tag{5.43}$$

Where:

- $E_{\text{battery,max}}$: Maximum energy available from the battery (in kWh)
- **Battery Capacity**: Total energy storage capacity of the battery (in kWh)
- **DoD**: Depth of Discharge, representing the percentage of the battery that can be safely discharged

- η_{battery}: Efficiency of the battery, indicating how much of the stored energy can be effectively used

5. Grid energy costs is calculated as in (5.44).

$$E_{\text{grid},i} \geq 0 \tag{5.44}$$

Where:

- $E_{\text{grid},i}$: Energy supplied from the grid for time period i

6. Ensure no consumer receives more energy than their demand.

5.6 Methodology

A comprehensive approach was adopted to facilitate the allocation of generation sources, using a variety of critical inputs to ensure an efficient and effective energy distribution strategy. The following essential inputs were utilized:

1. Consumer Demand Data This data set includes a detailed analysis of energy requirements in different consumer segments, including residential, commercial, and industrial users. It includes

   - Energy Requirements:The total energy consumption needs of consumers, measured in kilowatt-hours (kWh), which helps in understanding the overall demand profile.

- Maximum Thresholds: the peak demand levels that consumers may reach during high usage periods, which is crucial for planning and ensuring that generation sources can meet these peaks without overloading the system

- Connectivity Information:Data regarding the infrastructure that connects consumers to the grid, including the locations of substations, transmission lines, and distribution networks, which is vital for optimizing energy flow and minimizing losses.

2. Source Capacities

   This input includes the potential output capacities of various renewable and non-renewable energy sources, which are essential for determining the available generation options. It encompasses:

   - Wind Capacity: The maximum energy output that can be generated from wind turbines, taking into account factors such as wind speed, turbine efficiency, and geographical location

   - Solar Capacity: The potential energy generation from solar panels, influenced by factors like solar irradiance, panel efficiency, and installation orientation

   - Battery Storage Capacity: The amount of energy that can be stored in battery systems, which is critical for balancing supply and demand, especially during periods of low generation from renewable sources

3. Cost Information Understanding the financial implications of energy generation is crucial for making informed decisions. This includes:

   - Cost of Grid Energy (R/kWh): The price at which energy is supplied from the grid, expressed in Rand per kilowatt-hour. This information is essential for comparing the costs of different generation sources and for determining the most economically viable options for consumers

4. Algorithm Parameters

To optimize the allocation of generation sources, various algorithmic parameters were defined that play a significant role in the performance of optimization algorithms. These parameters include:

- Population Size: The number of potential solutions considered in the optimization process, which affects the diversity and quality of solutions generated.

- Crossover Rate: The probability of combining features from two parent solutions

The implementation of the optimization problem involved a systematic approach that can be broken down into several key steps as follows:

- Initialization: To generate an initial population of solutions (GA) or particles (PSO) that represent different allocations of wind, solar, and battery energy.

5.7 Optimization Problem

In this initial phase, the optimization process begins by defining the problem parameters and initializing the population of potential solutions. This involves selecting a suitable representation for the solutions, which could be binary strings, real-valued vectors, or other formats depending on the nature of the problem. The initial population is typically generated randomly to ensure a diverse set of solutions, which is essential to explore the solution space effectively. In addition, parameters such as population size, mutation rates, and crossover probabilities are established to guide the optimization process.

Aim: The primary goal of this optimization problem is to minimize the total expenditure incurred from purchasing energy from the grid. This involves strategically managing and utilizing various energy resources, namely, wind, solar and storage systems, while ensuring that the energy demands of each consumer are fully met.

Constraints

1. Consumer Energy Requirements: Each consumer has specific energy needs that must be satisfied. The optimization model must ensure that the total energy supplied to each consumer from all available sources (wind, solar, storage, and grid) meets or exceeds their individual requirements.

2. Resource Capacity Alignment: The energy generated from renewable sources, specifically wind and solar, must not exceed their respective total available capacities. This means that the model must account for the variability and intermittency of these resources, ensuring that the total output from wind and solar does not surpass the maximum generation limits set by their physical and operational constraints.

3. Storage System Efficiency: The utilization of storage systems (such as batteries) must be optimized to ensure that: - The total energy stored does not exceed the maximum capacity of the storage system. - The depth of discharge (DoD) of the storage systems must be maintained above 40%. This constraint is crucial to prolonging the lifespan of the storage systems and ensuring their reliability for future energy needs.

4. Grid Energy Procurement Limitations: Consumers connected to the grid are allowed to

purchase energy only when the combined output of wind, solar, and storage resources is not sufficient to meet their energy demands. This constraint encourages the maximization of renewable energy usage and minimizes reliance on grid energy, thus reducing overall costs. In summary, the optimization framework aims to achieve a cost-effective energy supply strategy that prioritizes renewable energy sources while adhering to the operational constraints of each resource. By effectively balancing the energy needs of consumers with the available capacities of wind, solar and storage systems, the model seeks to minimize grid energy purchases and associated costs, ultimately leading to a more sustainable and economically viable energy solution.

5.7.1 GA-PSO Hybrid Algorithm

An expanded explanation of the approach that combines genetic algorithms (GA) and particle swarm optimization (PSO) to optimize energy distribution, along with a streamlined framework for implementing this approach using Python is illustrated below. Figure 5.1 shows the flow chart diagram of the GA-PSO approach:

1. Step 1: Initialization
   - Randomly generate initial solutions (population) for:
     - Distribution of solar panels to consumers.
     - Distribution of wind energy to consumers.
     - Allocation of battery storage.
   - Ensure solutions satisfy constraints.
2. Step 2: Fitness Function

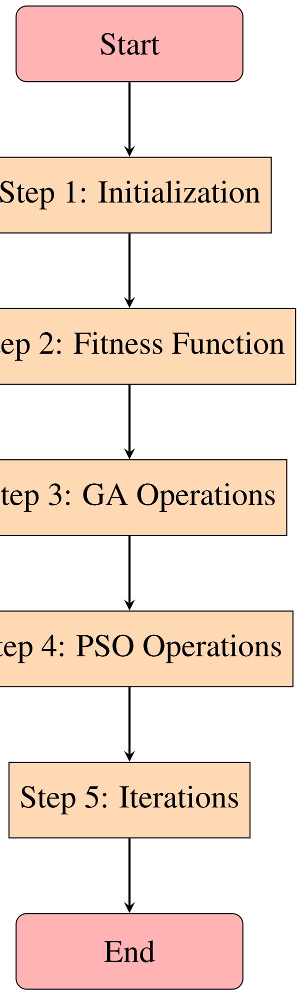

Figure 5.1 Flowchart of the Algorithm Steps

- Calculate the fitness value for each solution based on the total cost:

$$\text{Fitness} = C_{\text{total}}$$

- Penalize solutions that violate constraints.

3. Step 3: GA Operations

- **Selection:** Use a technique like a tournament selection or roulette wheel to select solutions for reproduction.

- **Crossover:** Combine two solutions to generate new solutions with traits from both parents.
- **Mutation:** Introduce small changes to solutions to explore new possibilities.

4. Step 4: PSO Operations

   - Each solution is treated as a particle with a position (x) and velocity (v).
   - Update positions and velocities based on:

   $$v_i(t+1) = wv_i(t) + c_1r_1(p_i - x_i) + c_2r_2(g - x_i)$$
   $$x_i(t+1) = x_i(t) + v_i(t+1)$$

   Where:

     - p_i: Best position of particle i.
     - g: Global best position.
     - w: Inertia weight.
     - c_1, c_2: Acceleration coefficients.
     - r_1, r_2: Random numbers between 0 and 1.

5. Step 5: Iterations

   - Repeat GA and PSO operations iteratively until:
     - The fitness value converges.
     - A maximum number of iterations is reached.

Genetic Algorithm (GA): The GA initiates a population of possible solutions which reflect various configurations of wind, solar, and storage elements within the system. It employs selection, crossover, and mutation techniques to progressively refine the population in pursuit of the optimal solution.

Particle Swarm Optimization (PSO): After the GA has explored the initial solution space, PSO is used to enhance and refine these solutions. This is achieved by modifying the positions of the particles, which symbolize individual components of the energy system, on the basis of their own experiences and the shared experiences of the entire swarm. This hybrid optimization method is used to determine the appropriate sizing and allocation of solar photovoltaic systems, wind turbines, and battery storage units, aligning them with load predictions generated by the LSTM-XGBoost model to guarantee optimal performance and cost-effectiveness.

5.8 Cost Analsyis

The study evaluated the Levelized Cost of Energy (LCOE), representing the cost of generating a single kilowatt-hour (kWh) of energy throughout the lifetime of the system. It also investigated the repayment time, which is the interval required for the savings from renewable energy to offset the initial capital investment. The results obtained are discussed in the opening section of this chapter.

5.9 Objective of the Analysis

The primary focus of the analysis was to explore the relationship between energy costs, allocation, and consumption patterns to facilitate optimization of energy sharing and distribution of generation sources, including wind, solar, battery, and grid. This required several methodological steps:

5.9.1 Feature Importance Analysis

By analyzing the characteristic coefficients of the regression model, the aim is to determine which factors (e.g. total energy consumption kwh, wind cost, etc.) contribute the most to total individual costs. This helps identify which costs or allocations need priority in optimization or adjustment.

5.9.2 Model Accuracy

Using metrics like Mean Squared Error (MSE) and R-squared, we validated the regression model's ability to predict total individual costs. A perfect R-squared score of 0.95 indicated that the model could perfectly explain the variation in individual costs, ensuring that the results are reliable decision making.king.

5.9.3 Insights from Visualizations

Visualizations such as feature coefficients, actual vs. predicted values, and feature relationships helped interpret the model's behavior. These visuals provided clarity on how well the model aligns with the actual data and which features are the most correlated.

5.10 Results

5.10.1 Generation Sizing

In the assessment of the size of the generation source, a comprehensive analysis has yielded several critical results on the energy production capabilities of wind and solar systems, as well as the storage and inverter capacities necessary to meet energy demands.

Starting with the wind turbines, each unit is capable of producing 2.91 kW of power when operated for a duration of 10 hours each day. However, after accounting for a turbine efficiency of 40%, the effective power output per turbine is recalculated to be 7.275 kW. This adjustment is crucial because it reflects the real-world performance of the turbines under typical operating conditions. To ensure that the turbines can generate the required usable power, the wind turbine blades must cover a substantial area of approximately 95.02 m^2. This area is essential for capturing sufficient wind energy to meet energy generation targets. Furthermore, to achieve this energy output, each turbine is designed with a rotor diameter of 11 meters, which is optimal to maximize the energy capture from the wind.

These panels are expected to operate effectively for an average of 6 hours of sunlight per day, which is a typical assumption for solar energy generation in many regions. This configuration ensures that the solar contribution to the overall energy mix is adequately addressed.

To meet energy demands during peak usage times, the system must also incorporate a battery storage solution. The battery is required to supply 203.74 kWh to cover 70% of the total daily energy demand during these peak periods. To ensure that the system can maintain autonomy for one full day, the battery system must be designed with a capacity that considers both the Depth of

Discharge (DoD) of 40% and a battery efficiency of 85%. Taking these factors into account, the required battery capacity is calculated to be approximately 598.04 kWh. This capacity ensures that the battery can provide sufficient energy while also prolonging its lifespan by preventing excessive discharge.

Furthermore, the grid experiences a peak power load of 8.732 kW sustained over a duration of 10 hours. To accommodate this maximum load, the inverter must be designed with a capacity that not only meets peak demand, but also includes a safety margin to account for potential fluctuations and inefficiencies. Applying a safety margin of 25%, the required inverter capacity is determined to be 10.915 kW.

5.10.2 Feature Importance

The variables of total energy consumption kwh and wind cost exhibited the highest coefficients in the analysis, indicating their role as the main contributors to overall costs. This indicates that fluctuations in energy consumption and wind-related expenses have a substantial impact on the total cost structure. In contrast, the presence of negative coefficients, such as those associated with grid cost, suggests potential avenues for cost savings or modifications in allocation strategies. This implies that by optimizing grid costs, organizations may be able to enhance their financial efficiency and redirect resources toward more impactful areas. Table 5.1 provides a detailed explanation and interpretation of the feature coefficient analysis, underscoring the relative importance of each predictor regarding the target variable, total individual costs.

The connections between features and costs indicate potential inefficiencies or an excessive dependence on particular energy sources, such as wind and solar. This information can inform redistribution strategies or the implementation of hybrid usage.

Table 5.1 Feature Coefficients and Interpretations

Feature	**Coefficient**	**Description**
total energy consumption kwh	0.415169	Strong positive influence on total individual costs. Higher energy usage significantly increases costs.
Wind cost	0.405776	Significant positive impact. Allocation to wind energy increases total costs.
Solar cost	0.134445	Moderate positive influence. Solar energy contributes less to overall costs compared to wind.
Battery cost	0.006652	Minimal impact. Battery costs have a negligible effect on total costs.
Grid cost	-0.013325	Slight negative effect. Indicates grid energy may reduce overall costs (or share energy).
Total cost	0.182977	Generalized total individual costs correlate with total individual costs.
Extra Factors	Varying positive effects	Represents external influences or unmeasured factors affecting total costs.

In summary:

- **Focus on optimizing:** wind cost and solar cost as they are significant cost drivers.
- **Grid energy:** grid cost may serve as a cost-saving measure; further analysis could explore this trade-off.

5.10.3 Energy Usage Behavior

- The correlation of total energy consumption (kwh) with individual costs emphasizes the need to manage or induce efficient consumption.

5.10.4 Solar Energy Potential

- Despite its smaller cost impact, solar energy remains a viable allocation source due to its scalability and moderate expense profile.

5.10.5 Battery Cost Impact

- Minimal contribution to costs suggests current battery usage levels are either negligible or highly efficient.

5.10.6 Allocation of Generation Sources Analysis

The analysis shows that Consumer 4 has not received any energy allocation, suggesting either a lack of connections to renewable or storage resources or insufficient capacity to meet their energy needs. Meanwhile, the other consumers have been allocated energy, with some allocations exceeding normal demand levels, which could imply a risk of over-allocation in the algorithm.

The cost of the grid of ZAR 8.65 was calculated, which indicates the daily cost of electricity acquired from the grid in the South African Rand (ZAR), using the specified starting rate of ZAR2.19 per kWh. This means that: the total energy shortfall (demand not supplied by wind, solar or battery) was around 3.95 kWh for the day. The grid demand verification and energy breakdown are shown below:

5.10.7 Step 1: Consumer Allocations of Wind, Solar and Battery

Each consumer's energy allocation is determined by:

- Their connection to specific energy sources (wind, solar, battery).
- Their actual demand versus the remaining capacity of each source.

5.10.8 Step 2: Shortfall Calculation

For each consumer, the shortfall (5.45) can be calculated as follows:

$$\text{Shortfall} = \text{Demand} - (\text{Wind Allocation} + \text{Solar Allocation} + \text{Battery Allocation}) \tag{5.45}$$

The shortfall represents the energy that must be provided by the grid.

5.10.9 Step 3: Total Grid Energy Demand

To calculate the total grid energy demand, sum up all shortfalls (5.46):

$$\text{Total Grid Demand (kWh)} = \sum (\text{Shortfall of All Consumers}) \tag{5.46}$$

5.10.10 Step 4: Grid Cost Calculation

The grid cost (5.47)can be calculated by multiplying the total demand of the grid by the cost per kWh:

$$\text{Grid Cost} = \text{Total Grid Demand (kWh)} \times \text{Grid Cost per kWh (R2.19)} \tag{5.47}$$

Table 5.2 provides the detailed breakdown analysis of the energy shortfalls for each consumer.

Table 5.2 Consumer Energy Demand, Allocation, and Shortfall

The total grid energy demand of 3.995 kwh matches the unmet demand for consumer 4, so the grid energy demand comes exclusively from Consumer 4's unmet demand of 3.95 kWh, and the rest of the consumers' demands are met through allocated wind, solar, and battery resources, which exceed their respective demands.

5.10.11 Optimal Sizing of System Components

The dependence on the grid has been significantly reduced, resulting in a system that operates highly independently without the need for external electricity purchases.

With about 500 kWh of battery storage, the system effectively retains energy produced from

Consumer	Demand (kWh)	Allocation (kWh)	Shortfall (kWh)
Consumer 1	17.49	57.46	0.00
Consumer 2	85.06	124.52	0.00
Consumer 3	14.78	211.67	0.00
Consumer 4	3.95	0.00	3.95
Consumer 5	17.79	77.22	0.00
Consumer 6	7.71	116.18	0.00
Consumer 7	8.47	41.82	0.00
Consumer 8	73.84	153.06	0.00
Consumer 9	45.84	112.46	0.00
Consumer 10	19.29	195.62	0.00

renewable sources for utilization during periods of low generation. Approximately 116 kWh of wind energy plays a crucial role, particularly for nighttime generation when solar energy is not accessible, while solar energy contributes around 87 kWh shown in figure 5.2.

