Machine Learning Techniques for Accurate Prediction and Improved Matching of Renewable Energy Production, Storage, and Consumption

A thesis

submitted in fulfilment

of the requirements for the degree

of

Doctor of Philosophy in Computer Science

at

The University of Waikato

by

Mostafa Pasandideh

THE UNIVERSITY OF

WAIKATO

Te Whare Wānanga o Waikato

2025

Abstract

In recent decades, the electricity industry has undergone considerable transformation, characterized by the expansion of competitive electricity markets, integration of renewable energy sources, and the deployment of digital systems. These developments have contributed to increased complexity and variability within power systems, highlighting the need for more accurate, adaptive, and intelligent methods for managing electricity generation, storage, and consumption. This thesis focuses on addressing these challenges, with a particular emphasis on advancing New Zealand's renewable energy ambitions through the application of machine learning (ML) and digital twin technologies for optimizing renewable energy forecasting and management in both industrial and residential contexts.

In the industrial domain, the research develops real-time solar power monitoring systems and predictive models employing convolutional neural networks (CNNs), recurrent neural networks (RNNs), and long short-term memory (LSTM) networks. These models are designed to enhance forecasting precision and system efficiency by accounting for the inherent variability in renewable energy generation. Complementing this, fuzzy logic-based control strategies are implemented to manage electricity consumption adaptively in industrial processes, including applications within the meat processing industry. The incorporation of digital twin technologies further facilitates real-time monitoring, predictive maintenance, and agile response to evolving operational conditions, thereby supporting system resilience and efficiency.

In residential and community energy systems, the thesis introduces an innovative adaptive energy management framework combining streaming machine learning (SML) with a fractal-structured microgrid architecture. This framework utilizes incremental learning algorithms and digital twin principles to enable real-time adaptive control of energy flows, local trading of electricity, and dynamic adjustments in response to fluctuating supply and demand. The approach aims to enhance battery management, support local energy autonomy, and potentially reduce dependency on centralized grids. Comparative evaluations of centralized, distributed, and hybrid battery storage configurations provide insights into the trade-offs between grid interaction, storage longevity, and operational flexibility, offering practical solutions for optimizing residential microgrids.

In conclusion, this thesis explores a combined, case-study based approach to enhancing renewable energy utilization through the combined application of advanced machine learning and digital twin technologies. By addressing both industrial and residential energy systems, the research

contributes to the broader vision of developing reliable, efficient, and sustainable power infrastructures. The findings align with New Zealand’s commitment to reducing carbon emissions and advancing renewable energy technologies, offering case-study evidence and potential approaches to support the transition toward a low-carbon energy future.

Acknowledgements

The past three years of my PhD journey have been transformative, offering me the opportunity to discover a research field that I am deeply passionate about. This experience not only helped me develop critical academic and technical skills but also allowed me to connect with inspiring individuals whose support has been invaluable throughout this journey.

First and foremost, I wish to express my sincere and heartfelt gratitude to my chief supervisor, Professor Mark Apperley. His calm and thoughtful guidance, combined with the space he allowed me to make decisions independently, created a truly supportive and trusting environment. Professor Apperley always encouraged me to think freely, while gently steering me back on track when needed—never with pressure, but always with wisdom and patience. His unwavering support, both academically and personally, has played a pivotal role in the completion of this thesis and in shaping me as a researcher.

I am also thankful to Dr. Jason Kurz, my co-supervisor, for his consistent support and helpful feedback across many parts of this research. I would also like to acknowledge Dr. Martin Atkins for his involvement in the supervision of this project and his input during various stages of the research process. I would like to acknowledge Dr. Shafiqur Rahman Tito, whose assistance with accessing electrical equipment and his technical support were essential during the practical stages of my research.

My sincere thanks go to the staff of the School of Computing and Mathematical Sciences, particularly Tim Elphick, who supported my role as a Doctoral Assistant while considering the workload I carried as a PhD student. This role not only eased the financial burden of my studies but also enriched my academic experience. I am also deeply grateful to the University of Waikato for awarding me a Doctoral Scholarship, which made this research journey financially feasible and academically focused.

I am grateful to Dr. Panos Patros for the opportunity to join the Ahuora Project, and for maintaining communication and support during the difficult COVID-19 period, ensuring I could still take part in this valuable research group and collaboration.

Lastly, I want to extend my heartfelt appreciation to the friends, colleagues, and fellow researchers who accompanied me on this journey. Your encouragement, advice, and companionship have meant more than words can express.

Abbreviations and Symbols

Abbreviation	Full Term
A2D	Analogue to Digital
ACDF	Augmented Dickey–Fuller
ACF	Auto Correlation Function
AFOLU	Agriculture, Forestry and Land Use
AI	Artificial Intelligence
ANN	Artificial Neural Networks
AR	Autoregressive
ARIMA	Auto Regressive Integrated Moving Average
ARIMAX	Autoregressive Integrated Moving Average with Exogenous Inputs
BESS	Battery Energy Storage Systems
BP	Back Propagation
CH_4	Methane
CNN	Convolutional Neural Networks
CO_2	Carbon Dioxide
DBN	Deep Belief Networks
DNN	Deep Neural Networks
DSM	Demand-Side Management
EMD	Empirical Mode Decomposition
EMS	Energy Management Systems
FF	Fill Factor
FIS	Fuzzy Inference Systems
FL	Fuzzy Logic
GARCH	Generalized AR Conditional Heteroskedasticity
GHG	Greenhouse Gas
HPS	High-Pressure Stage
HRES	Hybrid Renewable Energy Systems
IC	Integrated Circuit
IPCC	Intergovernmental Panel on Climate Change

LPS	Low-Pressure Stage
LSTM	Long Short-Term Memory
MA	Moving Average
MAE	Mean Absolute Error
MAP	Mean Absolute Percentage Error
MCFC	Molten-Carbonate Fuel Cells
MLP	Multi-Layer Perceptron
ML	Machine Learning
MPC	Model Predictive Control
MPE	Mean Percentage Error
MPPT	MPP Tracking
MPP	Maximum Power Point
MWSO	Modified War Strategy Optimization
NARX	Nonlinear AR Models with Exogenous Inputs
NDC	Nationally Determined Contributions
NLP	Natural Language Processing
OFMSW	Organic Fraction of Municipal Solid Waste
Op-Amp	Operational Amplifier
PACF	Partial ACF
PCA	Principal Component Analysis
PV	Photovoltaic
RBF	Radial Basis Function
RMSE	Root Mean Square Error
RNN	Recurrent Neural Networks
SARIMA	Seasonal ARIMA
SARIMAX	Seasonal ARIMAX
SML	Streaming Machine Learning
SNN	Simulated Neural Networks
SPI	Serial Peripheral Interface
STL	Seasonal-Trend Loess
MBIE	Ministry of Business, Innovation and Employment

Contents

List of Figures

List of Tables

1. Introduction

1.1 Growth of Renewable Energy Usage and the Role of Machine Learning

With increasing concerns relating to the increase in the production levels of Greenhouse Gas (GHG) produced by industries, and as a result of extensive climate changes, agreements have been made in recent years to control the production of carbon dioxide (CO_2) and GHG between countries. Following the Paris agreement [1] in 2015, a recent conference (COP27 - 27th Conference of the Parties to the United Nations Framework Convention on Climate Change) [2] was held in November 2022 in Egypt. One of the main targets set by this congress was to reduce carbon emissions to net zero by 2050.

The solution of using renewable energy to achieve the goal of net zero carbon emissions by 2050 is a critical step in the fight against climate change. Renewable electricity generation is a great source of low carbon energy, so that in order to reduce emissions, a key path is to shift processes away from direct fossil fuel energy to electricity, then to increase the total renewable electricity which is produced [2], [3].

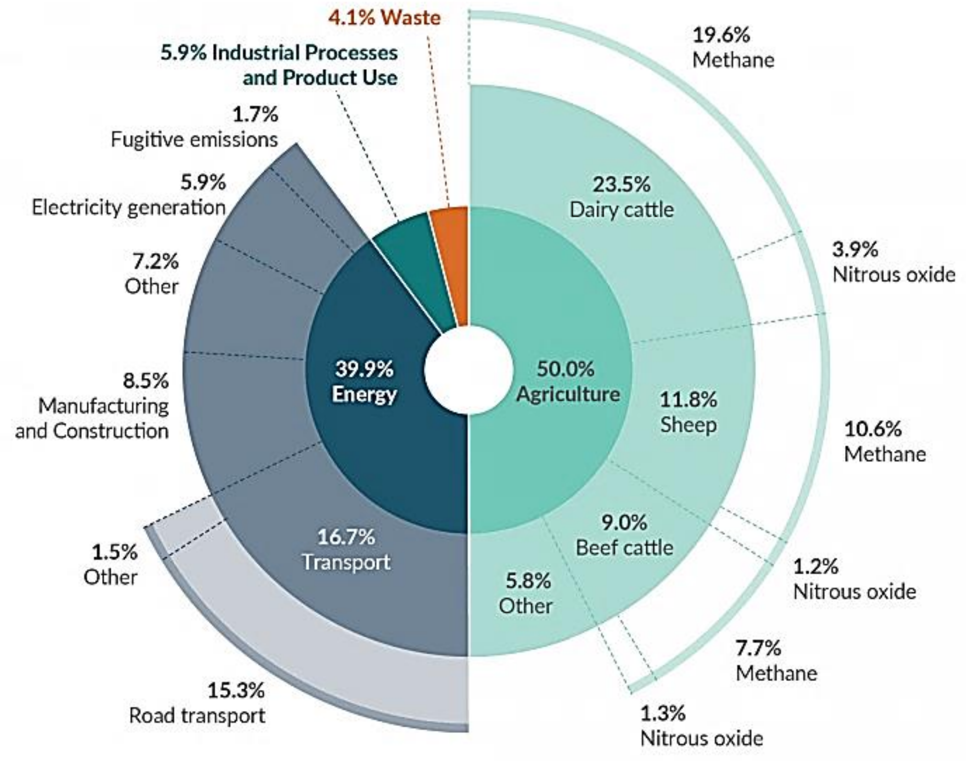


Figure 1.1: NZ's 2020 greenhouse gas emissions by sector, sub-category and gas type [3].

As shown in Figure 1.1, in 2020 energy production was responsible for a sizable portion of New Zealand's GHGs emissions, comprising 40% of total emissions. Within this, manufacturing and construction alone contributed 8.5%, reflecting emissions from industrial fuel use such as process heat for manufacturing and space heating. In addition, the Industrial

Processes and Product Use category contributed a separate 5.9%, representing non-combustion emissions from chemical and industrial activities.

A key component of these emissions is process heat, which refers to heat generated during industrial processes, such as in manufacturing, food processing, and pulp and paper production. In New Zealand, approximately 50% of the demand for this type of heat is met by burning coal or natural gas [4]. This is significant because the combustion of these fossil fuels releases CO_2 and other GHG, contributing to climate change. Reducing the use of fossil fuels in process heat generation is an important step towards transitioning to a low-carbon economy and meeting climate change goals. One approach to reach this goal is to increase the use of renewable energy, and other measures can also be taken to reduce carbon emissions, such as improving energy efficiency, promoting sustainable land use practices, and investing in carbon capture and storage technology.