5.10.12 Inputs for LCOE Calculation

Capital Costs

- **Wind Turbines:** $R283,876$ (one-time cost for all turbines)
- **Solar Panels:** $R81,165$ (including installation)
- **Battery Storage:** $R218,598$ ($R5,465.95 \times 40$ batteries)

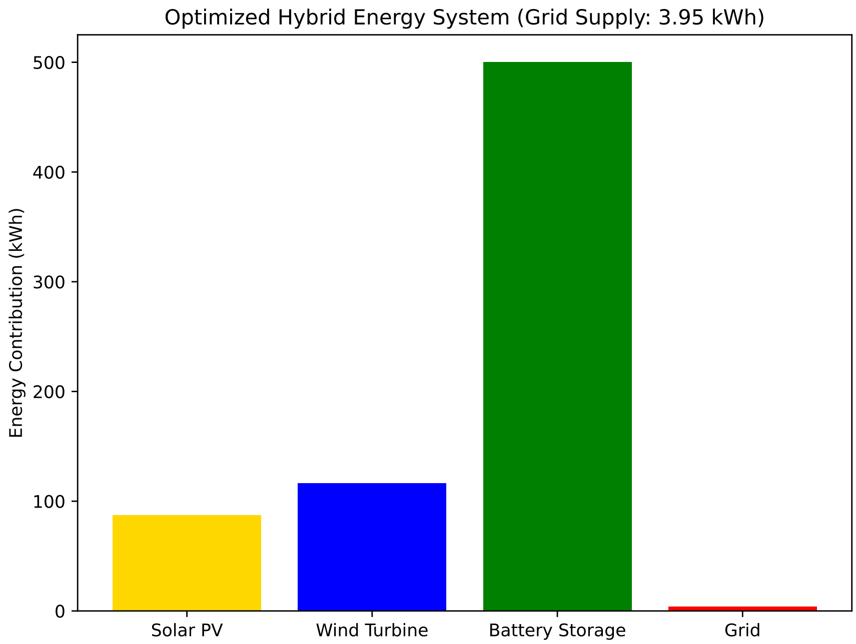

Figure 5.2 GA PSO Optimized Hybrid Energy Sytsem

Total Capital Costs:

$$\text{Total Capital Costs} = \text{R}283,876 + \text{R}81,165 + \text{R}218,598 = \text{R}583,639$$

O&M Costs

- **Solar:** R0 (as provided)
- **Wind:** R0 (assumed free energy)
- **Battery:** Negligible (assume R0 for simplicity)

Total O&M Costs:

$$\text{Total O\&M Costs} = \text{R}0$$

Lifetime Energy Output

- **Wind:**

$$116.42\,\text{kWh/day} \times 365\,\text{days/year} \times 20\,\text{years}$$

- **Solar:**

$$87.318\,\text{kWh/day} \times 365\,\text{days/year} \times 20\,\text{years}$$

- **Battery Efficiency Loss:** Accounted for during dispatch.

Total Energy: Sum of the wind and solar contributions.

5.11 Payback Time

The payback time is the time it takes for the savings from using renewable energy to equal the initial capital investment. It can be calculated as shown in equation 5.48:

$$\text{Payback Time} = \frac{\text{Total Capital Costs}}{\text{Annual Savings from Grid Cost Reduction}} \tag{5.48}$$

5.11.1 Inputs for Payback Time Calculation

Annual Savings from Grid Cost Reduction

The savings depend on how much energy demand is met from renewables rather than from the grid. It can be calculated using the total energy reduction of the grid and the cost of the grid per kWh (R2.19):

$$\text{Annual Savings} = \text{Total Grid Energy Reduction (kWh)} \times \text{Grid Cost per kWh (R2.19)}$$

The reliance of the grid was calculated using the Equation in(5.49) :

$$\text{Grid Reliance (\%)} = \left(\frac{\text{Grid Energy Supplied}}{\text{Total Energy Demand}}\right) \times 100 \tag{5.49}$$

Where:

$$\text{Grid Energy Supplied} = 3.95 \text{ kWh}$$

$$\text{Total RE Energy Production} = 87 + 116.42 = 203.42 \text{ kWh}$$

$$\text{Grid Reliance} = \left(\frac{3.95}{203.42}\right) \times 100 \approx 1.94\%$$

Thus, the grid supports approximately 1.94% of the total demand. This is a low-reliance signal, signaling a nearly self-sufficient hybrid system. in Equation (5.50), surplus energy is calculated: The formula for Surplus Energy (%) is given by:

$$\text{Surplus Energy (\%)} = \left(\frac{\text{Total Generation} - \text{Total Demand}}{\text{Total Demand}}\right) \times 100 \tag{5.50}$$

From the system, we calculate Total Generation as follows:

$$\text{Total Generation} = \text{Solar} + \text{Wind} + \text{Battery Storage} = 87 + 116.42 + 598 = 801.42 \text{ kWh}$$

The Total Energy Demand is:

$$\text{Total Energy Demand} = 203.42 \text{ kWh}$$

Now, substituting the values into the Surplus Energy is calculated as:

$$\text{Surplus Energy (\%)} = \left(\frac{801.42 - 203.42}{203.42}\right) \times 100 \approx 293.83\%$$

5.11.2 Summary of Results from GA-PSO Energy Allocation Optimization

All consumers, except Consumer 4, obtain adequate energy from the solar, wind and battery storage sources depicted in the energy distribution analysis shown in table 5.3,illustrates the energy distribution achieved through the GA-PSO method.

.

The unmet demand of Consumer 4, which stands at 3.95 kWh, aligns closely with the grid energy requirement of 3.995 kWh. This indicates that the grid dependency is minimized and the system operates in a largely isolated manner. Table 5.4 shows the hybrid system optimized by the GA-PSO hybrid optimization model.

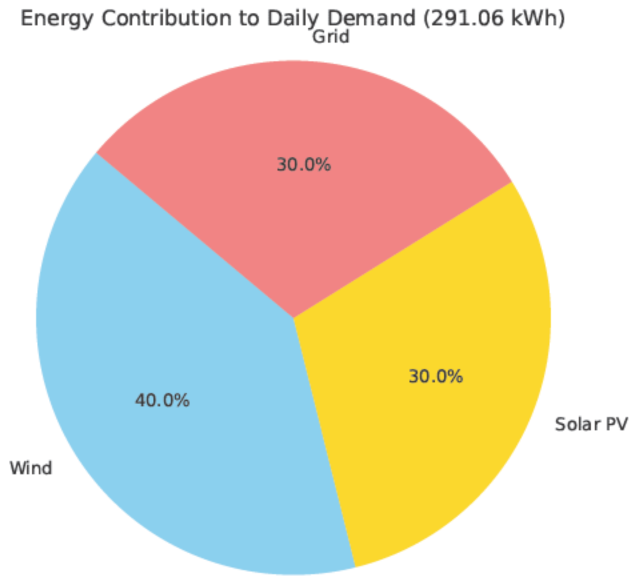

Figure 5.3 GA-PSO Energy Distribution

While the system is nearing complete independence from the grid, a minor deficit persists that requires further optimization. The system produces 476.58 kWh more than necessary, resulting in wasted excess energy unless it is effectively harnessed.

This situation supports the implementation of Linear Programming (LP) for the distribution of surplus energy, ensuring that this excess is efficiently allocated to other consumers or stored for future use.

Table 5.3 Energy Source Contributions

Energy Source	Contribution (kWh)	Purpose
Solar PV	80.00	Supplies energy during daylight hours
Wind Energy	100.00	Provides energy, especially at night
Battery Storage	500.00	Stores excess solar/wind energy for later use
Grid Energy Demand	3.995	Exclusively needed for Consumer 4
Total Demand Covered	**Fully met for all except Consumer 4**	**Hybrid system is effective including**
Surplus Energy	**476.58 kWh**	**Needed to be used/shared effectively**

Table 5.4 Summary of Hybrid Energy System Performance and Economic Metrics

Metric	**Value**	**Explanation**
Total Energy Demand	203.42 kWh/day	Total daily energy required from solar, wind, battery, and grid sources.
Solar Energy Production	87 kWh/day	Energy produced daily by the solar photovoltaic system.
Wind Energy Production	116.42 kWh/day	Daily energy contribution from the wind turbine system.
Battery Storage Capacity	598 kWh	Maximum energy the battery system can store and discharge.
Grid Energy Supplied	3.95 kWh/day	Remaining energy drawn from the grid to meet the shortfall.
Grid Reliance (%)	1.94%	Percentage of total demand sourced from the grid, minimizing reliance.
Surplus Energy (%)	293.83%	Extra energy generated beyond the demand, available for storage or export.
Wind Lifetime Energy Output	849,866 kWh	Total energy output from the wind turbine over 15 years.
Solar Lifetime Energy Output	637,421.4 kWh	Total energy generated by the solar system over its operational lifetime.
Total Lifetime Energy Output	1,487,287.4 kWh	Combined lifetime energy production from solar and wind systems.
Levelized Cost of Energy (LCOE)	R 0.392/kWh	Average cost per unit of energy produced over the system's entire lifespan.
Annual Savings from Grid Reduction	R 162,857.97/year	Yearly financial savings from reduced grid reliance, based on R2.19/kWh grid cost.
Payback Time	3.58 years	Time required for the system to recover its initial cost through grid savings.

5.12 Conclusion

The application of the Genetic Algorithm - Particle Swarm Optimization (GA-PSO) method facilitated the determination of the optimal sizing for renewable energy components and storage capacity, leading to the following outcomes:

- Solar Panels: 87.318 kW
- Wind Turbines: 116.428 kW
- Grid Connection: 87.318 kW
- Battery Storage: 600 kWh (with an inverter limit in kW)

This configuration was developed to ensure that the system could accommodate peak power requirements without risking equipment overload, thereby optimizing both renewable energy generation and battery usage. The GA-PSO method successfully achieved a balance between generation and storage capacities while addressing nonlinear and intricate constraints.

The application of GA-PSO for the optimal sizing and distribution of renewable energy sources has demonstrated its effectiveness in reducing costs and enhancing reliability within hybrid energy systems. Nevertheless, the intricate nature of real-world limitations, including fluctuating electricity prices, battery wear, and the management of excess energy, can be further refined through the use of Linear Programming (LP). LP models have the potential to improve the existing methodology by accurately distributing resources while adhering to various constraints, facilitating economical energy allocation among storage systems, the utility grid, and neighboring microgrids. Therefore, LP can act as a supplementary optimization instrument in the future developments of this research with:

- **Resource Allocation Optimization:** Resource Allocation Optimization: LP effectively reduces costs while maximizing the utility of surplus energy in storage systems, the grid, or adjacent microgrids.
- **Constraints Management:** LP adeptly addresses constraints such as battery degradation, peak load periods, and variations in grid energy pricing.
- **Efficiency Enhancement:** LP optimizes the charging and discharging cycles of storage systems to minimize losses, thus extending the battery lifetime.

It is important to acknowledge that the GA–PSO algorithm is computationally intensive, particularly for large-scale implementations with many consumers and multiple renewable sources. This complexity may pose challenges in rural settings with limited computational infrastructure or intermittent connectivity. To mitigate this, lightweight optimization approaches or edge-computing solutions could be integrated in future work to enable real-time or near-real-time decision-making without relying heavily on centralized high-performance computing resources.

Chapter 6

Optimal Energy Sharing Amongst Households

6.1 Introduction

In the preceding chapter, the GA-PSO hybrid technique was utilized to identify the optimal capacities of the essential elements of a renewable energy system, including solar panels, wind turbines, grid connections, and battery storage. The optimized configuration specified solar panels at 87.318 kW, wind turbines at 116.428 kW, grid connections at 87.318 kW, and battery storage at 600 kWh, all while complying with inverter limitations. This methodical approach ensured that the system could accommodate peak power requirements without overloading any individual components, thus improving reliability and overall system performance.In this study, surplus is defined as excess renewable generation not consumed immediately and not stored.

Building upon these findings, this chapter shifts to linear programming (LP) optimization as a standard for assessing performance. The emphasis transitions from determining capacities to refining energy-sharing strategies among consumers. In particular, when battery storage reaches its capacity and there is an excess of renewable energy, LP, augmented by game theory principles, will facilitate the redistribution of this surplus energy to consumers with immediate energy demands. This strategy seeks to reduce the reliance on the grid, promote energy equity, and promote environmental sustainability as follows.

- **Decrease Dependence on the Grid:** By optimizing the use of locally generated renewable energy, the system aims to reduce the dependence on the grid electricity. This not only contributes to energy independence, but also minimizes the costs associated with grid energy consumption.
- **Maximize Utilization of Solar and Wind Resources:** The optimization will focus on harnessing the maximum potential of solar and wind energy available throughout the day. By aligning energy production with consumption patterns, the system can ensure that renewable resources are utilized to their fullest extent, thereby reducing waste and increasing sustainability.
- **Store Surplus Energy in the Battery for Future Utilization:** The optimization process will also prioritize the effective management of surplus energy generated during peak production times. By directing excess energy into battery storage, the system can ensure that this energy is available for use during periods of low production or high demand, thus improving the reliability of the energy supply.
- **Ensure Energy Delivery to Consumers Remains Within Inverter Constraints :** A critical aspect of the optimization will be to ensure that the energy delivered to consumers does not

exceed the limitations set by the inverter's AC power capacity. This constraint is vital for maintaining the integrity of the system and preventing potential damage to the inverter or other components.

6.2 Methodology Based on Linear Optimization (LP) Approach

Objective Function:

The objective of the LP model calculated in equation (6.1) is to reduce the energy consumption from the grid while satisfying the energy requirements of each consumer through the utilization of renewable sources, efficiently managing any excess energy produced from these sources, and incorporating battery storage:-

Minimize

$$Z = \sum_{i=1}^{n} G_i \tag{6.1}$$

Where:

- G_i = Grid energy allocated to consumer i
- n = Number of consumers (10 in this case)

Decision Variables: The energy sources are defined as variables allocated to each consumer:-

$$S_i, \quad W_i, \quad G_i, \quad B_i \geq 0$$

Where:

- S_i = Solar energy allocated to consumer i
- W_i = Wind energy allocated to consumer i
- G_i = Grid energy allocated to consumer i
- B_i = Battery energy allocated to consumer i

Problem Constrains:

- Demand Satisfaction Constraint:-The total allocated energy must fulfill the energy demand of each consumer:

  $$S_i + W_i + G_i + B_i \geq D_i, \quad \forall i \in \{1, \dots, n\}$$

  Where:

  - S_i = Solar energy allocated to consumer i
  - W_i = Wind energy allocated to consumer i
  - G_i = Grid energy allocated to consumer i
  - B_i = Battery energy allocated to consumer i
  - D_i = Energy demand of consumer i

- Allocation of Maximum Renewable Contraints: The distribution of solar and wind energy resources is based on their availability:

$$S_i \leq S_{\max}, \quad W_i \leq W_{\max}, \quad \forall i$$

Where:

  - S_i = Solar energy allocated to consumer i
  - $S_{\max}$ = Maximum solar energy available
  - W_i = Wind energy allocated to consumer i
  - $W_{\max}$ = Maximum wind energy available

- Grid Usage Constraint: The utilization of grid energy occurs solely when renewable and battery resources are inadequate.

$$G_i \geq 0, \quad \forall i$$

Where:

  - G_i = Grid energy allocated to the consumer i

- Battery Availability Constraint: The capacity of the battery energy in (6.2) is constrained by the quantity of energy it can store:

$$B_i \leq B_{\text{total}}, \quad \forall i \tag{6.2}$$

Where:

  - B_i = Battery energy allocated to consumer i

  - B_{total} = Total battery energy available

- Calculation of Surplus Energy:- Consumer surplus is defined as the variation between the energy assigned and the corresponding demand for that energy calculated in equation (6.3):

$$\text{Surplus}_i = (S_i + W_i + G_i + B_i) - D_i, \quad \forall i \tag{6.3}$$

  Where:

  - Surplus_i = Surplus energy for consumer i
  - S_i = Solar energy allocated to consumer i
  - W_i = Wind energy allocated to consumer i
  - G_i = Grid energy allocated to consumer i
  - B_i = Battery energy allocated to consumer i
  - D_i = Energy demand of consumer i

- Other Contraints: It is required that all energy variables maintain nonnegative values, such as:

$$S_i, \quad W_i, \quad G_i, \quad B_i \geq 0, \quad \forall i$$

  Where:

  - S_i = Solar energy allocated to consumer i
  - W_i = Wind energy allocated to consumer i
  - G_i = Grid energy allocated to consumer i
  - B_i = Battery energy allocated to consumer i

6.3 Results Obtained from LP Model Approach

The analysis of the network graph shown in figure 6.1, reveals the distribution of surplus energy, which is the excess energy produced after meeting consumer demand, as well as the allocation of battery storage among different consumers. A detailed examination follows:-

Energy Distribution:- Each consumer has a specific energy requirement that is satisfied through a mix of Solar, Wind, Grid, and Battery resources. The surplus denotes the energy produced in excess of consumer demand, which remains unutilized at that moment. The battery serves to retain energy that is surplus or in excess when the demand is less than the available supply. Both Surplus and Battery contribute to the energy supply for consumers in different amounts, depending on individual energy requirements and the capacity of battery storage. Figure 6.1 shows the network graph showing the flow of energy from solar, wind, grid, and battery to each of the consumers as follows:

Energy Flow Allocation:

- The surplus node distributes energy to each consumer, either partially or completely fulfilling their outstanding demand.
- The energy stored in the battery helps to address any remaining energy requirements following the surplus energy distribution. This is particularly important when the surplus is negative or insufficient to meet demand, with the battery acting as a supplementary source.
- For instance, Consumer 1 has a Final Surplus of 11.6164 kWh, and since the Battery contribution is zero, it is likely that their energy needs are entirely satisfied by the Surplus.
- Conversely, Consumer 2 exhibits a negative surplus, indicating that their consumption has exceeded generation, thus the battery is likely essential in fulfilling their additional demand.
- The connections illustrated in the graph depict the energy flow among the surplus, battery, and consumer, which are optimized to reduce energy waste and enhance overall utilization.