The adoption of renewable energy as a solution to meet the goal of net zero carbon emissions by 2050 requires a collaborative effort by governments, businesses, and individuals. It is important to recognize the urgency of the issue and act now to transition to a low-carbon energy system and reduce the impacts of climate change.

With regard to renewable energy, the Ministry of Business, Innovation and Employment (MBIE) [5] plays a crucial role in promoting and supporting the development of renewable energy sources in New Zealand. This includes providing funding and support for research and development, facilitating the deployment of renewable energy technologies, and creating a supportive regulatory environment for renewable energy generation and integration.

One of the initiatives MBIE has undertaken to support the development of renewable energy in New Zealand is its support of the Ahuora Renewable Energy program [6]. The program concerns reducing carbon emissions from industrial process heat, through improved energy efficiency and greater use of renewable energy, so reducing GHG emissions. The goal of the Ahuora Renewable Energy program is to accelerate the transition to a low-carbon energy system and to help New Zealand meet its GHG emission reduction targets.

One of the solutions proposed by the Ahuora project to achieve this goal is to invest in renewable energy sources such as solar, wind, or hydro power, and integrating them into the factory's energy system (Aim 3 [1]). This will not only reduce the factory's carbon emissions but also create a source of locally generated power that can be shared with the local community. The aim to integrate renewable energy sources with factory-edge aligns with the growing global trend towards clean and sustainable energy production. By understanding the load profiles of the factory, community, and the electricity grid, a more efficient energy system can be created that reduces dependency on fossil fuels and minimizes the carbon footprint of the manufacturing process. Additionally, utilizing demand-response techniques, energy storage systems, and smart grid technologies, can help to better match energy demand and supply and ensure stability in the electricity grid. The success of this aim will not only benefit the factory

[1] Aim 3 of the Ahuora project focuses on provisioning renewable energy through factory-edge integration, enabling optimal coordination between local generation, factory load, and community demand.

and the local community but also contribute to a cleaner and more sustainable energy future for New Zealand.

Machine learning (ML) potentially has a vital role in maximizing the utilization of renewable energy sources in a factory processes [7]. By analysing energy consumption and generation data, ML algorithms can identify the most effective methods of integrating renewable energy sources into the factory's energy system. For example, ML can predict the energy output of a solar panel or wind turbine, enabling the factory to manage its energy supply and demand more efficiently. Additionally, ML can be used to design control algorithms that can regulate the usage of renewable energy sources in real-time, ensuring that the factory effectively utilizes all renewable energy generated. Utilizing ML in the integration of renewable energy sources into the factory's energy system can reduce energy costs, enhance energy efficiency, and decrease the factory's carbon footprint.

1.2 Problem Identification

The inherent variability of renewable electricity sources, particularly solar and wind energy, poses significant challenges for accurately predicting electricity production ,and understanding the behaviour of the generation sources [8]. This dissertation primarily focuses on solar energy, with a secondary emphasis on wind energy. Traditional methods, such as thermodynamic laws and equations, can be applied to estimate the electricity generated by these sources. However, factors such as dynamic shadowing caused by moving clouds or panel orientation, temperature fluctuations, and other environmental conditions can substantially affect the precision of these calculations [9].

This variability not only impacts the estimation of electricity production at lower levels but also influences higher-level decision-making processes. These include selecting the appropriate electricity source, managing battery charging and discharging, and optimizing overall electricity management strategies, especially in industries such as the meat processing industry.

To address these challenges, this research explores forecasting and electricity management techniques. While various methods have been proposed for these tasks, including physical, statistical, and experimental approaches, their inherent error rates often limit their effectiveness in mitigating the side effects of renewable electricity variability. Consequently, more advanced methods, such as deep learning models, are required. These models have demonstrated superior performance in solving complex problems compared to traditional ML techniques. Nonetheless, the impact of the volume and accuracy of input data on the performance of these models is a crucial factor that cannot be overlooked.

1.3 Aim of the Study

The aim of this study is to address the complexities of managing renewable electricity sources in industrial and community energy systems, focusing on the integration of Fuzzy Logic (FL) [10] for effective electricity management. By leveraging FL alongside advanced ML techniques such as (Convolutional Neural Networks) CNNs, Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks, this research seeks to develop a comprehensive framework for optimizing the prediction, generation, and consumption of renewable electricity.

Fuzzy Logic is utilized to manage the inherent uncertainty and variability of renewable electricity sources, facilitating more accurate and flexible energy distribution strategies. The study validates the proposed FL based models through case studies that include diverse scenarios such as gas turbines, solar energy systems, and the meat processing industry. These case studies highlight the adaptability of FL in managing electricity consumption and improving the efficiency of renewable electricity integration. By incorporating FL, the research aims to enhance the scalability and robustness of renewable electricity systems, ensuring optimal sizing and placement of energy storage while effectively balancing supply and demand under varying conditions. The ultimate objective is to contribute to the development of sustainable and carbon-neutral energy management solutions.

1.3.1 Research Questions

To achieve the aim stated above, this study was guided by the following research questions:

- RQ1: How can FL be effectively utilized to manage the variability and uncertainty in renewable electricity sources within industrial and community clusters?
- RQ2: How can ML methods, such as LSTM, CNN, and RNN, be used to accurately forecast renewable electricity production and consumption?
- RQ3: What are the optimal input parameters for enhancing the performance of the FL based electricity management system?
- RQ4: How can the potential renewable electricity sources for a given community/factory cluster be identified and prioritized, considering environmental factors, load profiles, storage implications, and costs?
- RQ5: How can LSTM or other ML methods be applied to optimally size, integrate and manage a renewable electricity system as a part of a factory/community cluster to maximise to use of renewable electricity while simultaneously minimising emissions and costs?

1.3.2 Research Objectives

The primary objective of this research is to address the inherent variability and uncertainty in renewable electricity sources by developing a robust energy management framework. This framework leverages advanced ML techniques, such as LSTM, CNN, and RNN, alongside FL to optimize electricity production, consumption, and storage within industrial and community clusters. Through this approach, the research seeks to enhance energy efficiency and reliability, particularly in contexts where renewable electricity sources such as solar and wind are utilized.

Additionally, the research demonstrates the practical application of these models through case studies, including solar energy monitoring and energy management in the meat processing industry. These case studies validate the integration of ML and FL, showcasing their effectiveness in real-world energy systems. By optimizing renewable electricity management, the study aims to reduce reliance on conventional energy sources and support sustainable energy practices.

1.4 Motivation for the Research

Technological advancements over the past century have significantly transformed industrial processes, energy consumption, and daily life. These advancements, however, have come at an environmental cost, primarily driven by the extensive use of fossil fuels such as coal, oil, and gas to meet energy demands. The burning of these fossil fuels has dramatically increased atmospheric CO_2 levels, contributing to the greenhouse effect and global climate change [11].

The consequences of climate change, including rising global temperatures, more frequent extreme weather events, and increasing sea levels, present a serious threat to ecosystems and human society. Reducing CO_2 emissions is therefore a global priority, and transitioning from fossil fuels to renewable electricity sources, such as solar and wind, is critical to addressing this challenge.

However, the inherent variability and unpredictability of renewable electricity sources, especially solar and wind, create significant challenges for their large-scale integration into power grids and industrial systems. The intermittency of renewable energy leads to difficulties in ensuring a continuous, reliable energy supply, especially in sectors with high and constant energy demand.

This research is motivated by the need to overcome these challenges and support the transition to a low-carbon future. By employing advanced ML techniques, this study aims to improve the prediction, management, and integration of renewable electricity in industrial systems. The approach leverages technologies such as FL and deep learning to address the variability of renewable sources, enabling more efficient energy distribution and consumption. Ultimately, this research provides case-study evidence to inform strategies for reducing carbon emissions and advancing sustainable energy solutions, in line with global climate goals.

1.5 Contributions of the thesis

This thesis presents several key contributions to the field of renewable electricity management and ML applications, particularly focusing on integrating renewable energy sources within industrial and community systems:

- Development of ML-Based Energy Forecasting Framework (Chapter 4)

  This research introduces an advanced ML framework combining CNN, LSTM, and RNN techniques to improve the prediction of renewable electricity generation and consumption. The models account for the stochastic nature of renewable resources, such as solar energy, while considering environmental factors and system variability.

- Fuzzy Logic-Based Electricity Management System (Chapter 5)

  A novel electricity management system is proposed, utilizing FL to manage the uncertainty and variability in renewable electricity generation. This system enables real-time decision-making for energy distribution, indicating ways to increase the flexibility of renewable energy integration.

- Case Studies and Real-World Applications (Chapter 3 & 5)

  This study includes practical case studies such as a simple hardware set-up for real-time solar energy analysis and meat processing industry systems, demonstrating the effectiveness of the proposed models in real-world scenarios. These case studies illustrate the scalability and applicability of ML techniques within the contexts examined.

- Contributions to Hybrid Renewable Energy Systems (HRES) Optimization (Chapter 6)

  The research investigates the sizing and configuration of HRES, addressing the challenges posed by the variability of renewable electricity sources and demand profiles. The integration of Demand-Side Management (DSM) strategies points to potential improvements in system efficiency and cost-effectiveness.

- Published Contributions

  The research outcomes have been disseminated through the following peer-reviewed publications:

  - "Design and Implementation of a Simple Low-Cost and Real-Time Solar Power Monitoring System" Presented at the IEEE International Conference on DC Microgrids (ICDCM) 2023 [9], (Chapter 3)

    This paper presents a cost-effective solution for monitoring solar power generation in real time, enabling improved energy management at the industrial level.

  - "Predicting Steam Turbine Power Generation: A Comparison of LSTM and Willans Line Model" Published in Energies 2024 [7], (Chapter 4)

    This study compares the efficacy of LSTM models against traditional methods in predicting steam turbine power generation, highlighting the potential of deep learning techniques in enhancing power prediction accuracy.

  - "Enhancing Industrial Energy Efficiency with Predictive Analytics and Fuzzy Logic: A Case Study of Renewable Energy Management in the Meat Processing Industry" Accepted for publication in the proceedings of the 31st International Conference on Neural Information Processing (ICONIP 2024), Springer LNCS, (Chapter 5)

    This paper presents the integration of predictive analytics using Artificial Neural Networks (ANNs) and real-time electricity consumption management through FL within an industrial meat processing context. It highlights how intelligent control strategies can support dynamic energy optimization under fluctuating solar availability.

  - "Energy Management and Edge-Driven Trading in Fractal-Structured Microgrids: A Machine Learning Approach" Published in Energies 2025, (Chapter 6)

This journal publication proposes an adaptive, streaming machine learning (SML) framework for real-time energy forecasting and control in residential microgrids. The work integrates hierarchical fractal architecture, incremental learning, and digital twin principles to manage battery systems and local grid interaction dynamically.

1.6 Organisation of the thesis

This dissertation is structured to progressively build the case for intelligent, data-driven energy management strategies, beginning with theoretical foundations and culminating in an advanced ML application for decentralized microgrids. The progression of chapters reflects a transition from foundational knowledge to practical implementations and advanced adaptive systems.

Chapter 2 presents a comprehensive literature review covering the fundamental concepts and research developments relevant to the thesis. It begins by examining time series data analysis, including statistical methods, ML approaches, and evaluation metrics. The chapter continues with an exploration of FL for decision-making, followed by an overview of the global and national context of GHG emissions and New Zealand's renewable energy potential. It concludes by reviewing the application of ML in various energy systems and storage scenarios, setting the conceptual stage for the applied research in the following chapters.