In figure 6.2, an optimal energy transfer bar chart is shown.The bar chart presents the final surplus energy for each consumer following the adjustments made for allocation and storage. Con-

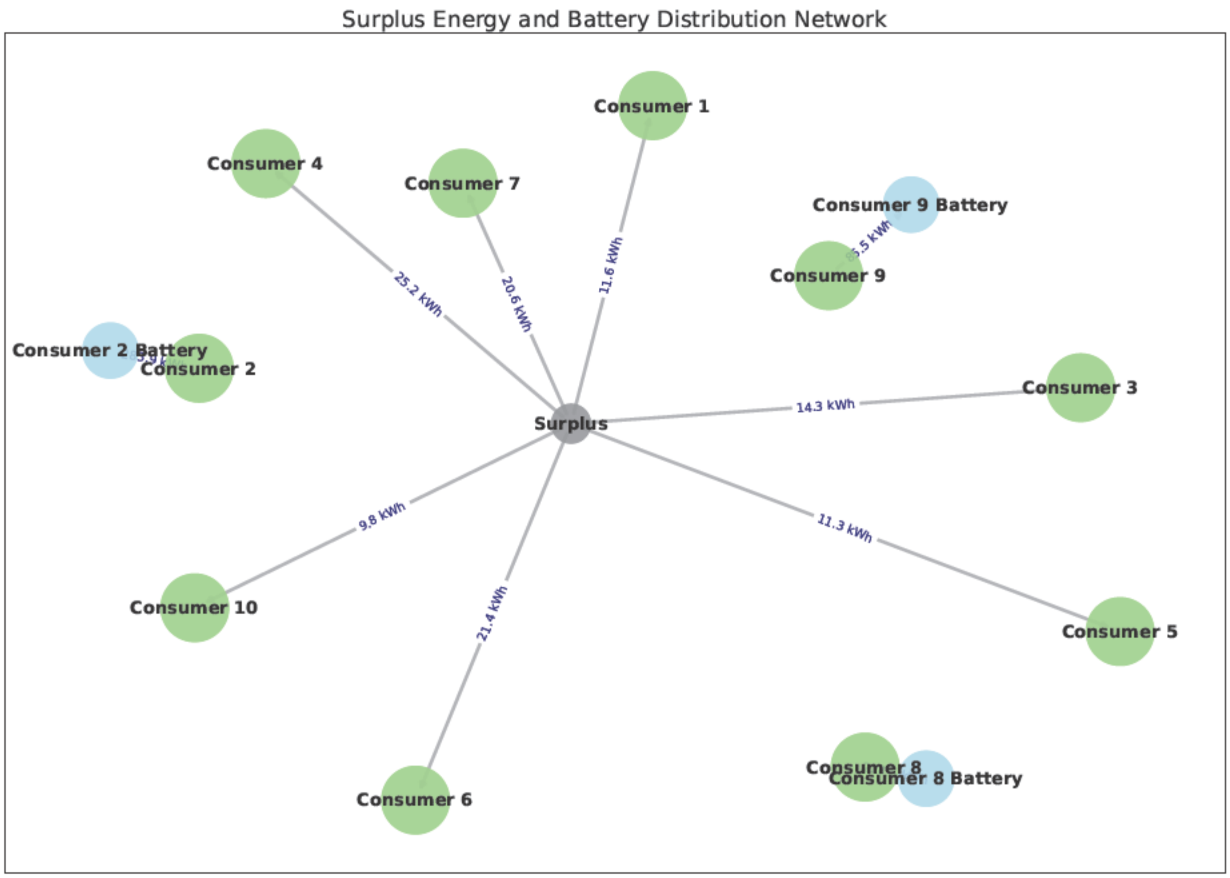

Figure 6.1 Energy Sharing Network Graph showing energy allocations from Battery, Surplus, to Consumers (LP).

sumers 3, 4, 6, and 7 exhibit the highest final surpluses, indicating that they used less energy or were supplied more energy than they required. In contrast, Consumers 2, 8, and 9, although they received energy from the battery, still report lower surpluses than their counterparts, indicating their higher energy consumption or demand.

Figure 6.3 shows the heatmap Tranfer Matrix, which represents the allocation of energy from Solar, Wind, Grid, and Battery sources to each consumer. The contributions of Solar, Wind, and Grid were uniformly distributed among all consumers, ensuring an equitable energy supply. In contrast, battery energy was exclusively used by consumers 2, 8, and 9, indicating that the system

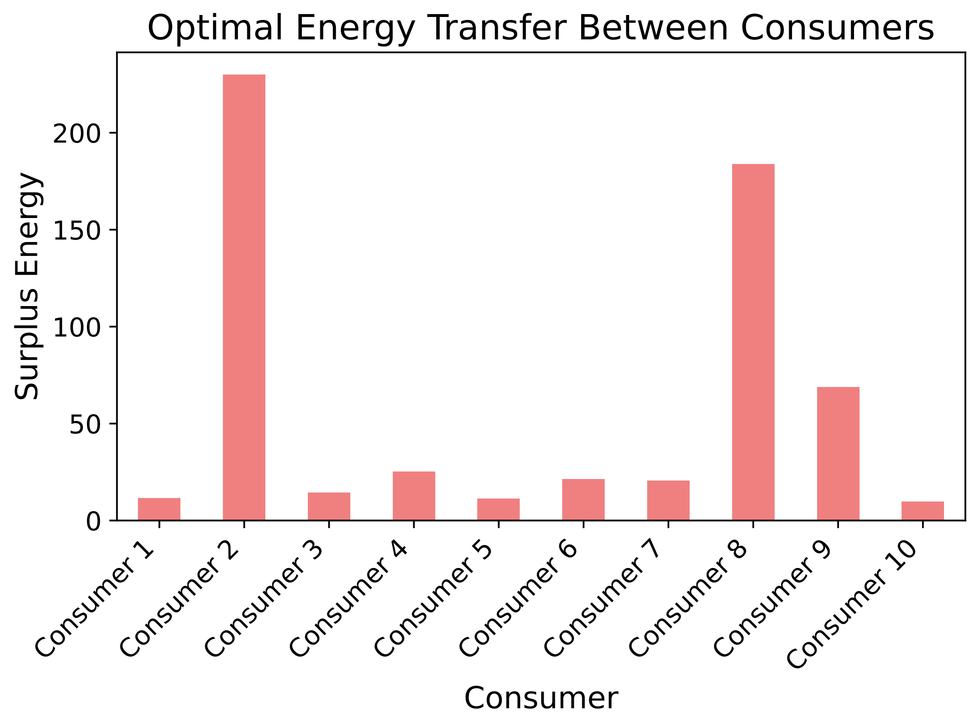

Figure 6.2 LP Optimal Energy Transfer (LP)

prioritized energy storage for those with high-energy deficiencies.

In figure 6.4 shows the overall energy distribution with the demand, total allocated, and final surplus. The majority of energy was designated to satisfy consumer requirements, with a lesser fraction directed toward surplus generation. This reflects an effective allocation strategy that reduces excess energy while adequately addressing demand. The proportions of final surplus and battery storage are compared in the pie chart shown in figure 6.5, where the mean battery per consumer is depicted in figure 6.6

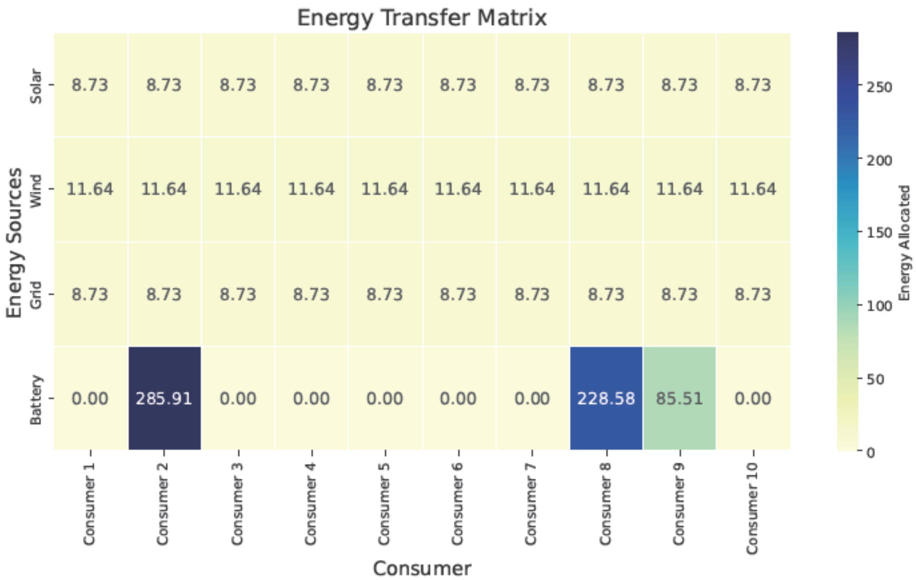

Figure 6.3 Heatmap Transfer Matrix (LP)

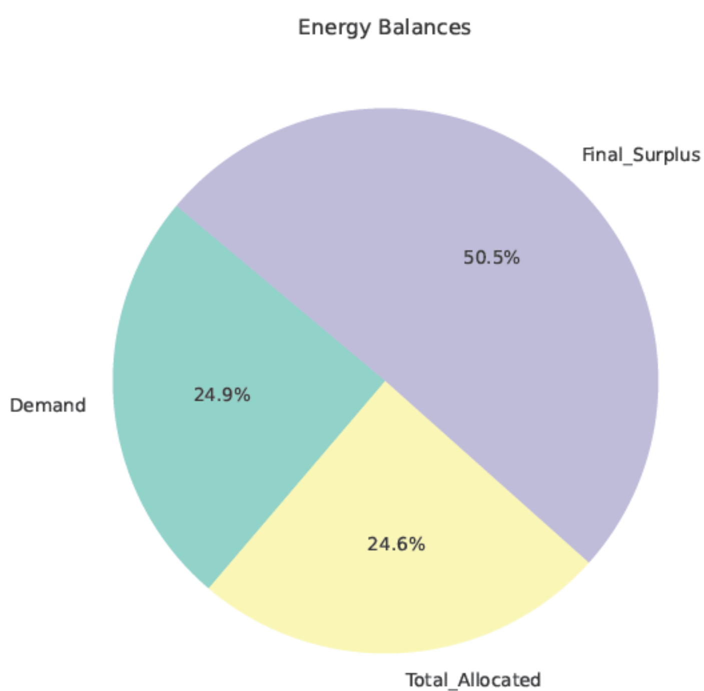

Figure 6.4 Energy Balance Pie Chart (LP)

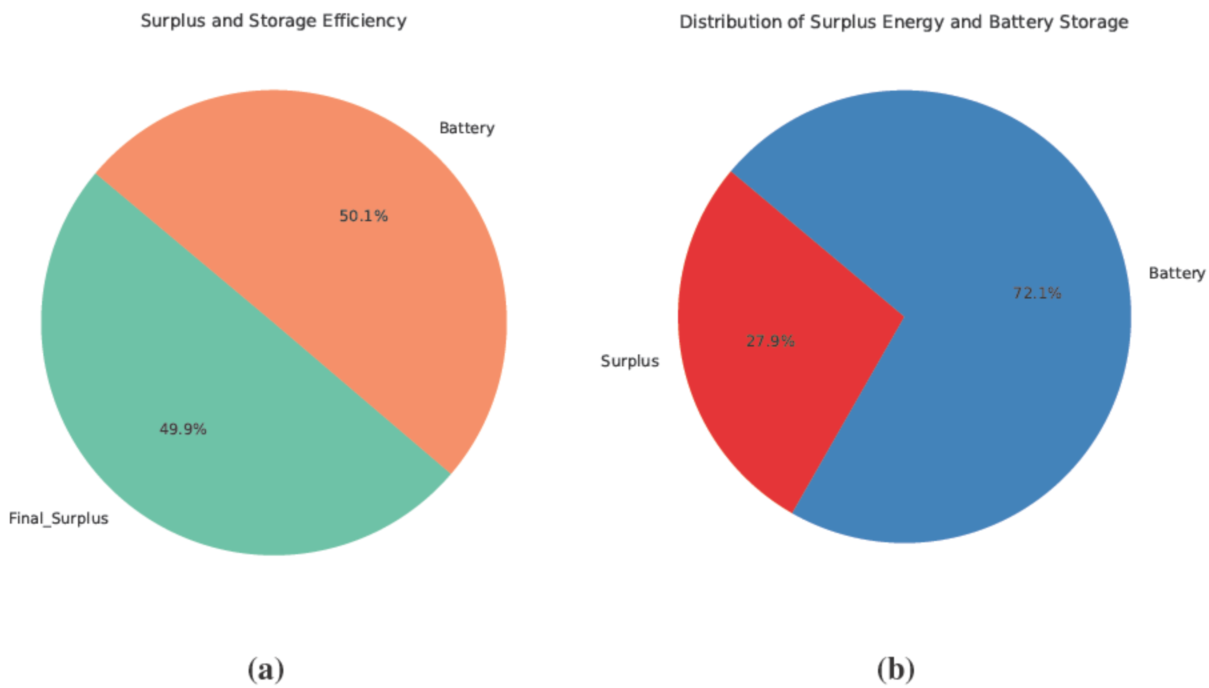

Figure 6.5 (a) Surplus Storage Efficiency (LP) and (b) Distribution of Surplus Energy and Battery Storage (LP)

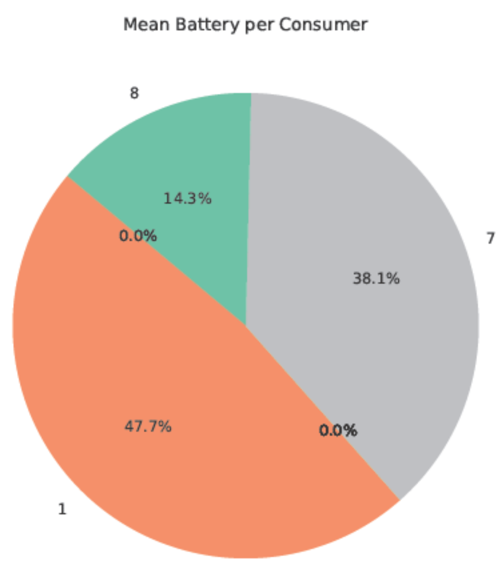

Figure 6.6 Pie Chart Mean Battery Per Consumer (LP)

6.3.1 Summary Of LP Optimization Model

Linear Programming (LP) is limited because it assumes centralized control, which ignores each participant's independent actions aimed at maximizing their individual gains. In the context of modeling a microgrid, competition among users for finite solar or battery energy might result in strategic behaviors that LP fails to capture since it assumes all participants are aligned with a single goal, ignoring the competitive dynamics at play.For instance, in a scenario involving battery storage sharing, consumers may collaborate to utilize a common battery, thereby lowering overall energy expenses. Conversely, if one consumer excessively draws energy, it results in reduced availability for others, potentially leading to conflicts that can be anticipated and addressed through game theory. This approach is particularly relevant given the unpredictability of renewable energy sources, such as solar and wind, which are influenced by varying weather conditions.

Furthermore, energy systems function in real time, with demand and supply constantly fluctuating. Game theory is adept at modeling multistage games, allowing players to modify their decisions based on prior results. In contrast, LP typically addresses a singular optimization problem and lacks the capacity to adapt to changing conditions. For example, in a smart grid scenario, consumers can modify their energy consumption in response to real-time pricing and battery status, creating a dynamic environment that game theory can effectively model.

In summary, while LP excels in centralized optimization with fixed parameters, game theory provides greater flexibility by accommodating decentralized decisions and making competitive and cooperative interactions. In the next section, game theory is explored.

6.4 Methodology Based on Game Theory Optimization (GT) Approach

The Game Theory model represents a significant advancement in energy distribution strategies when compared to traditional linear programming models. Using the principles of Game Theory, this approach fosters a collaborative environment among energy consumers, which leads to several key benefits.

Firstly, the model encourages energy sharing among participants. In a conventional linear programming framework, the energy distribution is often viewed as a one-way transaction, where consumers primarily depend on centralized energy sources. In contrast, the Game Theory model promotes a more interactive system where consumers can exchange energy with one another. The Game Theory framework encourages participants to strategically use their battery storage capabilities. By analyzing the energy consumption patterns and needs of various consumers, the model facilitates better coordination in battery usage. This ensures that energy is stored and released at optimal times, further enhancing the overall energy distribution process. This collaborative framework not only provides tangible benefits to consumers but also fosters a more sustainable energy framework.

Objective Function

The primary objective is to improve overall energy efficiency by redistributing surplus energy and optimizing battery utilization among consumers, as shown in the equation (6.4).

Maximize

$$U = \sum_{i=1}^{n} (S_i + B_i - C_i) \tag{6.4}$$

Where:

- S_i = Surplus energy received by consumer i
- B_i = Battery energy allocated to consumer i
- C_i = Cost of obtaining additional energy from the grid or losing surplus energy

Strategies of Players Here, players aim to maximize their returns while taking into account the consequences of their choices on other individuals where each consumer acts as a player in the game with the following strategies:-

- Surplus providers choose how much surplus energy to share
- Surplus Receivers decide how much surplus energy to accept based on demand while the battery users optimize battery discharge to minimize the impact of their decision on others.

The reward function (6.5) for each consumer is defined as follows:

$$\pi_i = (S_i + B_i) - C_i - \beta_i \tag{6.5}$$

Where:

- π_i = Payoff for consumer i

- C_i = Cost of grid energy used by consumer i
- β_i = Penalty for failing to meet energy demand or wasting surplus energy

This ensures that the consumers are motivated to utilize renewable energy and any surplus energy as a priority, reducing reliance on the grid and optimizing the efficiency of the system.

6.4.1 Problem Constraints

- **Demand Constraint**

  The demand constraint is calculated in (6.6) where every consumer is required to meet their energy needs utilizing surplus energy, battery storage, and a minimal amount of grid energy as:

$$S_i + B_i + G_i \geq D_i, \quad \forall i \in \{1, \ldots, n\} \tag{6.6}$$

  Where:

  - S_i = Surplus energy received by consumer i
  - B_i = Battery energy allocated to consumer i
  - G_i = Grid energy used by consumer i
  - D_i = Energy demand of consumer i

- **Redistribution of Surplus Energy Constraint:** The redistributed total surplus energy must not exceed the collective surplus.:

$$\sum_{i=1}^{n} S_i \leq \sum_{j=1}^{n} \text{Surplus}_j$$

  Where:

- S_i = Surplus energy received by consumer i
- Surplus_j = Total surplus energy available from all sources for consumer j

- **Battery Energy Constraint :** The discharge of a battery is constrained by the amount of energy it has stored:

$$B_i \leq B_{total}, \quad \forall i$$

Where:

- B_i = Battery energy allocated to consumer i
- B_{total} = Total available battery energy

- **Utilization of Grid Constraint:** The use of grid energy occurs exclusively when renewable and battery energy supplies are inadequate.

$$G_i \geq 0, \quad \forall i$$

Where:

- G_i = Grid energy used by the consumer i

- **Non-Negetivity Constraint :** It is required that all energy variables adhere to the non-negativity constraint, ensuring that they remain non-negative

$$S_i, B_i, G_i \geq 0, \quad \forall i$$

Where:

- S_i = Surplus energy received by consumer i

- B_i = Battery energy allocated to consumer i
- G_i = Grid energy used by consumer i

- **The Concept of Nash Equilibrium :** In this energy allocation game, Nash Equilibrium is defined as a state where no consumer can increase their payoff through a one-sided modification of their energy allocation strategy as :

$$\pi_i(S_i, B_i, G_i) \geq \pi_i(S_i', B_i', G_i'), \quad \forall i$$

Where:

- (S_i, B_i, G_i) = Optimal strategy profile for consumer i
- (S_i', B_i', G_i') = Alternative strategy profile for consumer i

At equilibrium, the surplus for each consumer and the utilization of batteries are optimized, thereby reducing the reliance on the grid and enhancing overall efficiency.

6.5 Cost Analysis Comparison Between LP and GT

The Levelized Cost of Energy (LCOE) for LP is calculated as (6.7):

$$\mathrm{LCOE}_{LP} = \frac{\text{Total Investment Cost} + \text{O\&M Cost}}{\text{Total Energy Produced Over Lifetime}} \tag{6.7}$$

Total Investment Costs:

The total investment is calculated in equation (6.8).

$$\text{Total Investment Cost} = \text{Solar} + \text{Wind} + \text{Battery} \tag{6.8}$$

$$= 81,165 + 283,876 + 21,859.80 = 386,900.80\,\text{ZAR}$$

Operation & Maintenance (O&M) Cost:

$$\text{O\&M} = 0.02 \times 386,900.80 \times 20 = 154,760.32\,\text{ZAR}$$

Total Energy Produced Over 20 Years:

$$\text{Total Energy} = (87 + 116.42) \times 365 \times 20 = 1,485,892\,\text{kWh}$$

LCOE for LP:

$$\text{LCOE}_{LP} = \frac{386,900.80 + 154,760.32}{1,485,892} = 0.3644\,\text{ZAR/kWh}$$

The Levelized Cost of Energy (LCOE) for GT is calculated as:

$$\text{LCOE}_{GT} = \frac{\text{Total Energy Produced}}{\text{Total Investment Cost} + \text{O\&M Cost}}$$

$$\text{LCOE}_{GT} = \frac{1,485,892}{386,900.80 + 154,760.32}$$

$$\text{Total Cost} = 386,900.80 + 154,760.32 = 541,661.12\,\text{ZAR}$$

$$\text{LCOE}_{GT} = \frac{1,485,892}{541,661.12}$$

Finally, calculating the LCOE for GT is given as:

$$\text{LCOE}_{GT} = 0.3648\,\text{ZAR/kWh}$$

LP Grid Price Calculation

The LP grid price is calculated as:

$$P_{LP} = (2.19 \times 0.2) + (0.995 \times 0.5) + (0.54 \times 0.3)$$

$$P_{LP} = 0.438 + 0.4975 + 0.162$$
$$= 1.0975\,\text{ZAR/kWh}$$

GT Grid Price Calculation

The GT grid price:

$$P_{GT} = (1.07 \times 0.2) + (0.7374 \times 0.5) + (0.41 \times 0.3)$$

$$\begin{aligned} P_{GT} &= 0.214 + 0.3687 + 0.123 \\ &= 0.7057\,\text{ZAR/kWh} \end{aligned}$$

Energy Cost Calculations

Using the formulas:

Total LP Cost

The total LP energy cost is calculated in equation (6.9).

$$\text{Total LP Cost} = \sum(\text{Grid}_{LP} \times P_{LP}) \tag{6.9}$$

$$\begin{aligned} \text{Total LP Cost} &= 16.46 + 87.80 + 13.17 + 3.29 + 17.56 + 6.59 + 7.68 + 76.83 + 47.20 + 18.66 \\ &= 295.24\,\text{ZAR} \end{aligned}$$

Total GT Cost

The total cost calculated based on the GT model is given as the equation in (6.10).

$$\text{Total GT Cost} = \sum(\text{Grid}_{GT} \times P_{GT}) \tag{6.10}$$

Calculating the total GT cost:

$$\begin{aligned}\text{Total GT Cost} &= 7.06 + 42.34 + 5.65 + 1.41 + 8.47 + 3.53 + 3.53 + 35.29 + 24.70 + 8.47 \\ &= 140.45\,\text{ZAR}\end{aligned}$$

Cost Savings from Grid Reduction and Surplus Energy Utilization

Grid Savings

The equation in (6.11) shows the savings in energy purchases from the grid, while in (6.12) is the energy saved by effectively utilizing surplus energy.

$$\text{Grid Savings} = \sum(\text{Grid}_{LP} - \text{Grid}_{GT}) \times P_{LP} = 154.79\,\text{ZAR} \tag{6.11}$$

Surplus Utilization Savings

$$\text{Surplus Utilization Savings} = \sum(\text{Surplus Used}_{GT} - \text{Surplus Used}_{LP}) \times P_{GT} \tag{6.12}$$

Calculating total surplus savings:

$$\begin{aligned}\text{Total Surplus Savings} &= 3.53 + 14.11 + 4.24 + 5.65 + 3.53 + 2.83 + 2.83 + 14.11 + 10.59 + 3.53 \\ &= 64.95\,\text{ZAR}\end{aligned}$$

Total Cost Savings (Grid + Surplus)

$$\text{Total Cost Savings} = 154.79 + 64.95 = 219.74\,\text{ZAR}$$

6.6 Results Obtained from Game Theory Optimization Model

Game theory facilitates improved local energy exchanges, decreasing reliance on the grid,resulting in cost savings. Figure 6.7 shows the energy-allocated network where each consumer node is sized based on their final surplus and battery storage,representing their energy advantage. The larger nodes indicate consumers who benefit more from energy redistribution, such as Consumer 2 and Consumer 8. Consumers with surplus energy can directly share it with others. The energy of the battery is selectively allocated to specific consumers, which helps to maintain the stability of the system. Regarding the distribution of energy, game theory enhances the network by effectively distributing surplus energy, thereby increasing resilience. Figure 6.8 illustrates the distribution of energy among consumers. In this scenario, consumers with a significant final surplus, such as Consumer 2 and Consumer 8, provide energy to those experiencing a deficit.

In Summary, the Game Theory model represents a notable improvement over conventional

Linear Programming (LP) techniques in the optimization of energy distribution within microgrids. In contrast to LP, which generally manages resource allocation in a centralized manner with limited adaptability, Game Theory promotes a decentralized and collaborative approach among consumers. In the LP framework, energy distribution tends to be inflexible, resulting in: a heightened reliance on grid energy for the majority of consumers; restricted adaptability to fluctuating consumer demands, particularly for those with substantial energy requirements; and inadequate sharing of surplus energy, leading to suboptimal use of local renewable resources. Conversely, the Game Theory-based model encourages active energy sharing among users. It motivates prosumers (consumers who generate or store energy) to work together and exchange surplus energy, diminishes dependence on external sources such as the grid by enhancing local energy equilibrium, and optimizes battery utilization throughout the network by examining usage patterns and coordinating charging and discharging strategies. This collaborative interaction guarantees that energy is allocated where it is most required, minimizes waste, and alleviates peak demand on the grid. As a result, Game Theory contributes to a more robust, economically viable, and sustainable energy ecosystem.

In Figure (6.9), the final surplus per consumer is illustrated in a bar chart after the energy relocation. Remarkably, those consumers who had previously experienced energy deficiencies (Consumers 2, 8, and 9) are now demonstrating a considerable positive surplus attributed to battery energy.The advantages for Consumers 2 and 8 are substantial, with final surpluses recorded at 229.96 kWh and 183.85 kWh, respectively. In addition, consumers who previously lacked any surplus now have adequate energy to meet their requirements, facilitated by the cooperative redistribution process.

The pie chart represented in figure 6.10, illustrates the relationships the efficient energy sharing between the battery and the surplus energy based on the Game theory model. The results shows

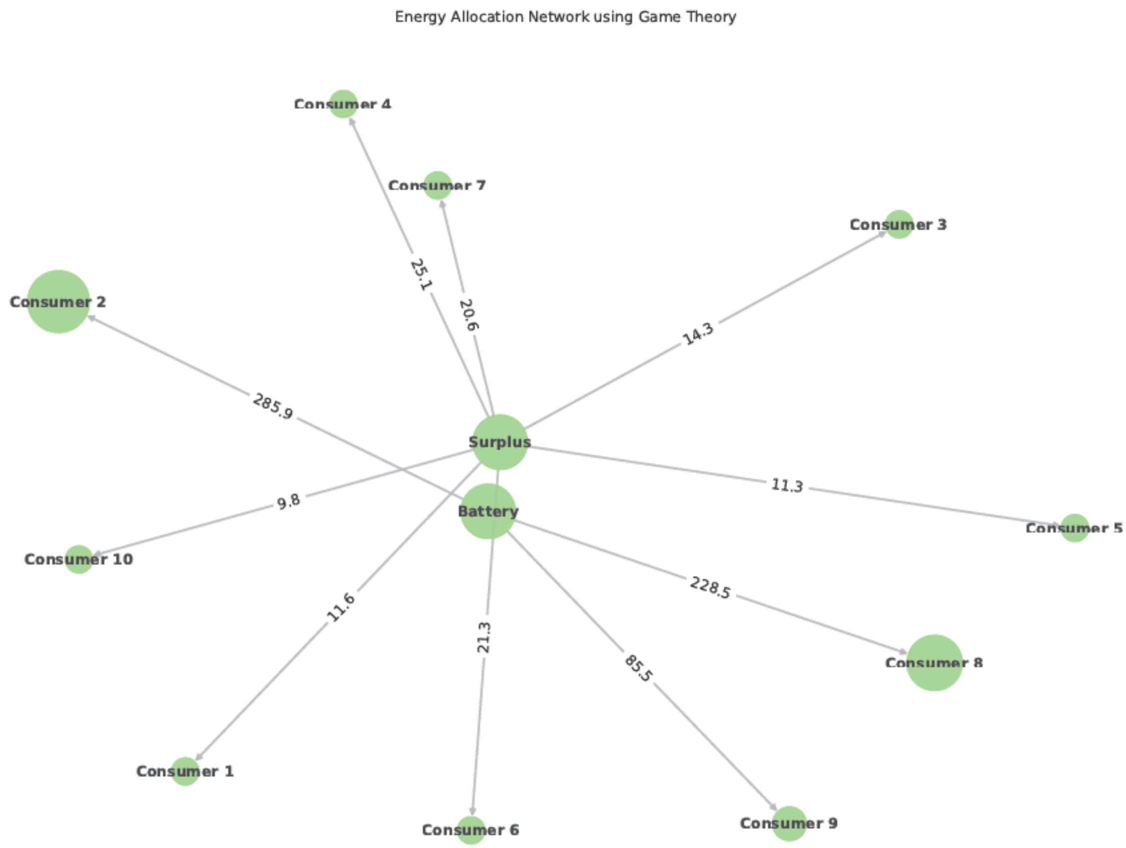

Figure 6.7 Energy Sharing Network Graph (GT).

a low battery efficiency (20.6%). The GT model might be more cost-effective despite its low battery efficiency because it has a large amount of remaining energy (79.4%). This suggests that the system stores energy efficiently for future use, which could reduce operational costs in the long term.

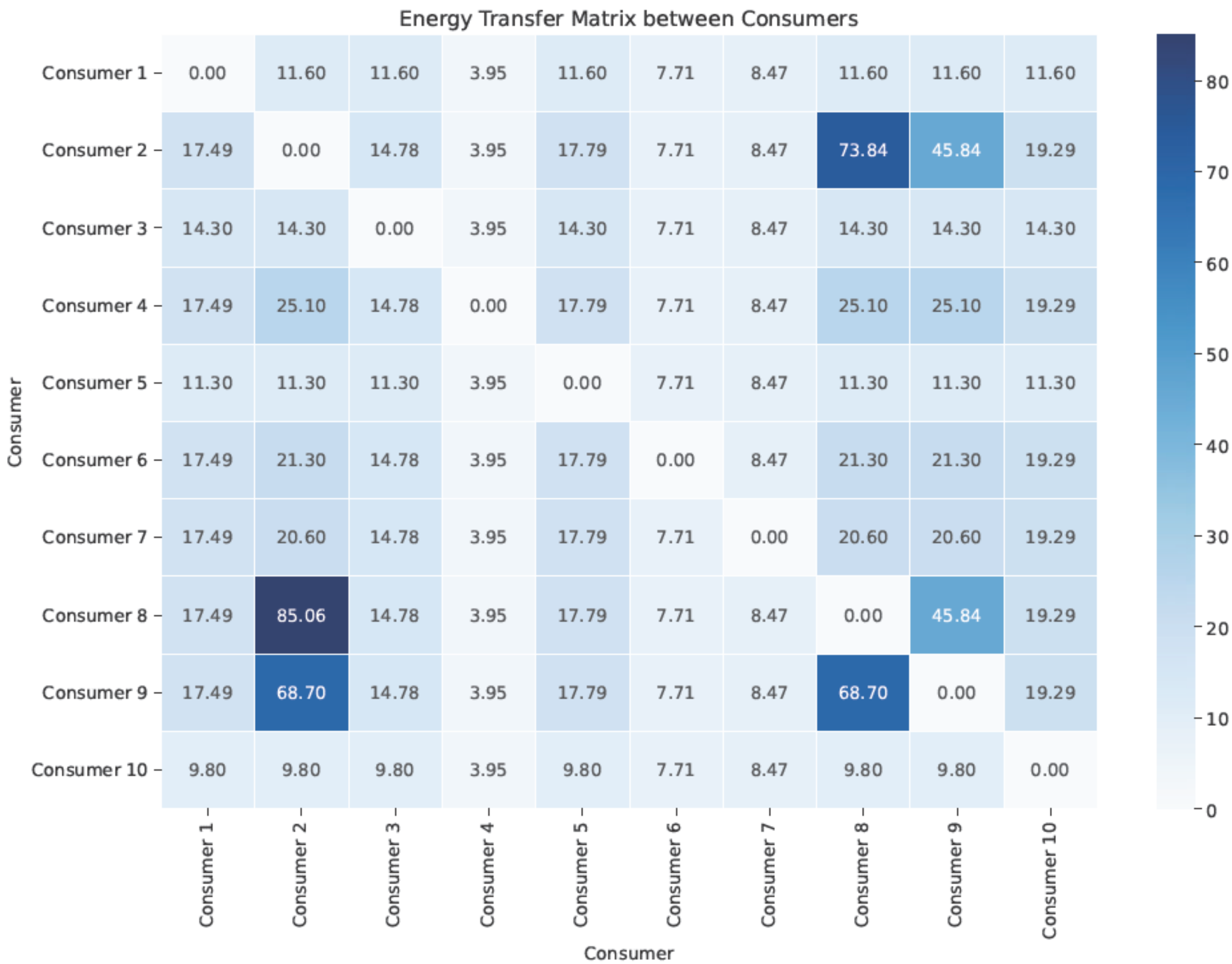

Figure 6.8 Optimal Energy Transfer Per Consumer (GT).