Building on this theoretical foundation, Chapter 3 details the design and deployment of a simple, low-cost, real-time solar power monitoring system. It outlines the system architecture, component selection, implementation, and performance analysis, thereby demonstrating a practical example of monitoring tools that can complement predictive energy management in renewable-based systems.

Chapter 4 shifts focus to an industrial setting, presenting a data-driven model for predicting the performance of steam turbines using operational data. The Willans line model—an empirical method that relates turbine power output to steam consumption under idealized isentropic conditions [12]—is employed alongside real-time monitoring inputs, highlighting the potential of predictive analytics in enhancing the operational efficiency of energy-intensive systems. Chapter 5 expands this perspective by integrating predictive analytics with FL to improve energy efficiency in the meat processing industry. Case studies are provided in which weather-based solar energy forecasting is combined with adaptive electricity consumption management strategies using fuzzy rule-based systems. This chapter emphasizes the synergy between data-driven forecasting and flexible control mechanisms in improving industrial energy performance.

Chapter 6 introduces a novel framework for edge-driven energy management in residential microgrids. It applies SML, particularly Hoeffding Trees —an incremental decision tree algorithm capable of updating its structure with each new data point [13]—to forecast electricity demand in real time. A hierarchical fractal architecture is proposed to facilitate local decision-making, enhance resilience, and optimize battery usage. The chapter also compares three battery sizing strategies—centralized, distributed, and hybrid—through simulation studies to assess their impact on grid dependence and battery lifecycle. Finally, Chapter 7 concludes the thesis by summarizing the key findings and discussing their practical implications. It outlines directions for future research, such as incorporating adaptive learning algorithms, accounting for dynamic pricing signals, and extending the framework to a wider range of geographic and climatic contexts. This closing chapter consolidates the thesis contributions and highlights the transformative potential of intelligent, adaptive systems in the future of sustainable energy management.

2. Literature Review

2.1 Literature Review Organisation

This chapter provides a comprehensive foundation for understanding the techniques and concepts that underpin this research. It begins with an overview of time series data analysis, including methods for identifying trends, seasonality, and random fluctuations. While traditional statistical approaches such as Auto Regressive Integrated Moving Average (ARIMA) models and Seasonal-Trend decomposition using LOESS (Locally Estimated Scatterplot Smoothing), commonly abbreviated as STL, are effective for stationary and linear datasets, their limitations in capturing the non-linear dynamics present in renewable energy systems are acknowledged [14].

The discussion then progresses to ML approaches for time series analysis. Fundamental concepts of ML are introduced, followed by a focus on neural network architectures. The structure and operation of ANNs are described, leading to CNNs for local feature extraction and RNNs for sequential data modelling. The challenges of RNNs in handling long-term dependencies are addressed, paving the way for a detailed explanation of LSTM networks. LSTM architectures are analysed with an emphasis on their gating mechanisms—input, forget, and output gates—and their ability to retain information across time steps.

The chapter continues by outlining the evaluation parameters used to assess model performance, such as Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), and highlights the role of sensitivity and correlation analysis in understanding relationships among energy variables. These metrics provide a basis for evaluating prediction accuracy and model reliability in dynamic environments.

Next, the chapter broadens the context by reviewing the global and national challenges related to GHG emissions, with a focus on New Zealand's renewable energy landscape. The country's capacity in hydro, wind, geothermal, and solar power is discussed, along with the integration challenges posed by variable renewable generation.

Finally, the review explores ML applications in energy systems, encompassing renewable energy forecasting, adaptive microgrid control, industrial energy efficiency, and predictive maintenance. This section emphasizes the transformative potential of ML in enabling smarter, more resilient energy systems. By combining these perspectives, this chapter sets a robust groundwork for the contributions and case studies presented in the subsequent chapters.

2.2 Time Series Data Analysis

The data analysed in this study is inherently sensitive with respect to time, where the sequence of observations carries critical information. Disrupting this order through techniques such as shuffling eliminates valuable insights about temporal dependencies. Research in this field generally falls into two broad categories: statistical methods and ML techniques. Each approach offers distinct advantages and limitations, making them suitable for addressing specific challenges in time series analysis [15], [16].

Statistical Data Analysis methods leverage historical data to uncover patterns and trends using well-established mathematical and probabilistic models. Techniques such as ARIMA, Seasonal Decomposition of Time Series, and Exponential Smoothing are commonly employed for their interpretability and efficiency in handling stationary and linear data [17]. These models are especially useful for short-term forecasting in energy systems where computational simplicity is paramount [18]. However, their reliance on assumptions such as linearity and stationarity often limits their effectiveness in capturing complex dynamics or non-linear relationships present in renewable energy systems [19].

In contrast, ML and Time Series Analysis has emerged as a transformative approach, particularly for non-linear and high-dimensional datasets. Algorithms such as ANNs, RNNs, and LSTM networks are designed to model intricate dependencies and adapt to dynamic systems. These methods have demonstrated superior performance in applications such as renewable energy forecasting, where variability and uncertainty pose significant challenges [20], [21]. The ability to incorporate vast amounts of data, including weather conditions and historical energy generation, makes ML approaches indispensable for modern energy system analysis [22].

To provide a comprehensive understanding of these complementary methodologies, the following two sections separately consider Statistical Data Analysis, and ML and Time Series. The first delves into the theoretical foundations and practical applications of statistical techniques, emphasizing their historical significance in energy systems modelling. The second explores the advancements brought about by ML, particularly in addressing the variability and non-linearity of renewable energy sources. This dual structure allows for a detailed examination of the strengths and limitations of both approaches, offering a robust framework for the subsequent discussions in this thesis.

2.2.1 Statistical Data Analysis

Statistical data analysis has long served as a foundational pillar in time series forecasting, offering structured, mathematically grounded approaches to extract predictive insights from sequential datasets. These models are particularly valuable in energy forecasting applications, where temporal patterns such as seasonality and trend significantly influence system behaviour. Time series data is typically characterized by four key components: trend, representing long-term progression; seasonality, referring to fixed-period patterns within a year; cyclical fluctuations, which may capture multi-year variations such as recurring climate influences or long-term demand shifts; and random noise, which captures irregular and unpredictable variations.

These components are generally analysed using additive or multiplicative decomposition frameworks, forming the conceptual foundation for many classical statistical models. Among the most widely used techniques are Auto Regressive (AR), Moving Average (MA), and Autoregressive Integrated Moving Average (ARIMA) models. These models offer transparency, ease of interpretation, and computational efficiency, especially under assumptions of linearity and stationarity[16], [17]. In energy systems, they are particularly effective for short-term load and generation forecasting, where temporal regularities are prominent.

In addition, methods such as Seasonal Decomposition of Time Series allow practitioners to isolate and analyse trend, seasonal, and residual components separately, enhancing model interpretability. Exponential smoothing methods, including Holt's linear method (which adds trend estimation to simple smoothing) and the Holt–Winters formulation (which further incorporates seasonality), are particularly well-suited for short- to medium-term forecasting in energy systems due to their adaptability and recursive structure [23], [24].

Beyond these foundational models, more advanced statistical techniques have emerged. For example, Autoregressive Integrated Moving Average with Exogenous inputs (ARIMAX) and Seasonal ARIMAX(SARIMAX) models extend ARIMA to include exogenous variables such as temperature or solar irradiance, which are highly relevant in renewable energy contexts. Similarly, Generalized AR Conditional Heteroskedasticity (GARCH)-family models, which are used to model and predict time-varying volatility in time series data, provide insight into uncertainty dynamics and an essential component in forecasting for energy markets and battery storage management.

Furthermore, diagnostic tools such as Auto Correlation Function (ACF) and partial ACF (PACF) guide model selection and validation by revealing lag dependencies and residual structure. These diagnostics ensure that the statistical assumptions underlying the models are met, bolstering forecast reliability and credibility [18].

The following subsections will explore these key statistical approaches—ARIMA, Seasonal ARIMA (SARIMA), ARIMAX/Nonlinear AR Models with External Inputs (NARX), GARCH, Holt-Winters, and STL—in detail. Their mathematical structures, assumptions, and practical implications for renewable energy forecasting will be discussed, supported by recent academic literature. This comprehensive review lays the groundwork for understanding when and how to apply each technique and paves the way for later discussion of ML-based approaches.

ARIMA (Autoregressive Integrated Moving Average)

The ARIMA model is a fundamental time series forecasting method combining AR and MA components on differenced data to handle non-stationarity [14]. An ARIMA model is typically denoted $\mathrm{ARIMA}(p, d, q)$, where p and q are the orders of the AR and MA terms, and d is the order of differencing (integration). In mathematical form, an ARIMA can be expressed as:

$$\Phi p(L)(1-L)^{d}y_{t}=\Theta q(L)\varepsilon t,\Phi p(L)(1-L)dyt=\Theta q(L)\varepsilon t \quad (1)$$

Where L is the lag operator, $(1\text{-}L)^d$ indicates differencing d times. By including the differencing term, ARIMA can model time series with trends or other forms of non-stationarity. Underlying assumptions of ARIMA include linearity of relationships and a roughly constant variance after differencing (though variance-stabilizing transforms like logarithms can be applied if needed). Residuals are assumed to be uncorrelated (white noise). When these assumptions hold, ARIMA effectively captures the temporal dependence structure using a small number of parameters [14] ARIMA models are often fitted using the Box-Jenkins methodology, which relies on ACF and PACF analyses for model identification and diagnostic checks on residuals to validate model assumptions.

ARIMA's appeal lies in its interpretability and proven accuracy for many stationary or near-stationary time series. It excels in short-term forecasting when patterns are stable, making it a common benchmark in energy systems forecasting [14]. ARIMA is computationally efficient and requires relatively little data to train (compared to data-hungry ML models), which is advantageous in energy applications where high-resolution data might be limited. Recent studies continue to use ARIMA as a baseline for renewable energy predictions [25]. ARIMA has been used to forecast monthly solar radiation and Photovoltaic (PV) output in various regions [26], and to predict wind speeds for wind power estimation. These studies confirm that ARIMA can capture the core dynamics of energy time series in the absence of complex nonlinearities.

The primary limitations of ARIMA stem from its linear structure and requirement for stationarity. Real-world renewable energy data (e.g. solar irradiance or wind speed) often exhibit nonlinear patterns -such as sudden spikes in irradiance due to cloud cover or threshold effects in wind turbine output - that deviate from a straight-line relationship between past and future values. ARIMA assumes that the current value is a linear combination of past values and errors, so such threshold effects or nonlinear interactions cannot be represented. In contrast, ARIMA performs adequately when the data show approximately linear dynamics, where changes in the series are roughly proportional to past changes. For instance, one study noted that for very short-term solar irradiance forecasting (minutes to hours ahead), ARIMA yielded large errors due to the highly volatile and rapid fluctuations in sunlight intensity [26]. Moreover, ARIMA handles seasonality only by explicit differencing; if a time series has a strong seasonal cycle, a non-seasonal ARIMA may perform poorly unless extended to SARIMA. In summary, while ARIMA is indispensable as a baseline and often performs well for short-term and roughly linear dynamics, its reliance on stationarity and linearity makes it less effective for the complex, nonlinear behaviour found in many renewable energy systems [14].