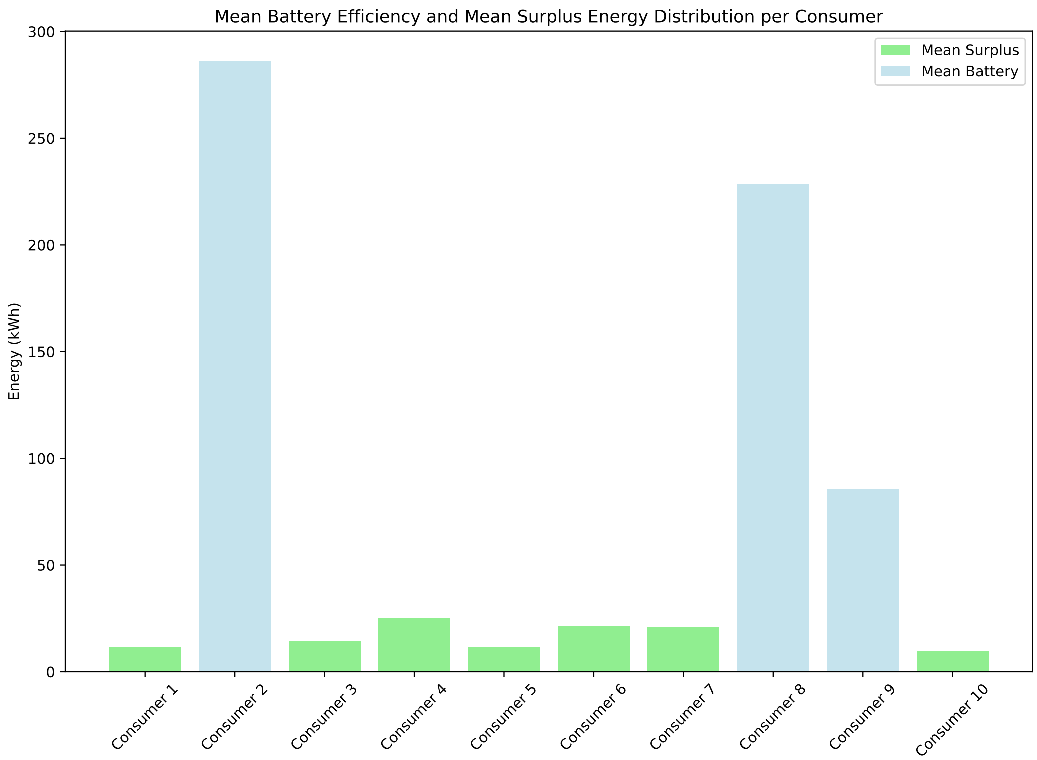

Figure 6.9 Mean Battery and Surplus Distribution (GT).

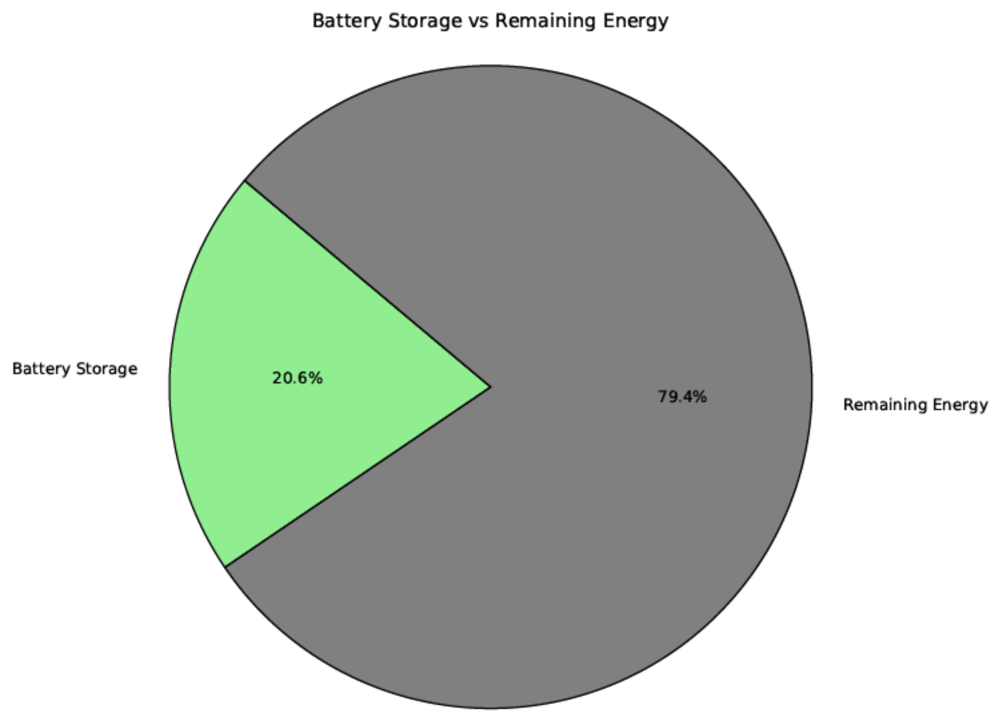

Figure 6.10 Battery Surplus Efficieny.

Costs Analysis

A summary of key findings, including LCOE calculations, grid price comparisons, and total cost savings between the LP and GT approaches, is illustrated in table 6.1. The total savings amount to of 219.74 ZAR is achieved when opting for GT over LP. This is achieved through the better utilization of surplus energy and the decreased dependence on the grid.

Game theory significantly reduces energy costs by decreasing grid expenses by 154.79 ZAR and generating savings of 64.95 ZAR from the use of surplus energy. This leads to a total savings of 219.74 ZAR. In summary, GT presents a more efficient and economical alternative to LP.

The Levelized Cost of Energy (LCOE) for Linear Programming (LP) is 0.3644 ZAR/kWh, while for Game Theory (GT) it is 0.3648 ZAR/kWh, indicating that the energy production costs for both methods are quite comparable. However, GT offers financial advantages through reduced expenses on grid purchases and better management of surplus energy, resulting in greater overall efficiency. Despite similar LCOE figures for LP and GT,GT is the most advantageous option as it contributes to lower overall energy costs and increased savings.

Table 6.1 Comparison of LP and GT Energy Economics

Description	**LP**	**GT**
LCOE (ZAR/kWh)	0.3644	0.3648
Weighted Average Grid Price (ZAR/kWh)	1.0975	0.7057
Total Energy Cost (10 Consumers, ZAR)	295.24	140.45
Surplus Energy Utilization Savings (ZAR)	–	64.95
Total Cost Savings (LP vs GT, ZAR)	–	154.79
Total Cost Savings (Grid + Surplus, ZAR)	–	219.74
Total Energy Produced (20 Years, kWh)	1,485,892	

6.7 Conclusion

This research presents multiple methods that improve the comprehension and advancement of energy management as well as the optimization of distributed generation systems.The first key contribution is the development of an integrated energy management framework that combines forecasting, optimized storage, and structured energy sharing. This holistic approach addresses key issues such as energy waste, grid dependency, and cost inefficiencies, making renewable energy systems more viable and reliable. The results demonstrate that this method significantly improves the efficiency of DG placement, cost efficiency, energy performance, and independence from the grid. The principal findings, along with their relevance to the study objectives, are outlined below:

- **Design of Hybrid System Utilizing HOMER Pro** Simulations conducted with HOMER Pro yielded a Levelized Cost of Energy (LCOE) of 0.9889 ZAR/kWh, which is considerably higher than the calculated values for Linear Programming (LP) and Game Theory (GT). This indicates a need for further optimization of system components and cost factors to achieve more economical energy generation.
- **Optimization of Component Sizing through GA-PSO**

  The study employs Genetic Algorithm-Particle Swarm Optimization (GA-PSO) to optimize the sizing of renewable energy components, including solar panels, wind turbines, and batteries. This optimization process ensures a harmonious balance between cost and performance, contributing to the creation of an efficient and economically sustainable hybrid system.
- **LP-Based Framework for Energy Sharing** A Linear Programming (LP) model was used to evaluate cost optimization and energy distribution. The findings demonstrates that GT facilitates a more cost-effective energy allocation.

- **Game Theory for Improved Energy Sharing** The application of Game Theory (GT) enhances the utilization of surplus energy, resulting in increased savings and a reduced dependence on the grid. The results revealed that GT-based energy sharing strategies yield superior economic benefits compared to LP-based strategies.

The analysis indicated that GT not only results in lower overall costs but also generates greater savings than LP, establishing it as a more effective approach.

Chapter 7

Conclusions and Recommendations

7.1 Conclusion

The results of this study indicate that the proposed energy sharing model not only improves energy efficiency, but also actively contributes to Sustainable Development Goal 7 (SDG7) by facilitating access to affordable, reliable, and sustainable energy for everyone.

By emphasizing the use of renewable energy sources, the model significantly decreases reliance on fossil fuels, thereby lowering carbon emissions and fostering a more sustainable and resilient energy environment. This strategy is in line with the SDG7 objective of improving infrastructure and advancing technology for sustainable energy in developing areas, such as South Africa.

Key conclusions related to the principles of SDG7 include:

1. **Enhanced Utilization of Renewable Energy:** The model focuses on the distribution of

surplus renewable energy, which reduces dependence on the grid and encourages sustainable energy practices, directly supporting SDG7's aim to increase the global share of renewable energy.

2. **Battery Storage as a Secondary Priority:** After the surplus energy has been effectively distributed to meet immediate demands, battery storage systems play a crucial role in managing any residual energy deficiencies that may arise. These systems are designed to capture and store excess energy for later use, thereby providing a reliable backup during periods of low energy generation or high consumption. It is important to note that each battery unit has specific storage limits that must be adhered to in order to maintain optimal performance and safety. By carefully monitoring and managing these limits, it can be ensured that the battery storage systems operate efficiently and effectively, ultimately contributing to a more stable and resilient energy supply.

   The incorporation of battery systems guarantees a consistent energy supply during times of low generation or high demand, thereby enhancing energy reliability—an essential target of SDG7 for modern energy access.

3. **Grid Energy as the Last Resort:** The model is designed to utilize grid energy exclusively when there is a lack of both surplus energy and battery resources, effectively reducing reliance on external power supplies. This strategic approach ensures that the system prioritizes internal energy sources before resorting to external options.

   By implementing this method, the model not only conserves energy but also enhances its efficiency in managing power resources. The careful allocation of grid energy serves to optimize performance while maintaining a sustainable energy framework.

   By intentionally minimizing reliance on the grid, the model improves energy security and affordability, addressing SDG7's objective of ensuring universal access to reliable and modern energy services.

4. **Fair Energy Distribution:** By implementing this method, the allocation process is made just, allowing for the efficient fulfillment of consumer requirements and the promotion of sustainable energy practices. This method ensures that resources are distributed based on fairness and need, which helps to eliminate inefficiencies and waste that can occur in less equitable systems. As a result, consumers receive the energy they require in a timely manner, which enhances their overall satisfaction and quality of life. Additionally, this approach encourages the integration of sustainable energy solutions, such as solar, wind, and other renewable sources, into the allocation framework. By doing so, it not only meets current energy demands but also supports the transition to a more sustainable energy landscape. This dual focus on justice and sustainability ultimately leads to a more resilient energy system that can adapt to future challenges while meeting the needs of all consumers.

   The results of this study indicate that the suggested energy-sharing model enhances energy efficiency while simultaneously advancing Sustainable Development Goal 7 (SDG7) by facilitating access to affordable, reliable, and sustainable energy for everyone.

   Utilizing Linear Programming and Game Theory, the energy-sharing framework guarantees equitable distribution of resources, encouraging collaborative energy sharing. This methodology aligns with SDG7's commitment to providing affordable energy, especially for marginalized and underserved populations.

7.2 Recommendations and Future Work

To enhance the model's influence and provide additional support for SDG7, the following improvements are suggested:

1. **Real-Time Data Integration Process:**

   The integration of a dynamic monitoring system will enable stakeholders, including energy providers, consumers, and policy makers, to respond more effectively to fluctuations in energy requirements and resource availability. With real-time data at their fingertips, stakeholders can make informed decisions that align with current conditions, such as adjusting energy production levels, redistributing resources, or implementing demand-response strategies. This agility in management will lead to a more resilient energy infrastructure, capable of adapting to unexpected changes in demand or supply. Furthermore, by fostering collaboration among various stakeholders, this system will facilitate a more coordinated approach to energy management, ultimately resulting in improved efficiency and reliability of energy resources. The outcome will be a more sustainable energy ecosystem that not only meets the needs of consumers but also supports broader environmental and economic objectives.

2. **Demand Response Strategies:**

   Implement adaptable pricing structures or incentive initiatives to motivate consumers to modify their energy consumption behaviors, thus decreasing dependence on the grid during periods of high demand. Such strategies can effectively promote energy efficiency and encourage users to shift their usage to off-peak times, ultimately contributing to a more balanced and sustainable energy system.

3. **Advanced Optimization Techniques:**

   The use of advanced methodologies is a significant advancement in the field of resource management that will provide a structured approach to solving optimization problems that involve both integer and continuous variables, making it particularly useful for complex allocation scenarios where multiple factors must be considered. By Applying these advanced Technologies, organizations can systematically evaluate different allocation strategies and identify the most effective solutions that align with their goals. On the other hand, machine learning-driven predictive analytics offers a dynamic approach to forecasting by analyzing vast amounts of data to uncover hidden patterns and correlations. This enables organizations to make informed decisions based on data-driven insights, thereby elevating the accuracy of their forecasts. By delving into these advanced methodologies, organizations can not only optimize their resource allocation processes, but also enhance their overall operational effectiveness, leading to improved results and a competitive edge in their respective industries.

4. **Expansion into Smart Grid :** The implementation of smart meters enabled by the Internet of Things (IoT), combined with blockchain technology for energy trading, facilitates a decentralized and automated system for energy exchange among consumers. This innovative approach improves the efficiency and transparency of energy transactions, enabling users to participate seamlessly in energy trading. This will accelerate progress towards SDG7's target of universal, modern energy access.

   Using these advanced technologies, consumers can participate in a more dynamic energy market, where they can buy and sell energy directly with each other. This not only promotes energy efficiency, but also fosters a sustainable energy ecosystem that benefits all participants involved in the exchange.

This study presents a comprehensive energy sharing framework aligned with the Sustainable Development Goal 7 (SDG7), aimed at enhancing energy generation, distribution, and consumption. Facilitates South Africa's shift towards a low-carbon, resilient, and cost-effective energy future by :

- **Emphasizing the use of renewable energy sources** to promote sustainability.

- **Utilizing advanced forecasting and optimization techniques** to improve cost effectiveness.
- **Fostering equitable energy sharing** to guarantee fair access and encourage energy democracy.

By integrating technological advancements with a commitment to social responsibility, this research not only aids South Africa's energy transition but also contributes to global initiatives aimed at achieving SDG7, which seeks to provide affordable, reliable, sustainable, and modern energy for everyone.

Bibliography

[1] SAJN Nikolina. International renewable energy agency (irena). 2016.

[2] THOKOZANI CHILENGA-BUTAO and W Holl. The role of the department of mineral resources and energy in south africa's just energy transition. 2024.

[3] Henerica Tazvinga, Xiaohua Xia, and Jiangfeng Zhang. Minimum cost solution of photovoltaic diesel battery hybrid power systems for remote consumers. *Solar Energy*, 96:292–299, 2013.

[4] ATD Perera, RA Attalage, KKCK Perera, and VPC Dassanayake. Designing standalone hybrid energy systems minimizing initial investment, life cycle cost and pollutant emission. *Energy*, 54:220–230, 2013.

[5] T Ahmed, KM Muttaqi, and AP Agalgaonkar. Climate change impacts on electricity demand in the state of new south wales, australia. *Applied Energy*, 98:376–383, 2012.

[6] P IEA. World energy outlook 2022. *Paris, France: International Energy Agency (IEA)*, 2022.

[7] Thokozile Mazibuko, Katleho Moloi, and Kayode Akindeji. An assessment to determine energy surplus in an isolated hybrid energy system. In *2024 IEEE PES/IAS PowerAfrica*, pages 1–6, 2024.

[8] Barun K Das and Mahmudul Hasan. Optimal sizing of a stand-alone hybrid system for electric and thermal loads using excess energy and waste heat. *Energy*, 214:119036, 2021.

[9] Meryeme Azaroual, Mohammed Ouassaid, and Mohamed Maaroufi. Model predictive control-based energy management strategy for grid-connected residential photovoltaic–wind–battery system. In *Renewable Energy Systems*, pages 89–109. Elsevier, 2021.

[10] Ranjay Singh, Ramesh C Bansal, Arvind R Singh, and Raj Naidoo. Multi-objective optimization of hybrid renewable energy system using reformed electric system cascade analysis for islanding and grid connected modes of operation. *IEEE Access*, 6:47332–47354, 2018.

[11] Chih-Ta Tsai, Teketay Mulu Beza, Emiyamrew Minaye Molla, and Cheng-Chien Kuo. Analysis and sizing of mini-grid hybrid renewable energy system for islands. *IEEE Access*, 8:70013–70029, 2020.

[12] Eric C Okonkwo, Ifeoluwa Wole-Osho, Olusola Bamisile, Muhammad Abid, and Tareq Al-Ansari. Grid integration of renewable energy in qatar: Potentials and limitations. *Energy*, 235:121310, 2021.

[13] Benjamin Boeckl and Thomas Kienberger. Sizing of pv storage systems for different household types. *Journal of Energy Storage*, 24:100763, 2019.

[14] Solomon E Uhunamure and Karabo Shale. A swot analysis approach for a sustainable transition to renewable energy in south africa. *Sustainability*, 13(7):3933, 2021.

[15] SA Stats. Statistical release p0318: General household survey 2018. *Pretoria: ISIbalo House*, 2020.

[16] Koech Cheruiyot, Ezekiel Lengaram, and Mncedisi Siteleki. South africa's energy landscape amidst the crisis: Unpacking energy sources and drivers with 2022 statistics south africa census data. 2023.

[17] S Pillay, JHC Pretorius, and D De Canha. The selection of renewable energy technologies and their cost implications for a developing country: the case of south africa. In *2017 Australasian Universities Power Engineering Conference (AUPEC)*, pages 1–6. IEEE, 2017.

[18] Anton Eberhard, James Leigland, and Joel Kolker. *South Africa's renewable energy IPP procurement program*. World Bank Publications, 2014.

[19] Thomas Chrostek and Vittoria Favaro. The role of decentralized energy communities in supporting telecom infrastructure, 2024.

[20] Muhammed Y Worku. Recent advances in energy storage systems for renewable source grid integration: a comprehensive review. *Sustainability*, 14(10):5985, 2022.

[21] Mohammad Amin Vaziri Rad, Alibakhsh Kasaeian, Xiaofeng Niu, Kai Zhang, and Omid Mahian. Excess electricity problem in off-grid hybrid renewable energy systems: A comprehensive review from challenges to prevalent solutions. *Renewable Energy*, 212:538–560, 2023.

[22] Sumit Sharma, Yog Raj Sood, Naveen Kumar Sharma, Mohit Bajaj, Hossam M Zawbaa, Rania A Turky, and Salah Kamel. Modeling and sensitivity analysis of grid-connected hybrid green microgrid system. *Ain Shams Engineering Journal*, 13(4):101679, 2022.

[23] Fabian Eze, Julius Ogola, Reuben Kivindu, Munonyedi Egbo, and Chioma Obi. Technical and economic feasibility assessment of hybrid renewable energy system at kenyan institutional building: A case study. *Sustainable Energy Technologies and Assessments*, 51:101939, 2022.

[24] Fanny Boulaire, Jonathan Love, and Ian Mackinnon. An adaptive renewable energy plant (arep)-to power local premises and vehicles with 100% renewables. *Energy Strategy Reviews*, 38:100703, 2021.

[25] Mark Z Jacobson and Mark A Delucchi. Providing all global energy with wind, water, and solar power, part i: Technologies, energy resources, quantities and areas of infrastructure, and materials. *Energy policy*, 39(3):1154–1169, 2011.

[26] Tahmid Mostafa Hasan, S Islam, M Hridoy, Humaun Parvez, M Sharif-Al-Amin, Nazmul Huda Chowdhury, Ma Haque, Naimul Arif, and Bangladesh Dhaka. Demand side management of electricity for controlling peak demand in bangladesh. *Bangladesh Power Management Institute (BPMI), Dhaka, Bangladesh*, 2020.

[27] Luigi Passos, José M Cardemil, and Sergio Colle. Feasibility study of using domestic solar hot water systems as alternative to reduce the electricity peak demand in brazil. *Energy Procedia*, 57:2487–2495, 2014.

[28] A Ananthasingam and A Atputharajah. Forecast daily night peak electric power demand in sri lankan power system. In *2015 IEEE 10th International Conference on Industrial and Information Systems (ICIIS)*, pages 238–243. IEEE, 2015.

[29] Diogo Santos Pereira and António Cardoso Marques. An analysis of the interactions between daily electricity demand levels in france. *Utilities Policy*, 76:101368, 2022.

[30] Samson Obu Showers and Atanda K Raji. Benefits and challenges of energy storage technologies in high penetration renewable energy power systems. In *2019 IEEE PES/IAS PowerAfrica*, pages 209–214. IEEE, 2019.

[31] Xing Luo, Jihong Wang, Mark Dooner, and Jonathan Clarke. Overview of current development in electrical energy storage technologies and the application potential in power system operation. *Applied energy*, 137:511–536, 2015.

[32] Vellanki Mehar Jyothi, T Vijay Muni, et al. An optimal energy management system for

pv/battery standalone system. *International Journal of Electrical and Computer Engineering*, 6(6):2538, 2016.

[33] Irtaza M Syed and Kaamran Raahemifar. Predictive energy management and control system for pv system connected to power electric grid with periodic load shedding. *Solar Energy*, 136:278–287, 2016.

[34] Nsilulu T Mbungu, Ramesh C Bansal, Raj M Naidoo, and Mukwanga W Siti. Analysis of a grid-connected battery energy storage based energy management system. In *2020 First International Conference on Power, Control and Computing Technologies (ICPC2T)*, pages 1–5. IEEE, 2020.

[35] Brian Tarroja, Fabian Mueller, Joshua D Eichman, and Scott Samuelsen. Metrics for evaluating the impacts of intermittent renewable generation on utility load-balancing. *Energy*, 42(1):546–562, 2012.

[36] Anshu Prakash Murdan. Modeling and analysis of load balancing and demand response in renewable energy infrastructures. In *2023 IEEE Electrical Design of Advanced Packaging and Systems (EDAPS)*, pages 1–3. IEEE, 2023.

[37] Walid Saad, Zhu Han, H Vincent Poor, and Tamer Basar. Game-theoretic methods for the smart grid: An overview of microgrid systems, demand-side management, and smart grid communications. *IEEE signal processing magazine*, 29(5):86–105, 2012.

[38] Dharmaraj Kanakadhurga and Natarajan Prabaharan. Demand side management in microgrid: A critical review of key issues and recent trends. *Renewable and Sustainable Energy Reviews*, 156:111915, 2022.

[39] Devinder Kaur, Shama Naz Islam, Md Apel Mahmud, Md Enamul Haque, and ZhaoYang

Dong. Energy forecasting in smart grid systems: A review of the state-of-the-art techniques. *arXiv preprint arXiv:2011.12598*, 2020.

[40] Yi Wang, Qixin Chen, Tao Hong, and Chongqing Kang. Review of smart meter data analytics: Applications, methodologies, and challenges. *IEEE Transactions on Smart Grid*, 10(3):3125–3148, 2018.

[41] Devinder Kaur, Rahul Kumar, Neeraj Kumar, and Mohsen Guizani. Smart grid energy management using rnn-lstm: A deep learning-based approach. In *2019 IEEE global communications conference (GLOBECOM)*, pages 1–6. IEEE, 2019.

[42] Devinder Kaur, Shama Naz Islam, and Md Apel Mahmud. A bayesian deep learning technique for multi-step ahead solar generation forecasting. *arXiv preprint arXiv:2203.11379*, 2022.

[43] Yao Zhang, Jianxue Wang, and Xifan Wang. Review on probabilistic forecasting of wind power generation. *Renewable and Sustainable Energy Reviews*, 32:255–270, 2014.

[44] Bidong Liu, Jakub Nowotarski, Tao Hong, and Rafał Weron. Probabilistic load forecasting via quantile regression averaging on sister forecasts. Technical report, 2015.

[45] Yandong Yang, Wei Li, T. Aaron Gulliver, and Shufang Li. Bayesian deep learning-based probabilistic load forecasting in smart grids. *IEEE Transactions on Industrial Informatics*, 16(7):4703–4713, 2020.

[46] Shivam Sharma and Sudhir Mishra. *Electricity Demand Estimation Using ARIMA Forecasting Model*. 01 2023.

[47] Mohamad As'ad. Finding the best arima model to forecast daily peak electricity demand. 2012.

[48] Ying Zhang and Gong Meng. Simulation of an adaptive model based on aic and bic arima predictions. *Journal of Physics: Conference Series*, 2449(1):012027, mar 2023.

[49] Agbassou Guenoukpati, Adekunlé Salami, Kodjo Koffi Mawugno, and Kossi Napo. Short-term electricity generation forecasting using machine learning algorithms: A case study of the benin electricity community (c.e.b). *TH Wildau Engineering and Natural Sciences Proceedings*, 1, 06 2021.

[50] Ping-ping Xiong, Yao-guo Dang, Tian-xiang Yao, and Zheng-xin Wang. Optimal modeling and forecasting of the energy consumption and production in China. *Energy*, 77(C):623–634, 2014.

[51] Leo Breiman. Bagging predictors. *Machine learning*, 24:123–140, 1996.

[52] Thomas G Dietterich. An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization. *Machine learning*, 40:139–157, 2000.

[53] Leo Breiman. Using adaptive bagging to debias regressions. Technical report, Technical Report 547, Statistics Dept. UCB, 1999.

[54] Yuanyuan Wang, Shanfeng Sun, Xiaoqiao Chen, Xiangjun Zeng, Yang Kong, Jun Chen, Yongsheng Guo, and Tingyuan Wang. Short-term load forecasting of industrial customers based on svmd and xgboost. *International Journal of Electrical Power & Energy Systems*, 129:106830, 2021.

[55] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794, 2016.

[56] Zeliha Ergul Aydin and Zehra Kamisli Ozturk. Performance analysis of xgboost classifier with missing data. In *1st Int. Conf. Comput. Mach. Intell., no*, 2021.

[57] H.M. Al-Hamadi and Soliman Soliman. Long-term/mid-term electric load forecasting based on short-term correlation and annual growth. *Electric Power Systems Research*, 74:353–361, 06 2005.

[58] Rashmi Bareth, Anamika Yadav, Shubhrata Gupta, and Mohammad Pazoki. Daily average load demand forecasting using lstm model based on historical load trends. *IET Generation, Transmission & Distribution*, 18(5):952–962, 2024.

[59] Diksha Jain, N Sandeep, Arun Kumar Verma, and Udaykumar R Yaragatti. A simple methodology for sizing of stand-alone pv-battery system. In *2019 8th International Conference on Power Systems (ICPS)*, pages 1–5. IEEE, 2019.

[60] Guido Carpinelli, Fabio Mottola, and Daniela Proto. Probabilistic sizing of battery energy storage when time-of-use pricing is applied. *Electric Power Systems Research*, 141:73–83, 2016.

[61] Hamza Alloui, Mohamed Becherif, and Khoudir Marouani. Modelling and frequency separation energy management of fuel cell-battery hybrid sources system for hybrid electric vehicle. In *21st Mediterranean Conference on Control and Automation*, pages 646–651. IEEE, 2013.

[62] Ehsan Hajipour, Mokhtar Bozorg, and Mahmud Fotuhi-Firuzabad. Stochastic capacity expansion planning of remote microgrids with wind farms and energy storage. *IEEE transactions on sustainable energy*, 6(2):491–498, 2015.

[63] Islam Safak Bayram, Mohamed Abdallah, Ali Tajer, and Khalid A Qaraqe. A stochastic sizing approach for sharing-based energy storage applications. *IEEE transactions on smart grid*, 8(3):1075–1084, 2015.

[64] S Jayashree and K Malarvizhi. Methodologies for optimal sizing of battery energy storage in microgrids: A comprehensive review. In *2020 International Conference on Computer Communication and Informatics (ICCCI)*, pages 1–5. IEEE, 2020.

[65] Thongchart Kerdphol, Yaser Qudaih, and Yasunori Mitani. Battery energy storage system size optimization in microgrid using particle swarm optimization. In *IEEE PES Innovative Smart Grid Technologies, Europe*, pages 1–6, 2014.

[66] Amir Toliyat and Alexis Kwasinski. Energy storage sizing for effective primary and secondary control of low-inertia microgrids. In *2015 IEEE 6th International Symposium on Power Electronics for Distributed Generation Systems (PEDG)*, pages 1–7, 2015.

[67] Gokturk Poyrazoglu, Burak Arda Taskiran, and Naim Kayabolen. Optimal battery sizing in micro-grid by linear programming. In *2020 2nd Global Power, Energy and Communication Conference (GPECOM)*, pages 220–225, 2020.

[68] Ahmed Al Ameri, Cristian Nichita, T Riouch, and R El-Bachtiri. Genetic algorithm for optimal sizing and location of multiple distributed generations in electrical network. In *2015 Modern Electric Power Systems (MEPS)*, pages 1–7. IEEE, 2015.

[69] Malti Baghel, Shikha Agrawal, and Sanjay Silakari. Survey of metaheuristic algorithms for combinatorial optimization. *International Journal of Computer Applications*, 58, 10 2012.

[70] James Kennedy and Russell Eberhart. Particle swarm optimization. In *Proceedings of ICNN'95-international conference on neural networks*, volume 4, pages 1942–1948. ieee, 1995.

[71] John G. Vlachogiannis and Kwang Y. Lee. Economic load dispatch—a comparative study on heuristic optimization techniques with an improved coordinated aggregation-based pso. *IEEE Transactions on Power Systems*, 24(2):991–1001, 2009.

[72] Yuhui Shi and Russell C Eberhart. Fuzzy adaptive particle swarm optimization. In *Proceedings of the 2001 congress on evolutionary computation (IEEE Cat. No. 01TH8546)*, volume 1, pages 101–106. IEEE, 2001.

[73] D. Kavitha, P. Renuga, and S. Muthamil Priya. Optimal placement and location of distributed generators in distorted distribution system. In *2014 International Conference on Circuits, Power and Computing Technologies [ICCPCT-2014]*, pages 533–539, 2014.

[74] Lefeng Cheng and Tao Yu. Game-theoretic approaches applied to transactions in the open and ever-growing electricity markets from the perspective of power demand response: An overview. *IEEE access*, 7:25727–25762, 2019.

[75] Alihan Kabalak, Elena Smirnova, and Jürgen Jost. Non-cooperative game theory in biology and cooperative reasoning in humans. *Theory in Biosciences*, 134:17–46, 2015.

[76] Anwar Hood Ahmed, Henry Bwisa, Romanus Odhiambo Otieno, and Kabare Karanja. Strategic decision making: Process, models, and theories. 2014.

[77] Waqas Amin, Qi Huang, M Afzal, Abdullah Aman Khan, Khalid Umer, and Syed Adrees Ahmed. A converging non-cooperative & cooperative game theory approach for stabilizing peer-to-peer electricity trading. *Electric power systems research*, 183:106278, 2020.

[78] Takako Fujiwara-Greve. *Non-cooperative game theory*. Springer, 2015.

[79] Georgios Chalkiadakis, Edith Elkind, and Michael Wooldridge. Cooperative game theory: Basic concepts and computational challenges. *IEEE Intelligent Systems*, 27(3):86–90, 2012.