SARIMA extends ARIMA by incorporating seasonal trends, making it suitable for energy systems with cyclical patterns. A study analysed the effectiveness of ARIMA and SARIMA models in forecasting wind farm energy production, indicating that these models are effective, but their performance is influenced by the specificity of the data and seasonal patterns [14].

ARIMAX models integrate external influencing factors, such as temperature or solar irradiance, to improve forecasting accuracy. For example, a study developed a self-adaptive ARIMAX model optimized for very-short-term wind speed prediction, incorporating exogenous inputs from weather simulations to enhance performance [27].

NARX are nonlinear counterparts to ARIMAX, often implemented using neural networks. These models are particularly suitable for capturing complex, nonlinear relationships in energy systems. A study utilized NARX models for short-term forecasting of energy production in PV systems, demonstrating their capability in handling nonlinearities in solar energy data [28].

GARCH models focus on modelling the variance of time series, making them useful for capturing volatility in energy-related financial time series, such as electricity market prices. A study employed a finite mixture GARCH approach to model energy price volatility, indicating its viability in mitigating associated risks in energy-related predictions [29].

Each of these models has distinct assumptions and strengths. While ARIMA and SARIMA are suitable for relatively stable and linear energy data, ARIMAX and NARX extend this capacity by incorporating external variables and capturing nonlinearities, respectively. GARCH models address the dynamic nature of variance, making them appropriate for financial and high-frequency energy data. Although these statistical approaches offer interpretability that is often lacking in complex machine learning systems and remain useful benchmarks during early model selection, machine learning methods have generally been shown to outperform them in terms of flexibility and predictive accuracy. Therefore, the focus of this research is placed on ML-based approaches.

2.2.2 Machine Learning and Time Series

ML has emerged as a powerful tool for analysing complex datasets, particularly those with temporal dependencies, such as time series data. This section explores the intersection of ML and time series analysis, focusing on the techniques and methodologies used to model, predict, and interpret temporal data. ML offers distinct advantages over traditional statistical methods by its ability to capture nonlinear relationships and adapt to diverse patterns in data, making it particularly effective for forecasting and anomaly detection.

Time series data, characterized by its sequential nature, poses unique challenges that require specialized methods. Key components such as trend, seasonality, and noise must be accurately identified and accounted for to ensure reliable predictions. ML algorithms, ranging from traditional approaches such as support vector machines to advanced deep learning models such as RNNs and CNNs, provide the flexibility to handle these complexities. Their capacity to learn from historical patterns and adapt to new data makes them an invaluable asset in time series forecasting.

This section begins with an overview of ML categories and their relevance to time series analysis. Following this, the fundamental components of time series data are discussed, highlighting how ML models address challenges such as data sparsity, missing values, and irregular sampling. Additionally, forecasting methods are reviewed, with a particular emphasis on how ML enhances the accuracy and robustness of predictions in various domains, including energy management and industrial applications.

By examining these topics, this section aims to provide a comprehensive understanding of how ML techniques are applied to time series data, setting the stage for more detailed discussions on neural networks and advanced deep learning models in the subsequent section.

Machine Learning

ML is the process by which a system creates models that can learn from specific examples, leveraging prior knowledge and experience, to improve its performance on a defined task without requiring reprogramming. This capability enables machines to generalize from data and adapt to new information, making ML a transformative tool in addressing complex challenges [30],[31].

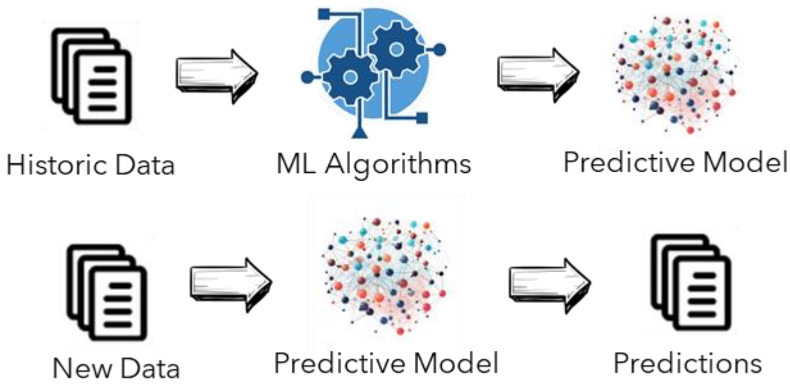


Figure 2.1: Conceptual illustration of the main stages in a machine learning workflow, including data preprocessing, model selection, training, evaluation, and deployment.

At its core, ML attempts to mimic human learning processes. Just as humans organize and relate experiences into structured mental patterns based on pre-existing knowledge, ML systems develop models by identifying patterns in data. These models are then refined through iterative learning, enabling the system to make predictions or decisions in new scenarios. The stages of the ML framework typically include data preprocessing, model selection, training, evaluation, and deployment, as illustrated schematically in Figure 2.1. As shown in Figure 2.1, the stages of the ML workflow begin with raw data preprocessing (documents), proceed through model selection and training (central processing stage), and are refined by evaluation before deployment into practice.

The interdisciplinary achievements of ML span numerous fields, such as Data Mining, Probability and Statistics, Information Theory, Numerical Optimization, Complexity Theory, Control Systems, Neurobiology, and Linguistics. Each of these areas contributes to the theoretical foundations and practical advancements of ML, making it a pivotal discipline in modern technology. In the context of this thesis, ML serves as a cornerstone for optimizing renewable electricity systems, leveraging these diverse influences to address challenges such as predicting solar irradiance, managing energy consumption, and integrating renewable electricity sources efficiently [30],[31].

Artificial Neural Networks (ANN)

Neural networks, which are a subset of ML and integral to deep learning algorithms, are also referred to as ANNs or simulated neural networks (SNNs). They are modelled after the human brain, with their name and structure inspired by the way that biological neurons communicate with one another [32]. One of the most basic neural models available is the Multi-Layer Perceptron (MLP) model, which simulates the transmission function of the human brain.

ANNs are composed of interconnected neurons that are organized into different layers. MLP is a widely used neural network architecture that includes an input layer for receiving external data to conduct pattern recognition, an output layer that produces the solution to the problem, and a hidden layer that functions as an intermediate layer, separating the other layers [33]. This behaviour continues until a specific outcome is achieved, which is likely to eventually lead to a decision, processing, thinking, or action. A typical structure corresponds to MLP, shown in Figure 2.2.

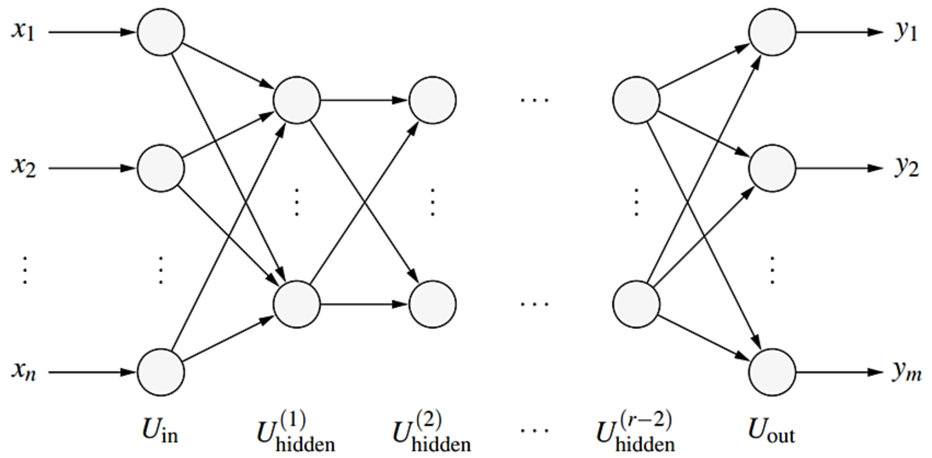


Figure 2.2: Typical structure of an r-layered perceptron [34].

The parameters that define ANNs consist of three main components: firstly, the interconnection pattern established between different layers of neurons, secondly, the learning process utilized to update the weights of the interconnections, and thirdly, the activation function that facilitates the conversion of a neuron's input to its corresponding output activation [35].

ANNs posse many brain-like features based on their structure and foundations. As an example, ANNs possess various advantages such as the ability to learn from experience, generalize previous items to new ones, extract vital features from inputs that represent irrelevant information, and more. Therefore, this type of technology is utilized in numerous areas. The advantages of using ANNs, as outlined in [36] are as follows:

i. Adaptive learning, which enables the network to learn to perform tasks based on prior training or experience.
ii. Self-organization, where neural networks can create an organization or representation of the information it receives through a learning process.
iii. Fault tolerance, which means that even if a network is partially destroyed, some of its capabilities can be preserved.

iv. Real-time operation, as neural computations can be done in parallel, machines with specialized hardware can be designed and built to achieve this capacity.
v. Easy insertion into existing technology, specialized chips can be added to neural networks to improve their ability to perform specific tasks, enabling modular integration into existing systems.

Convolutional Neural Networks (CNNs)

CNNs have increasingly been employed in analysing time-series data due to their capability to automatically detect local temporal patterns through convolutional operations. Unlike traditional statistical methods or simpler neural network architectures, CNNs do not require extensive manual feature engineering, significantly streamlining model development. The convolutional layers in CNNs scan input sequences with filters designed to capture various short-term temporal dynamics and periodic patterns, making them particularly effective in energy-related forecasting tasks where local variations significantly impact system performance [37].

Recent studies highlight the strong capability of CNN-based models, especially when integrated with optimization techniques or hybridized with recurrent architectures such as LSTM networks. For example, research applying CNNs optimized by metaheuristic algorithms, such as the Modified War Strategy Optimization (MWSO), has demonstrated superior performance over traditional time-series forecasting approaches, effectively reducing prediction errors in electricity demand scenarios [38]. Furthermore, CNN-LSTM hybrid architectures have frequently outperformed classical methods, such as SARIMAX models, in electricity load forecasting. This improved performance is primarily attributed to CNNs' inherent ability to extract and generalize local features from large temporal datasets and their efficient parallel computational architecture, facilitating faster training and real-time prediction capabilities essential for effective Energy Management Systems (EMS).

Although CNNs are effective at capturing local temporal features, they typically lack mechanisms to model long-term temporal dependencies inherent in many energy forecasting problems. Therefore, integrating CNNs with RNNs, particularly LSTM networks, has gained attention. LSTM networks specifically address the issue of capturing long-term temporal dependencies, which CNNs alone may not sufficiently model. Recent studies by Chung and Jang [39] and Al-Ali et al. [40] have demonstrated the effectiveness of hybrid CNN-LSTM architectures in significantly improving forecasting accuracy by combining local feature extraction and temporal dependency modelling. Additionally, Salman et al. [41] highlighted the advantage of such hybrid models in solar power forecasting scenarios, emphasizing the complementary capabilities of CNNs and LSTMs. The following section provides a detailed discussion of LSTM networks, highlighting their unique capabilities in handling sequential data with long-range temporal relationships.

Long Short-Term Memory networks (LSTM)

Hochreiter and Schmidhuber introduced the LSTM neural networks in 1997 [20], and it has since undergone numerous improvements by various individuals in the subsequent years.

LSTM networks are a relatively old technique, but it is used in a wide range of issues and is still very popular, Figure 2.3 shows this.

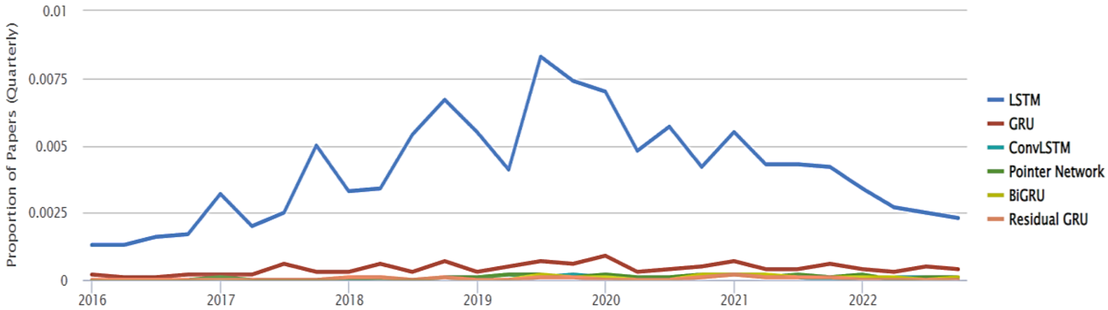
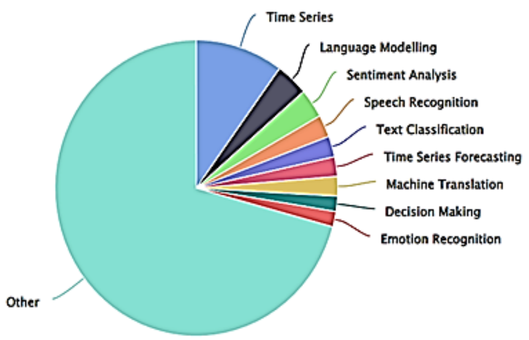


Figure 2.3: Usage of LSTM neural networks in recent times[42].

Figure 2.4 indicates the extent of LSTM usage in different applications, where Other refers to different and lesser-known tasks. Apart from Other, we see that LSTM is most used in time series data.



Task	Papers	Share
Time Series	66	10.05%
Language Modelling	24	3.65%
Sentiment Analysis	21	3.20%
Speech Recognition	16	2.44%
Text Classification	15	2.28%
Time Series Forecasting	14	2.13%
Machine Translation	14	2.13%
Decision Making	11	1.67%
Emotion Recognition	11	1.67%

Figure 2.4: Application rate of LSTM neural networks in different fields [42].

LSTM neural networks are considered a type of RNN, in which long-term time dependence is considered for the input data. In order to have a proper understanding of the function of neural networks, a brief explanation about RNNs will be given first, and then the reasons for the creation and structure of LSTM neural networks will be examined.

Recurrent Neural Networks

RNNs are a type of ANN that is used in speech recognition, Natural Language Processing (NLP) and also in sequential data processing [43]. Many deep networks such as Convolution Neural Networks (CNN) are feed-forward networks, that is, the signal in these networks moves in only one direction from the input layer to the hidden layers, and then to the output layer, and the previous data is not stored in the memory. However, RNNs are equipped with a feedback layer that feeds the output of the network back into itself along with the next input. This feature allows RNNs to retain its previous input due to its internal memory and utilize this memory to process a sequence of inputs. In simpler terms, the recurrent loop in RNN prevents information obtained from previous moments from being lost, allowing it to remain in the network [44].

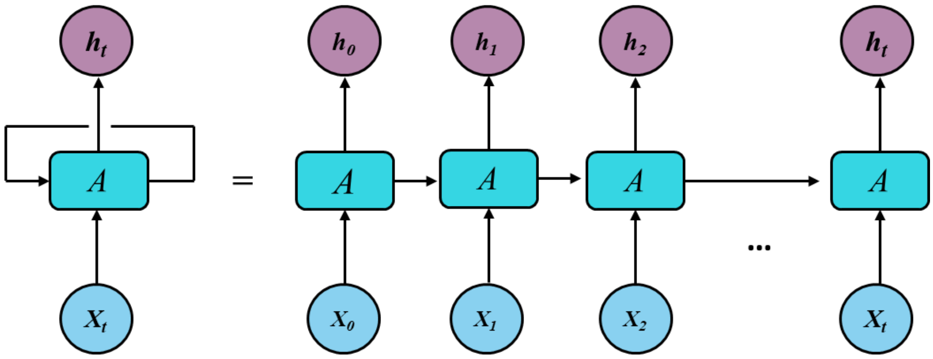


Figure 2.5: Visualisation of RNN that have been unrolled and displayed in a chain-like structure.

As shown in Figure 2.5, X_0 is taken from the input sequence, the output h_0 along with X_1 is the input of the next stage. In the next stage, the output h_1 and X_2 are input the next step is this way, the network will be able to remember previous inputs while training.

Long term dependency problem in RNN

RNN has long-term dependency problems , which means that they cannot perform well in sentences, paragraphs, and all long sequences of data. Consider the Figure 2.6 as a sentence used to train a RNN.

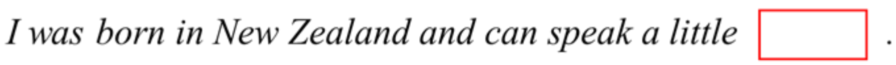


Figure 2.6: Example of long-term dependency problem.

There is no need to refer to the previous sentence or the previous paragraph to fill in the blank space in the sentence in Figure 2.6 rather, only by looking at the words "New Zealand" and "speak" we conclude that the appropriate word for the empty place is "Māori". In order to find the right word for the blank, it was not necessary to look at the neighbouring words before the blank and after the blank, such as "little". It is possible that there is a large gap between related words in a sentence.

Unfortunately, the conventional RNN cannot learn such a distance relationship. This is a big disadvantage. However, looking at the structure of the RNN, it seems that it can transfer even the information from the first words to the last words. However, in practice, this does not happen and RNN has a weakness of long-term dependence.

The solution considered to solve this problem is the use of small memories above the important words in the target sentence, then use these partial memories to fill the empty space, as shown in Figure 2.7. In the mentioned example, a long-term memory is considered for each of the phrases of words New Zealand and born, and it is used to fill the blank.

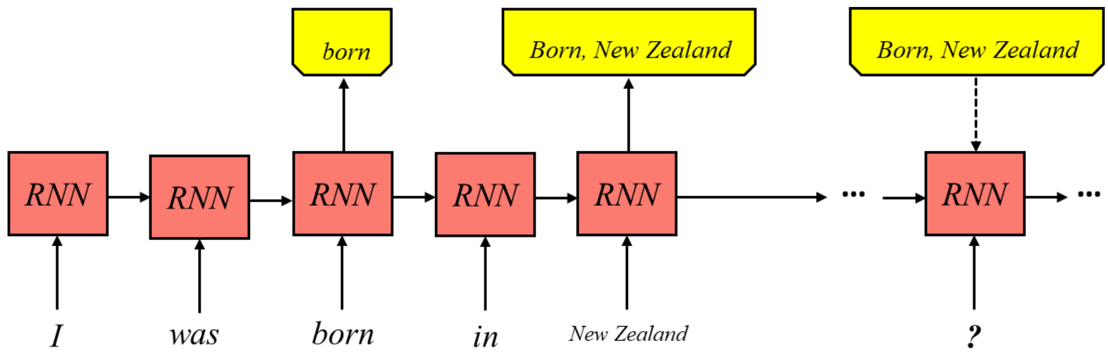


Figure 2.7: Long term memory solution for RNN.

LSTM networks do have this long-term memory. This type of network has the ability to store important information that may need to be accessed in the near (short) future in its long-term memory. This is why this kind of network is LSTM.

LSTM networks architecture

RNN neural networks have a single input and output, with a path connecting the two. However, the LSTM networks differ in that they possess two inputs and outputs, as shown in Figure 2.8.

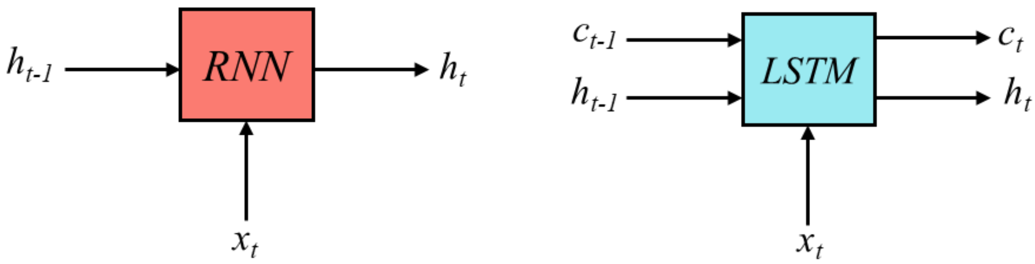


Figure 2.8: Comparing the number of inputs and outputs in RNN and LSTM networks.

Between LTSM cells, the output of one is connected to the input of the next. As it is depicted in Figure 2.8, C_{t-1} input is directly connected to C_t output. This simple connection continues from the beginning to the end of the sequence. *C* stands for Cell State and is a key component in LSTMs. Cell State is also called long-term memory and has the same function as the yellow shape in Figure 2.9.

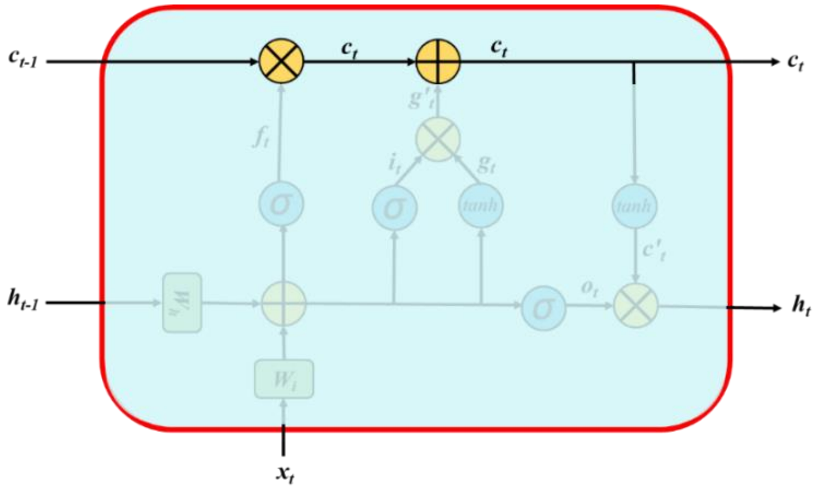


Figure 2.9: Cell State line.

This input-output line connects all LSTM blocks at different time steps. Over time, information is stored or removed from C. This long-term memory has two important properties:

- Information can be erased, the same as forgetting.
- Information can be added, the same as remembering.

As depicted in Figure 2.9 the $\otimes$ and $\oplus$ signs on the $C_{t-1}C_t$ line correspond to forgetting and remembering, respectively. In continue of this study, the process of this forgetting/remembering is studied.

Forgetting in LSTM networks

Gate $\otimes$ on the long-term memory line in Figure 2.10 has two inputs, one of which is C_{t-1}. The second input is an input that goes through a Sigmoid function before being applied. Sigmoid is a function whose output is a number between 0 and 1. As can be seen in Figure 2.10, both inputs are vectors and the multiplication between these two is digit by digit. The colours are meaningfully chosen to allow comparison of input and output.