[80] Jingtao Shi, Guangchen Wang, and Jie Xiong. Linear-quadratic stochastic stackelberg differential game with asymmetric information. *Science China Information Sciences*, 60:1–15, 2017.

[81] Csaba Fogarassy. Theory of cooperative and non-cooperative games in practice of environmental or sustainable process modeling. *HUNGARIAN AGRICULTURAL ENGINEERING*, 26:54–61, 2014.

[82] Wenhui Zhao, Haonan Yan, and Wei He. Equilibrium model of electricity market based on non-cooperative game of wind farms, thermal power plants and power grid company. *Power System Technology*, 42(1), 2018.

[83] Xinze Zheng, Qiang Li, Cong Bai, Yongquan Nie, and Chuhong Huang. Energy trading management based on stackelberg game theory to increase independence of microgrids. *Energy Reports*, 8:771–779, 2022.

[84] Yolanda Matamala and Felipe Feijoo. A two-stage stochastic stackelberg model for microgrid operation with chance constraints for renewable energy generation uncertainty. *Applied Energy*, 303:117608, 2021.

[85] M.S. Ismail, M. Moghavvemi, T.M.I. Mahlia, K.M. Muttaqi, and S. Moghavvemi. Effective utilization of excess energy in standalone hybrid renewable energy systems for improving comfortability and reducing cost of energy: A review and analysis. *Renewable and Sustainable Energy Reviews*, 42:726–734, 2015.

[86] Farzad Ghayoor, Andrew Swanson, and Hudson Sibanda. Optimal sizing for a grid-connected hybrid renewable energy system: A case study of the residential sector in durban, south africa. *Journal of Energy in Southern Africa*, 32:11–27, 12 2021.

[87] Omar Hafez and Kankar Bhattacharya. Optimal planning and design of a renewable energy based supply system for microgrids. *Renewable Energy*, 45:7–15, 2012.

[88] John A Duffie, William A Beckman, and Nathan Blair. *Solar engineering of thermal processes, photovoltaics and wind*. John Wiley & Sons, 2020.

[89] Mamoon Ur Rashid, Irshad Ullah, Muhammad Mehran, MNR Baharom, and Faisal Khan. Techno-economic analysis of grid-connected hybrid renewable energy system for remote areas electrification using homer pro. *Journal of Electrical Engineering & Technology*, 17(2):981–997, 2022.

[90] Yashwant Sawle, Siddharth Jain, Sanjana Babu, Ashwini Ramachandran Nair, and Baseem Khan. Prefeasibility economic and sensitivity assessment of hybrid renewable energy system. *IEEE Access*, 9:28260–28271, 2021.

[91] Dispersed Generation Photovoltaics and Energy Storage. Ieee guide for array and battery sizing in stand-alone photovoltaic (pv) systems. 2008.

[92] Jean De Dieu Niyonteze, Fumin Zou, Godwin Asemota, Samuel Bimenyimana, and Gilbert Shyirambere. Key technology development needs and applicability analysis of renewable energy hybrid technologies in off-grid areas for the rwanda power sector. *Heliyon*, 6:e03300, 01 2020.

[93] James F Manwell, Jon G McGowan, and Anthony L Rogers. *Wind energy explained: theory, design and application*. John Wiley & Sons, 2010.

[94] Fahad Ali, Muahnmmad Ahmar, Yuexiang Jiang, and Mohammad Alahmad. A techno-economic assessment of hybrid energy systems in rural pakistan. *Energy*, 215:119103, 10 2020.

[95] Meryl Jagarnath, Tirusha Thambiran, and Michael Teweldemedhin Gebreslasie. Identifying opportunities for low carbon emission zones in south africa: a case study of durban. 2015.

[96] Boon L. Lee, Clevo Wilson, Paul Simshauser, and Eucabeth Majiwa. Deregulation, efficiency, and policy determination: An analysis of australia's electricity distribution sector. *Energy Economics*, 98:105210, 2021.

Appendices

Appendix A

Raw Data and Preprocessing

A.1 Data Source and Accessibility

The raw dataset utilized in the study was too vast to provide in the appendix. However, it has been made available at the following place for reference and reproducibility [96]:

- **File Format:** CSV file Database
- **Size:** Dataset File Size: 46.70 MB,

A.2 Criteria for data reduction and selection

To make computation practicable, the dataset was reduced from 300 consumers to 10 representative clusters, using the following criteria:

- Diversity Factor and Load Factor Analysis for each Consumer:

A.2.1 Summary Statistics of Selected Data

Instead of the entire raw data set, the following is a summary of relevant statistics for the top ten users, identifying significant energy habits for optimal analysis.