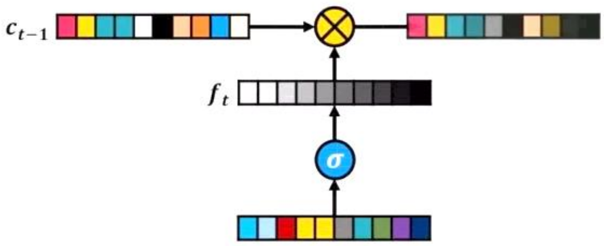


Figure 2.10: Multiplication between two vectors f_t and C_{t-1}.

As shown in Figure 2.10 the output is the same as C_{t-1} with a few changes that the f_t input makes on it. The equation of this process is showed in equation **(3)**. These changes (changing C_{t-1} to C_t) may be as follows:

- Any digit of f_t if it is 0, that means it does not allow its corresponding digit in C_{t-1} to go to the output.

- ❖ Any digit of f_t if 1 means all input C_{t-1} goes to output and remains untouched.
- ❖ If it is a number between 0 and 1, it naturally affects the input to some extent.

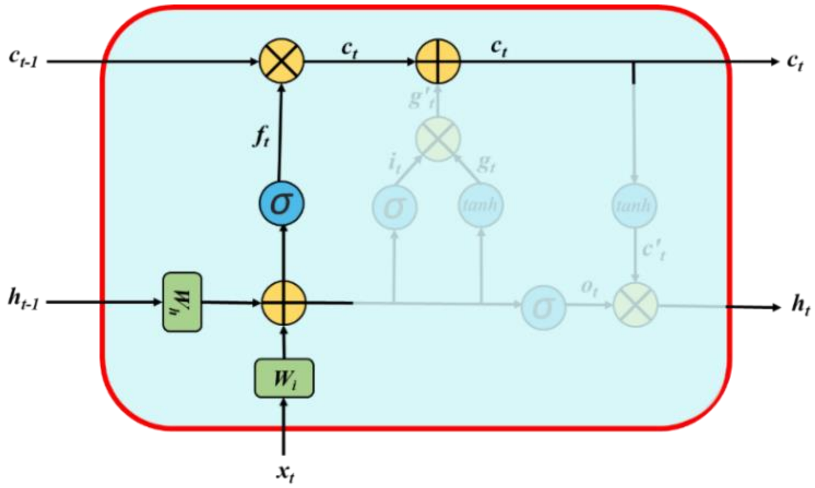


Figure 2.11: Forgetting gate in LSTM neural networks.

$$f_t = \sigma\,(W_{hf}h_{t-1}+W_{if}x_t+b_{hf}+b_{if}) \tag{2}$$

$$C_t = C_{t-1} \otimes f_t \tag{3}$$

The input f_t is formed by small neural networks with two inputs x_t and h_{t-1}. These small neural networks have the task of forgetting part of the information in long-term memory. This neural network is called the forget gate.

Figure 2.11 illustrates how the two input vectors, x_t and h_{t-1}, are combined using MLP. As discussed in the earlier Recurrent Neural Networks section, it is enough to give these two vectors to two fully connected layers and then add these two together. In the forget gate, we have two fully connected layers with weights w_{hf} and w_{aif}. The first layer is intended for the input value h_{t-1}, while the second layer is for input value x_t. The abbreviation *hf* refers to the *hidden* and *forget* components, where *i* represents *input* and *f* represents *forget*. The corresponding formulation is presented in equation **(2)**.

Remembering in LSTM networks

After the ⊗ operator is applied, the path $C_{t-1}C_t$ shown in Figure 2.11, leads to the ⊕ operator. At this point, an additional value is combined with the input value *C*, resulting in the integration of new information. This process is essentially equivalent to memorization in practice. There are two inputs two inputs, one of which is $C_{t,}$ and it is not known what the second entry is, although the vector is the same size as the input *C*. At the current time step (t) some processing has been done, and that now needs to transfer to the memory cell to hold.

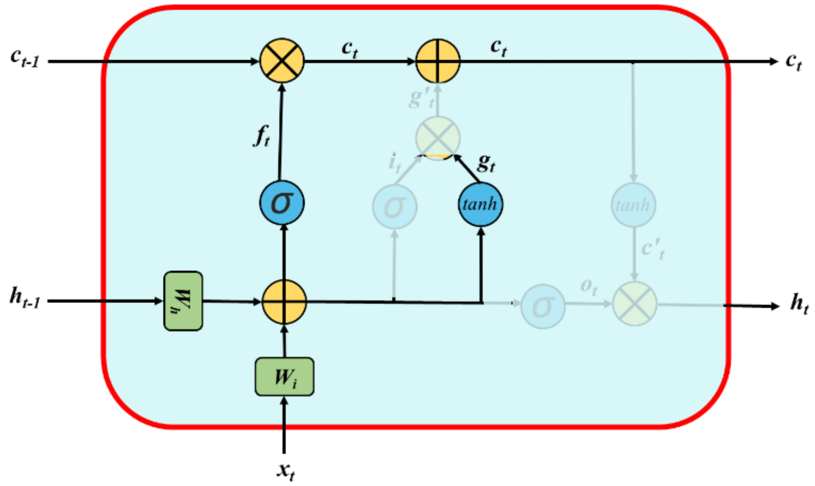


Figure 2.12: Computation of g_t from x_t and h_{t-1} for updating the LSTM memory cell.

Neural networks are employed to compute the information stored in long-term memory at time (t). Similar to the input gate, this neural network has two inputs, namely x_t and h_{t-1}, which are passed through two fully connected layers (w_{ig} and w_{hg}) before being summed. This resulting output is then passed through a hyperbolic tangent function. The resulting value, g_t, which falls within the range of –1 to 1, serves as the input to the next stage. This range is desirable to mitigate the impact of certain components in C. That is, with values between -1 and 1, we can increase or decrease the effect of some components. The structure of this computation is illustrated in Figure 2.12, and the corresponding formula is shown in Equation **(4)**:

$$g_t = tanh\ (W_{hg}\, h_{t-1} + W_{ig}\, x_t + b_{hg} + b_{ig}) \tag{4}$$

However, there is a problem: some information in the g_t output may not be valuable enough to justify updating C_{t-1} . While this input does contain information, it might not be significant enough to commit to long-term memory.

If a gate similar to the forget gate is applied to the output path of g_t, it becomes possible to regulate how much influence this output should have. Therefore, a gate or valve is still required. This new gate is referred to as the input gate.

Input gate in LSTM neural networks

The input gate is an evaluator of the value of the information in the g_t. This is because the input gate which its output is represented by i_t uses a sigmoid function as the activation function. And as previously explained, the output of the sigmoid function has a value between zero and one, which can be said to act as a filter. The function of the input gate is to scrutinize the introduction of a sequence of novel information into the long-term memory. This gate is given the name *entrance gate* due to its role in allowing information to enter. Similar to the forgetting gate, the values within the i_t vector can be in close proximity to zero, causing the effect of g_t to be reduced. Alternatively, if the values of the i_t vector are approximately equal to 1, then g_t is more likely to be stored in long-term memory. Equation **(5)** shows the way of calculation of i_t by use of sigmoid function. In equations **(6)** & **(7)** steps of calculation of C_t as the memory of the cell is explained.

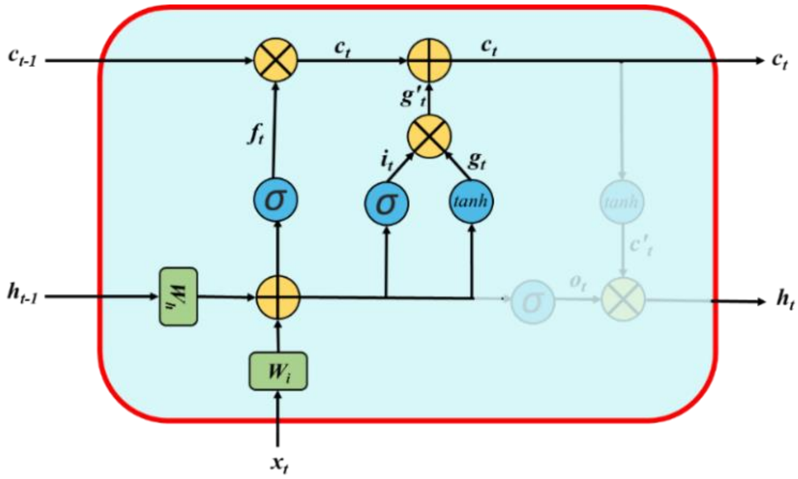


Figure 2.13: Adding an input gate to the internal structure of the LSTM networks.

$$i_t = \sigma\,(W_{hi}\,h_{t-1}+W_{ii}\,x_t+b_{hi}+b_{ii}) \tag{5}$$

$$g'_t = g_t \otimes i_t \tag{6}$$

$$C_t = C_t + g'_t \tag{7}$$

Output structure in LSTM networks

The recurrent LSTM networks connect the newly updated C_t to the output to produce the output h_t. In fact, the effect of long-term memory C_t is transferred to the output. So, we pass the output C_t through a hyperbolic tangent, then we are ready to connect it to the output h_t. The effect on C_t on h_t is showed in equation **(*8*)**.

It would be better to take as much of the information in C_t that is needed and transfer it to the output of h_t. This is done using another gate called *output* gate. Equation **(*9*)** demonstrates how the amount of long-term memory to be transferred to the output is determined by the *output* gate.

Finally, to transfer the required amount of output to the output h_t, it is necessary to multiply the output of the sigmoid function with the output generated from the output gate o_t. In the following, a comprehensive picture of the performance of each of the sections in the LSTM structure, the formulas related to the output section are shown.

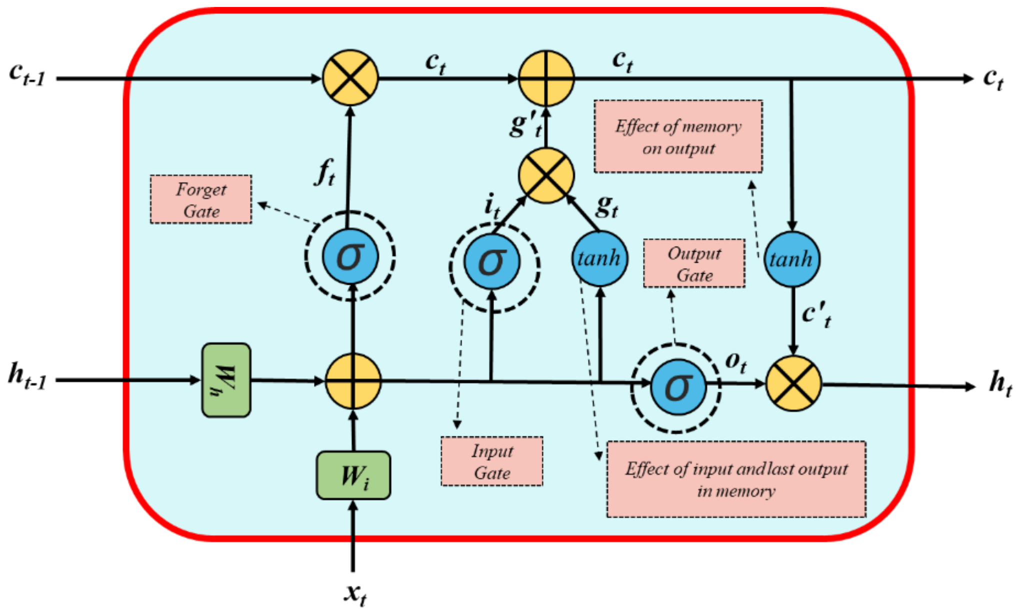


Figure 2.14: Final structure of LSTM neural networks.

$$o_t = \sigma\,(W_{ho}\,h_{t-1} + W_{io}\,x_t + b_{ho} + b_{io}) \tag{8}$$

$$h_t = o_t \otimes tanh(c_t) \tag{9}$$

Table 2.1 compares RNNs and LSTM, which is an improved model compared to RNN [45].

Table 2.1: Comparison between RNN and LSTM.

Model	Pros	Cons
RNN	Simpler than LSTM, short term processing of time-series input data	Limited accuracy for long term forecasting
LSTM	Considering long term dependencies characteristic of the input data and applicable for pattern recognition	High run time and complexity for long term forecasting

2.2.3 Evaluation Parameter

The evaluation of ML models is a fundamental aspect of predictive analytics, as it determines the reliability and effectiveness of the algorithms in addressing specific tasks. In regression problems, such as forecasting renewable electricity parameters, evaluation metrics serve as critical benchmarks for assessing the quality of predictions. These metrics enable the identification of potential shortcomings in the models, allowing researchers to refine algorithms and improve their accuracy. The choice of evaluation parameters depends on the nature of the dataset, the intended application, and the specific challenges associated with the prediction task. For instance, while some metrics emphasize the magnitude of errors, others focus on their relative or proportional aspects. This diversity in evaluation criteria ensures that different aspects of model performance can be comprehensively analysed and compared [46], [47].

In recent years, the focus on evaluation metrics has become increasingly important, particularly in applications involving renewable energy forecasting. These domains require high levels of precision due to the variability of environmental factors such as solar irradiance, wind speed, and temperature. As a result, selecting suitable metrics is not merely a technical consideration but a critical step toward achieving meaningful and actionable insights from predictive models [16].

The selection of evaluation parameters is also an important topic in the field of ML. When the purpose of ML is to predict the value of a specific variable, Mean Absolute Percentage Error (MAPE), RMSE, Mean Percentage Error (MPE), and Absolute Percentage Error (APE) are evaluation parameters that usually used [48]. MPE measures the average difference between predicted and true values over N samples using equation (10). RMSE measures the square root of the average of the squared differences between predicted and true values over N samples using equation (11).

The APE of each sample is calculated using equation (12), which measures the absolute percentage difference between predicted and true values. The MAPE is then calculated by averaging the APEs over N samples using equation (13). By selecting appropriate evaluation parameters, ML models can be optimized to improve the accuracy of the predictions.

$$MPE = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i) \tag{10}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \tag{11}$$

$$APE_i = \left|\frac{(\hat{y}_i - y_i)}{y_i}\right| \times 100\% \tag{12}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}(APE_i) \tag{13}$$

Beyond the metrics discussed, other advanced methods such as Mean Bias Error (MBE), R-squared (coefficient of determination), and normalized metrics are frequently utilized in specific scenarios. For instance, R-squared is particularly effective for evaluating how well predictions align with observed data trends, while MBE offers insights into systematic biases in model outputs [49]. These metrics complement traditional measures, offering additional layers of analysis that help address the unique challenges of diverse datasets and prediction tasks.

Furthermore, as ML applications in renewable energy forecasting become increasingly sophisticated, hybrid approaches combining multiple evaluation parameters are gaining prominence. For example, the simultaneous use of RMSE and MAPE ensures a balanced evaluation, addressing both absolute and relative errors in predictions [41]. This holistic approach enables a deeper understanding of model strengths and weaknesses, paving the way for more targeted improvements and innovations in predictive modelling.

In summary, selecting and applying appropriate evaluation parameters is a cornerstone of ML research. By employing a combination of traditional and advanced metrics, researchers can effectively assess model performance, refine algorithms, and address the unique demands of prediction tasks in renewable electricity and beyond.

2.2.4 Sensitivity and correlation analysis

Sensitivity analysis is a critical component in the development and refinement of predictive models, particularly within the realm of renewable energy forecasting. It enables researchers and practitioners to discern how variations in input variables influence model outputs, thereby identifying the most impactful parameters and facilitating model optimization. This process not only enhances the accuracy of predictions but also contributes to more efficient energy management and planning.

One fundamental technique employed in sensitivity analysis is correlation analysis, which assesses the strength and direction of linear relationships between variables. The Pearson correlation coefficient, denoted as corr(X,Y) , is commonly used for this purpose and is calculated using the following equation:

$$corr(X,Y) = \frac{\frac{1}{N}\sum_{i=1}^{N}(X_i - \bar{X})(Y_i - \bar{Y})}{\sigma_X \times \sigma_Y} \tag{14}$$

In this equation, N represents the number of observations, Xi and Yi are individual sample points, $\bar{X}$ and $\bar{Y}$ denote the mean values of X and Y, respectively, and σ_X and σ_Y are the standard deviations of X and Y. A coefficient value close to +1 indicates a strong positive linear relationship, while a value near -1 signifies a strong negative linear relationship. A coefficient around zero suggests no linear correlation between the variables. This method is extensively utilized in energy forecasting to evaluate how different environmental or operational factors influence key parameters such as energy generation or demand.

Another pertinent technique is autocorrelation analysis, which examines the correlation of a variable with its own past values. This is particularly relevant in time-series data, where understanding temporal dependencies is crucial for accurate forecasting. By identifying patterns and periodicities within the data, autocorrelation analysis aids in constructing models that effectively capture temporal dynamics, thereby improving the reliability of predictions. The significance of autocorrelation in time-series analysis has been well-documented in studies focusing on renewable energy systems.

Principal Component Analysis (PCA) is also a valuable method in sensitivity analysis, especially when dealing with high-dimensional datasets. PCA reduces the dimensionality of the data by transforming correlated variables into a set of uncorrelated components, known as principal components. This transformation retains the most significant features of the data, allowing for simplification of the modelling process without substantial loss of information. In renewable energy forecasting, PCA has been effectively applied to identify the most informative variables, thereby enhancing model efficiency and interpretability. For instance, studies have demonstrated the application of PCA in improving the accuracy of wind power forecasts by capturing the underlying patterns in meteorological data [50].

The importance of sensitivity analysis extends beyond the identification of key variables; it also provides a framework for evaluating model robustness and understanding the dynamics of input-output relationships. Advanced methods, such as variance-based sensitivity analysis, offer a quantitative assessment of the contribution of each input variable to the overall output variability. These methods are particularly beneficial in complex systems where interactions between variables can significantly affect performance. Comprehensive reviews of sensitivity analysis practices in energy modelling highlight the necessity of such techniques in developing reliable and efficient predictive models [51].

Sensitivity analysis and its associated techniques such as correlation analysis, autocorrelation analysis, and PCA are integral to the advancement of predictive modelling in renewable energy systems. By elucidating the relationships between input variables and model outputs, these methods contribute to the creation of models that are not only accurate but also robust and efficient, thereby supporting the effective integration and management of renewable energy sources.

2.3 Fuzzy Logic and Decision Making

Fuzzy systems, rooted in fuzzy set theory pioneered by Zadeh in 1965 [10], offer a revolutionary approach to managing and modelling complex systems. Zadeh's seminal work laid the foundation for controlling intricate nonlinear systems by leveraging expert systems. Two critical theoretical contributions underpinning FL success include the Generalized Modus Ponens Scheme for imprecise reasoning and the Compositional Rule of Fuzzy Inference. These contributions enable fuzzy systems to simulate human reasoning and decision-making processes in uncertain environments, establishing them as vital tools in engineering and system control applications [52], [53].

FL is especially significant due to its ability to formalize human reasoning capabilities in two keyways: first, by enabling rational decision-making within environments characterized by imperfect information, and second, by performing physical and mental tasks without relying on exact measurements or extensive computations [10]. As a rule-based control mechanism, FL operates by formulating mappings between inputs and outputs through linguistic variables, which are defined by membership functions. These membership functions classify variables into fuzzy sets, with each fuzzy set representing degrees of membership between 0 and 1. The mapping process underpins decision-making systems and enables the control of complex, real-world systems.

Fuzzy Inference Systems (FIS) utilize linguistic rules to model relationships between variables. For instance, "IF-THEN" rules form the backbone of fuzzy reasoning, allowing systems to process imprecise input data to produce actionable outputs. These rules are applied using inference mechanisms such as Mamdani or Sugeno models. The Mamdani model, first introduced in 1975, uses fuzzy sets for output membership functions and is often applied in decision-making processes requiring nuanced outputs. In contrast, the Sugeno model employs constant or linear output membership functions, offering computational simplicity and making it particularly suitable for optimization tasks [54], [55].

Fuzzy systems typically consist of four components: fuzzification, a rule base, an inference engine, and defuzzification. Fuzzification converts precise numerical inputs into fuzzy values, while the rule base encapsulates expert knowledge through "IF-THEN" statements that define relationships between variables. The inference engine applies these rules to evaluate fuzzy inputs, and defuzzification transforms the fuzzy output into precise, actionable commands. Membership functions, such as triangular, trapezoidal, and Gaussian shapes, play a pivotal role in defining fuzzy sets. Among these, triangular membership functions are often preferred for their simplicity and ease of implementation [56].

In the domain of energy systems, FL control has demonstrated significant potential. Rule-based FL control systems are increasingly employed to enhance the management and optimization of energy consumption and production across various applications, including industrial facilities, smart grids, and systems harnessing renewable energy [57]. These systems leverage FL's ability to simulate human decision-making in the face of uncertain or imprecise data. This makes FL particularly valuable in complex operational environments where traditional control systems often fall short due to the difficulty of securing precise input data.

A practical example of FL's application in energy management is its use in renewable electricity systems, as illustrated in Figure 2.15. The fuzzy control process begins by identifying and categorizing input and output variables into fuzzy sets. Membership functions are then defined for each set, followed by the development of "IF-THEN" rules that link inputs to outputs. Through a fuzzy inference process, these rules evaluate input values against fuzzy sets to generate fuzzy outputs, which are subsequently defuzzified into actionable commands. This methodology enables adaptive and responsive energy management strategies that are crucial for optimizing resource use in dynamic and uncertain conditions [58].

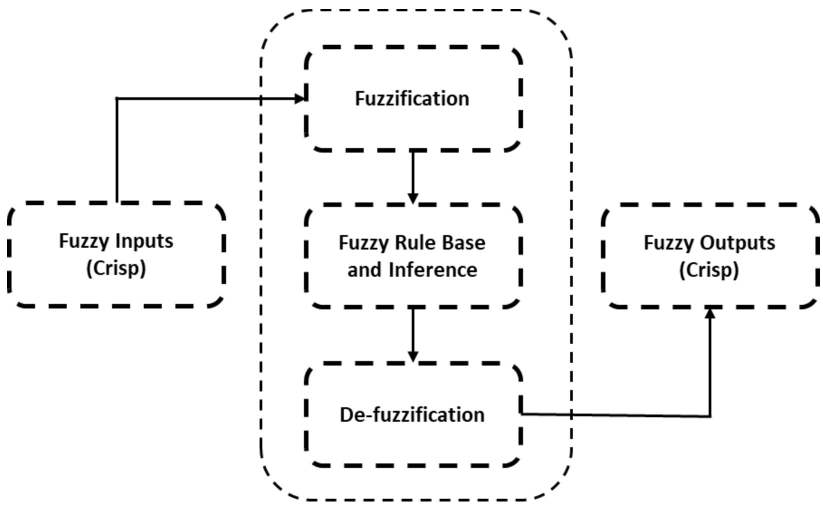


Figure 2.15: Fuzzy rule base control system structure.