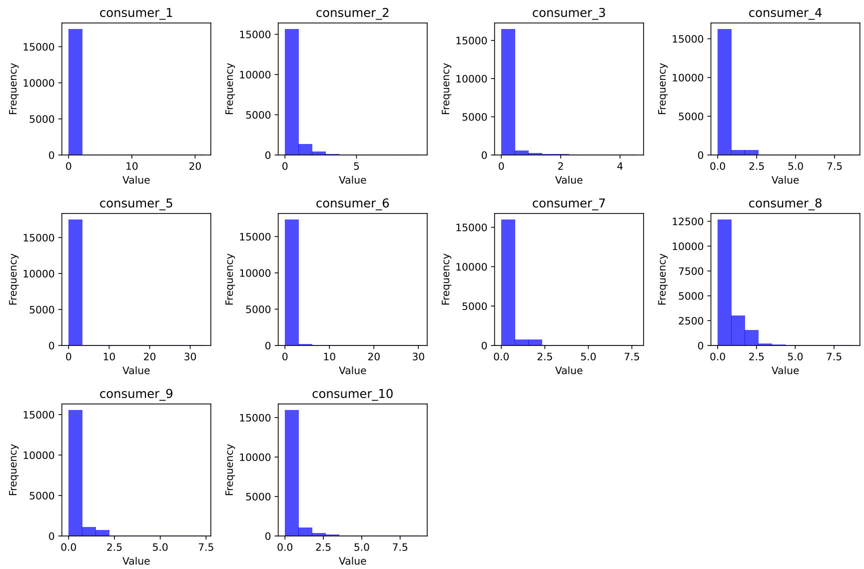

Figure A.1 Statistical Sampling

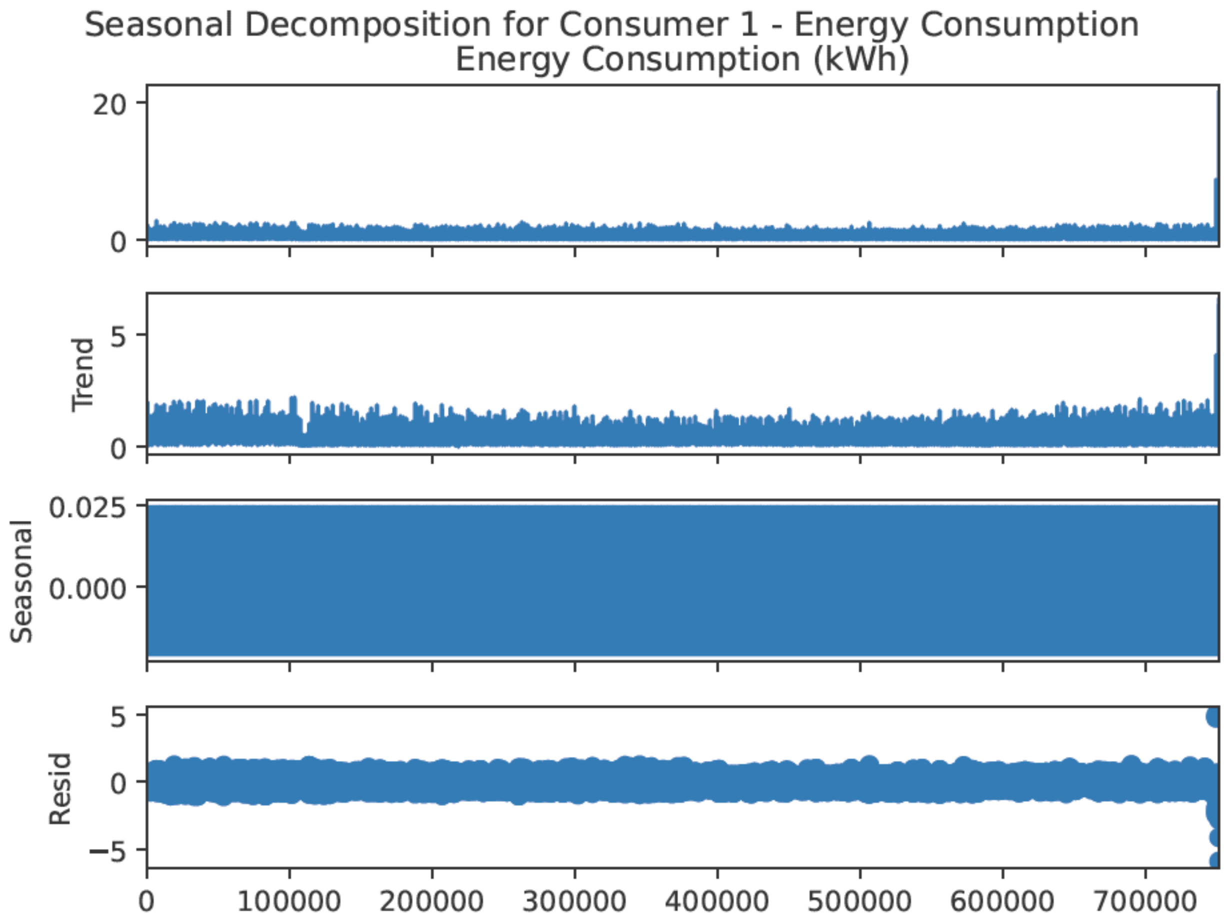

Figure A.2 Seasonal Decomposition Consumer 1

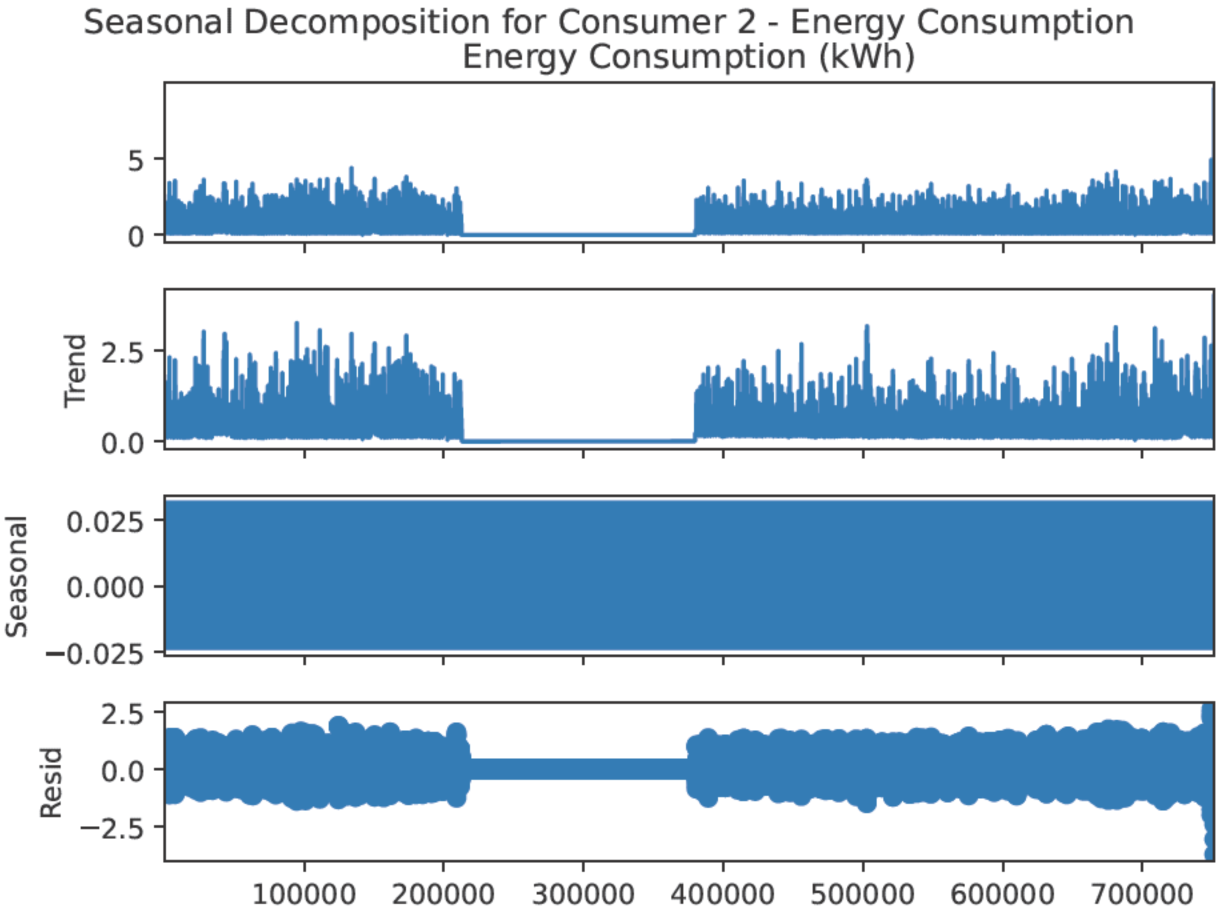

Figure A.3 Seasonal Decomposition Consumer 2

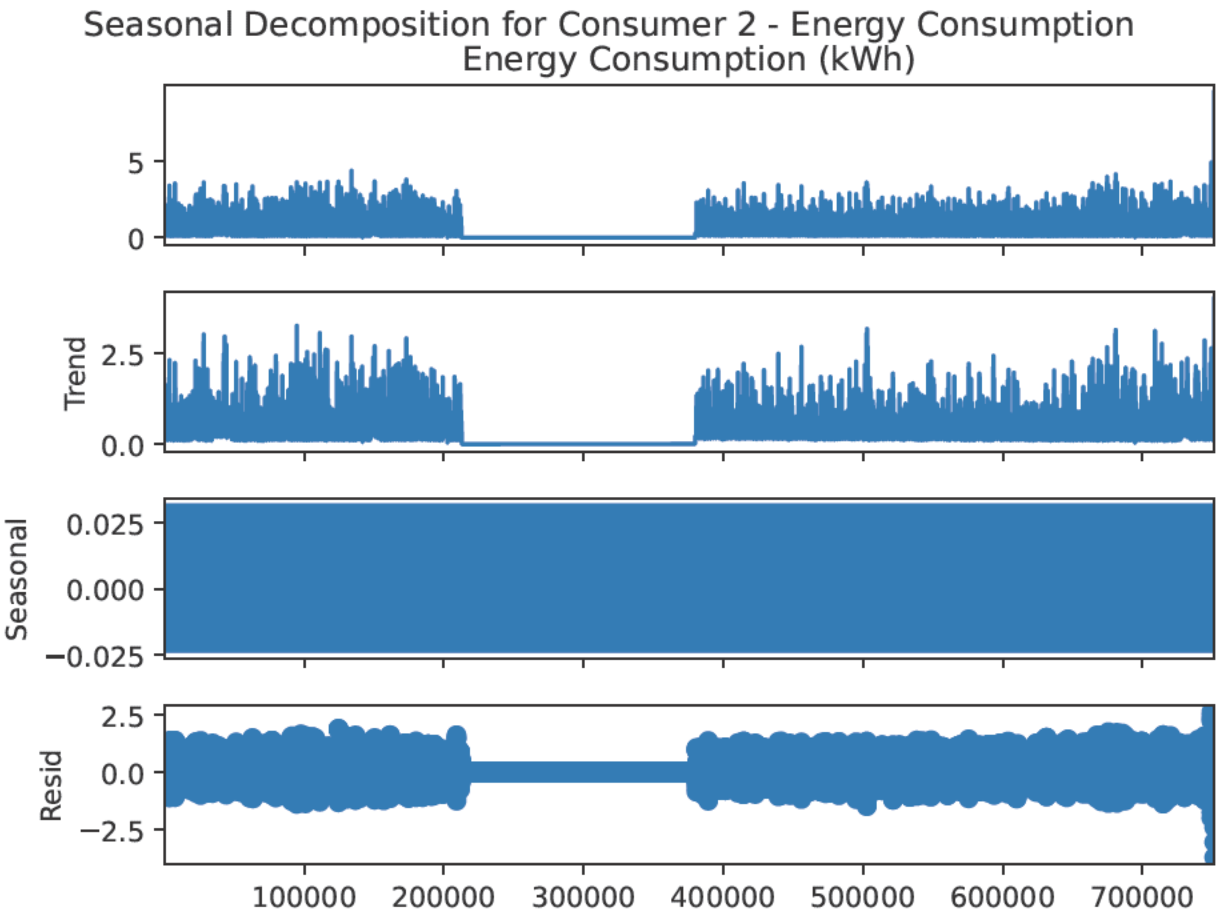

Figure A.4 Seasonal Decomposition Consumer 3

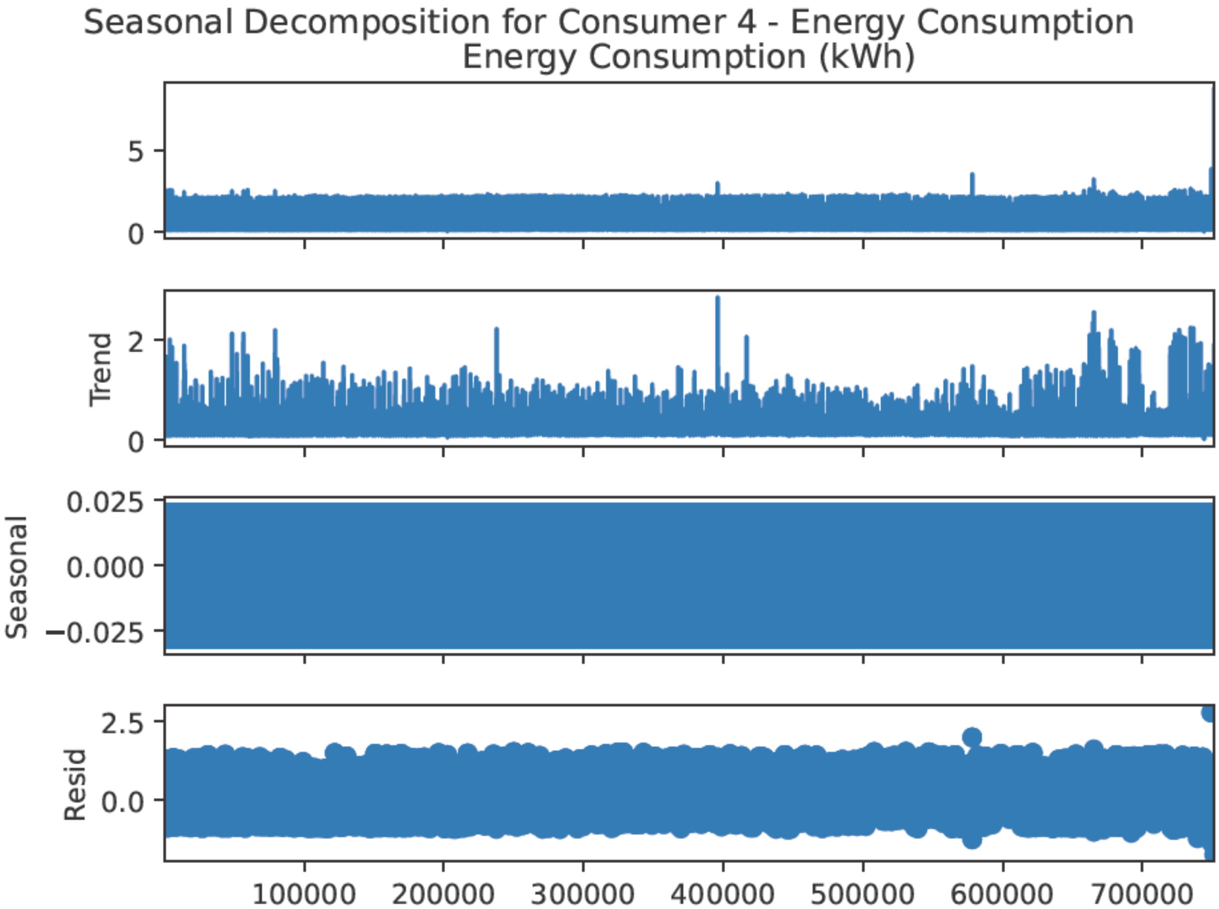

Figure A.5 Seasonal Decomposition Consumer 4

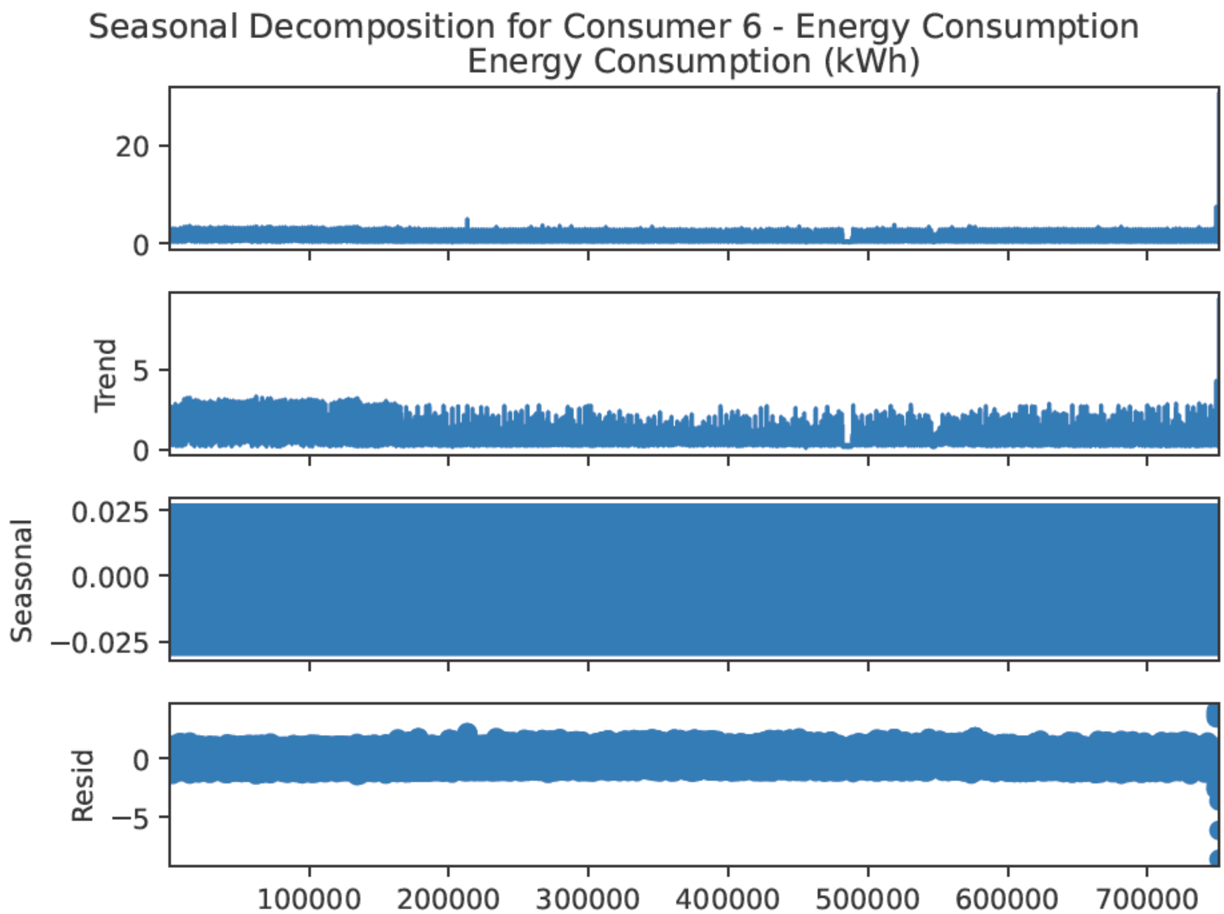

Figure A.6 Seasonal Decomposition Consumer 6

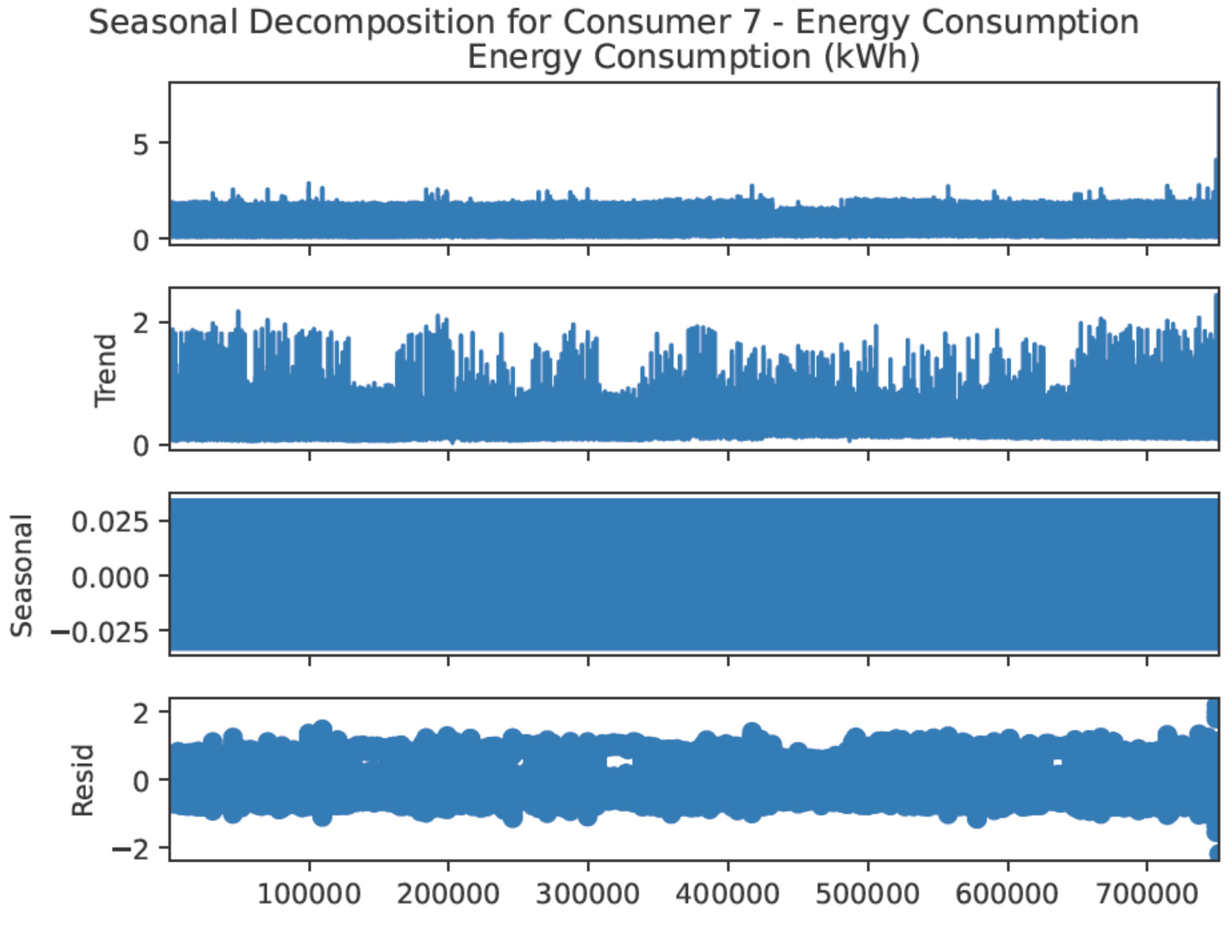

Figure A.7 Seasonal Decomposition Consumer 7

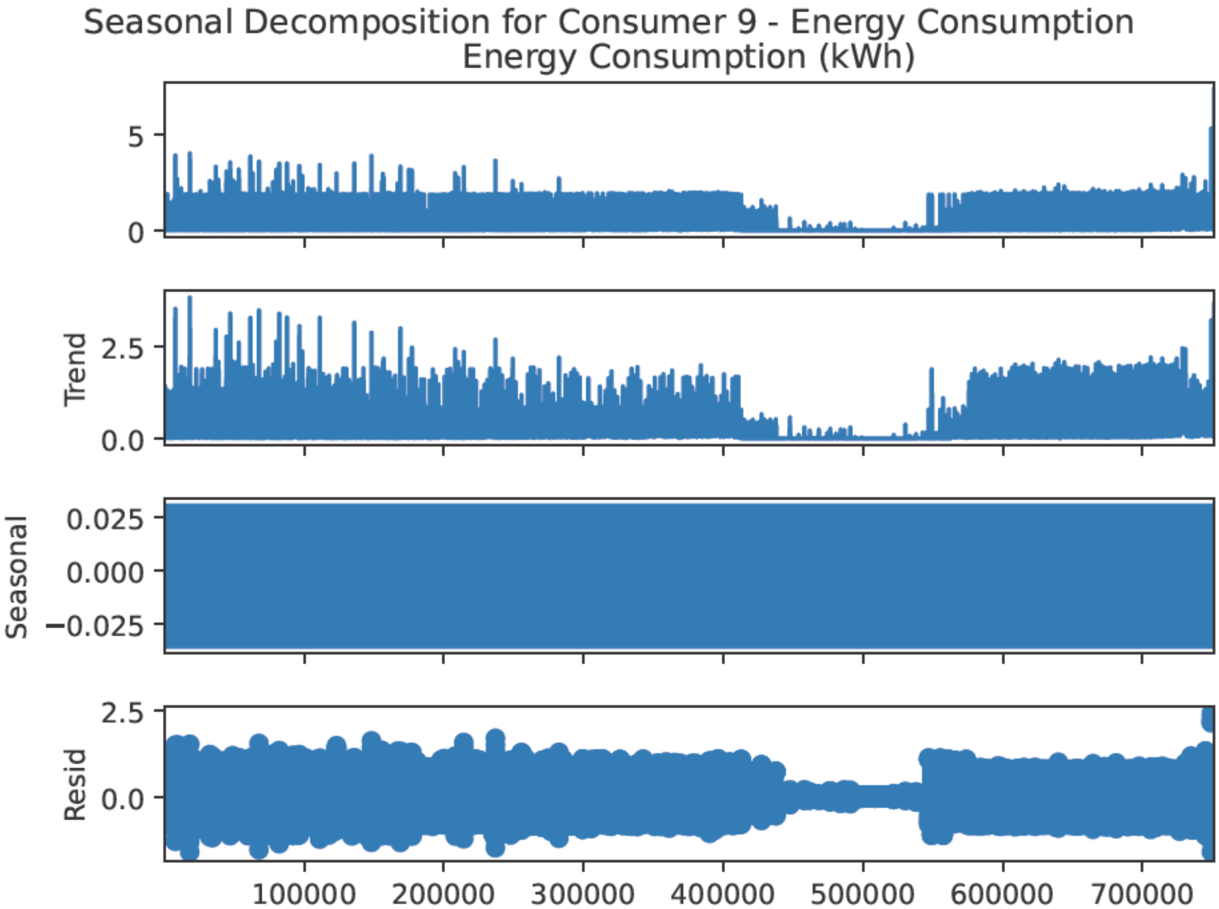

Figure A.8 Seasonal Decomposition Consumer 9

Figure A.9 Seasonal Decomposition Consumer

Appendix B

Load Prediction Models and Performance Metrics

This appendix presents the LSTM, RF, XGBoost, and ARIMA models used for energy consumption forecasting:

B.1 Specified Hyperparameters Used

- **LSTM**: Learning rate = 0.001, Batch size = 32, Epochs = 10
- **Random Forest :** Number of trees = 100, random
- **XGBoost :** Default Learning rate = 0.3, n-estimators=100, max depth=5, random state=42
- **ARIMA :** ARIMA was fiited with the auto-arima function from pmdarima library,hence auto-Arima function automatically determined the best values for ARIMA parameters (p,d,

and q) based on the data.

language=Python, basicstyle=, keywordstyle=, commentstyle=, stringstyle=, numbers=left, numberstyle=, stepnumber=1, numbersep=5pt, frame=single, tabsize=4, captionpos=b,

B.2 LSTM Model Code

Here is the code for the LSTM model:

```
import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

from sklearn.model_selection import train_test_split

from sklearn.preprocessing import StandardScaler, LabelEncoder

import tensorflow as tf

from tensorflow.keras.models import Sequential

from tensorflow.keras.layers import LSTM, Dense

from tensorflow.keras.callbacks import EarlyStopping

from sklearn.metrics import mean_absolute_error, r2_score


file_path = r'C:\Users\thokozilem1\OneDrive - Durban University of
Technology\Desktop\XGBoost\categorized_data_consumption_with_timestamp.csv'

df = pd.read_csv(file_path)

print("Initial dataset:")

print(df.head())

print("\nColumn data types:")

print(df.dtypes)

 Encoding categorical features

print("\nEncoding categorical features...")

label_encoder = LabelEncoder()

df['Consumer'] = label_encoder.fit_transform(df['Consumer'])

df = pd.get_dummies(df, columns=['load'], prefix='load')

 Print the dataset after encoding

print("\nDataset after encoding:")

print(df.head())

feature_columns = ['Consumer', 'Energy Consumption (kWh)'] + [col for col in df.columns if 'load_' in
col]

target_column = 'total_energy_consumption_kWh'


X = df[feature_columns].values  # Input features
```


Figure B.1 LSTM Model