Fuzzy systems also excel in integrating renewable energy sources, which are characterized by their inherent variability and unpredictability. By employing fuzzy inference mechanisms, these systems adapt to changes in environmental variables such as solar irradiance and wind speed, allowing for more stable and efficient integration into the grid. For example, FL can prioritize energy storage or direct consumption based on real-time conditions, balancing supply and demand more effectively.

In the context of energy systems, hybrid fuzzy systems have been employed to optimize the allocation of renewable resources and manage electricity consumption effectively. For example, integrating FL with ANNs enables systems to model complex, non-linear relationships between variables such as solar irradiance and energy demand. This hybrid approach enhances prediction accuracy and improves decision-making in real-time energy management scenarios [59].

Despite their strengths, fuzzy systems are not without limitations. Challenges include the selection of appropriate membership functions and the design of comprehensive rule bases, which can significantly influence system performance. Moreover, achieving a balance between interpretability and accuracy is critical. While interpretability ensures that fuzzy systems can model complex phenomena in an understandable manner, accuracy is essential for reliable decision-making. Researchers continue to refine FL methodologies to address these challenges, ensuring their applicability in increasingly complex environments [60].

The adaptability and robustness of FL systems have established them as a cornerstone of modern decision-making technologies. By simulating human reasoning and handling uncertainty with precision, these systems provide a powerful framework for managing complex, dynamic systems. Whether applied independently or as part of hybrid approaches, FL continues to drive innovation across domains, particularly in the integration and management of renewable energy systems[61].

2.4 Greenhouse Gas Emissions: A Global and National Perspective

Global GHG emissions have climbed steadily over the past decades, driving unprecedented climate change. Contextualizes these trends, illustrating global emissions by sector since 1990 and the regional distribution of emissions in 2019. Energy-related emissions (from electricity, heat, and transport) dominate the global total, followed by industry, Agriculture, Forestry and Land Use (AFOLU), transport, and buildings [62]. Notably, different regions show wide disparities in per-capita emissions. Developed regions (e.g. North America) have some of the highest per-capita emissions, whereas regions in Asia and Africa emit far less per person [63]. This imbalance underscores the varied economic and developmental contexts worldwide. The Intergovernmental Panel on Climate Change (IPCC) warns that “unabated GHG emissions pose profound risks to global stability,” linking rising emissions to more frequent extreme weather, ecosystem disruption, and socio-economic impacts [64]. Cutting emissions thus requires a concerted global effort, balancing the common but differentiated responsibilities of each nation under frameworks such as Paris Agreement.

On the world stage, New Zealand is a relatively small emitter—contributing roughly 0.17% of global CO_2 emissions. However, New Zealand’s GHG emissions profile is distinct among developed nations. Over half of New Zealand’s gross emissions originate from agriculture, mainly Methane (CH_4) from livestock and nitrous oxide from soils, while the energy sector—

including power generation and transport—accounts for approximately 37% [65]. In 2022, agriculture emitted around 41.7 million tonnes CO_2-equivalent ($MtCO_2e$), compared to approximately 28.7 $MtCO_2e$ from energy [66]. This high agricultural share is rare among developed countries: New Zealand remains the only OECD nation where CH_4 and other non-CO_2 gases form the majority of emissions [67].

Enteric fermentation in ruminant livestock—including dairy, beef, and sheep farming—is the dominant source of New Zealand's GHG emissions, accounting for approximately 35% of the national total. This makes CH_4 reduction a critical challenge for the country [68]. In contrast, CO_2 emissions from fossil fuel combustion, which are the primary focus in many other countries, constitute a smaller fraction of New Zealand's total GHG output. This is largely due to New Zealand's substantial reliance on renewable electricity sources, resulting in per-capita CO_2-only emissions of around 6.7 tonnes per person, aligning closely with the developed country average [68]. However, when CH_4 and other non-CO_2 gases are included, New Zealand's total per-capita GHG emissions rise to approximately 15 tonnes CO_2-equivalent per person, ranking significantly higher than many of its peers [69].

Since 1990, New Zealand's GHG emissions have increased by approximately 24%, primarily driven by the expansion of the dairy herd and increased road transport activity [70]. Net emissions, which account for carbon sequestration by forestry, have also risen by about 57% over the same period, influenced by fluctuations in forest harvesting and afforestation rates. To address these trends, New Zealand enacted the Climate Change Response (Zero Carbon) Amendment Act in 2019, committing to net-zero CO_2 emissions by 2050 and a reduction of biogenic CH_4 emissions by 24–47% below 2017 levels by 2050 [71].

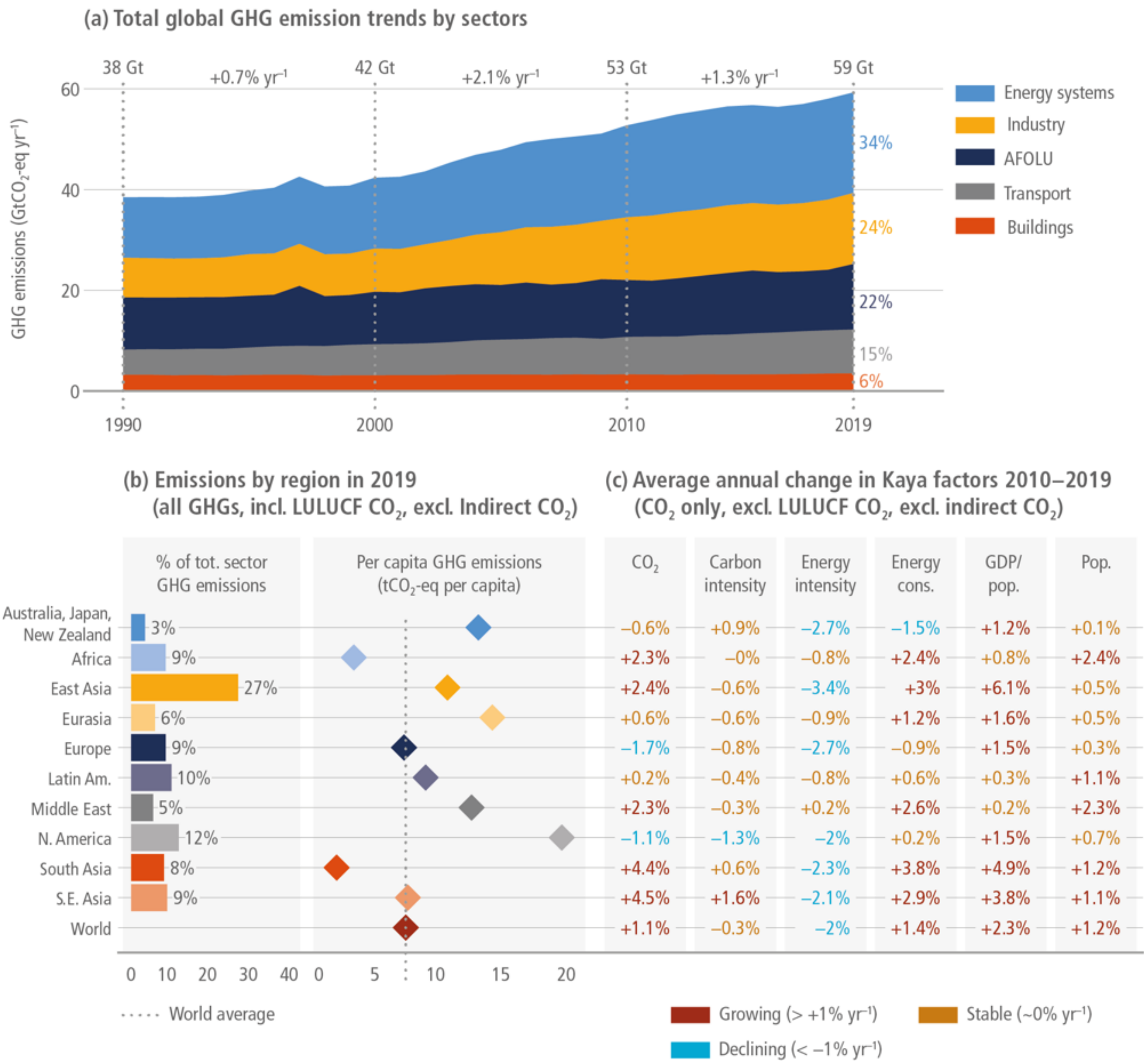


Figure 2.16: Global GHG Emissions Overview — (a) Sectoral emission trends (1990–2019), (b) Regional shares and per-capita GHG emissions (2019), (c) Kaya factor changes in CO_2 drivers (2010–2019) [72].

The implication for energy system modelling is that New Zealand must tackle a multi-gas problem – requiring innovative CH_4 mitigation (e.g. livestock feed additives, CH_4 vaccines, improved manure management) alongside CO_2 reduction. For ML applications, this broadens the scope of forecasting and optimization to include agricultural emissions and sinks (e.g. soil carbon, forestry) in addition to energy use. Models that can integrate diverse emission sources and predict outcomes of combined interventions (like CH_4 inhibitors plus electrification of transport) would be highly valuable.

2.4.1 The Broader Picture and Path Forward

Addressing climate change is a global challenge that requires both worldwide coordination and local action. New Zealand's mitigation efforts thus operate within a larger "broader picture" of international agreements, technological trends, and collective targets. Globally, the Paris Agreement aims to limit warming to well below 2°C, with a stretch goal of 1.5°C. The IPCC's latest assessment indicates that to meet 1.5°C, global CO_2 emissions need to decline about 45% from 2010 levels by 2030 and reach net-zero around 2050 [73]. For New Zealand, contributing to this path means aligning its domestic policies with carbon budgets and leveraging its strengths (high renewables) to compensate for its challenges (agricultural gases). The country has submitted progressively stronger Nationally Determined Contributions (NDCs), and in 2023 announced an intended NDC aiming for a 50% reduction in net emissions by 2030.

There is increasing global momentum in clean energy transitions. One notable trend is the worldwide push for energy sector decarbonization. Many countries are rapidly expanding renewable electricity (solar, wind) and electrifying transportation. New Zealand is in a fortunate starting position – over 80% of its electricity already comes from renewable sources (hydro, geothermal, wind). Indeed, as of 2022, about 87% of NZ's electricity was renewable, one of the highest shares on the planet. This provides a solid foundation for deep decarbonization: as other sectors (such as transport and industry) electrify, the emissions benefits are maximized if that electricity is green. New Zealand's government has accordingly set an ambitious goal of 100% renewable electricity by 2035 (with an interim 95% by 2030). Achieving this will require further growth in wind and solar generation, improved grid management, and possibly emerging technologies such as energy storage or demand response to maintain reliability [74]. From a modelling perspective, reaching 100% renewables introduces complexities around intermittency and resource adequacy – areas where ML can help. For example, ML algorithms can forecast renewable output (sun, wind) with high precision, enabling better scheduling of backup power or storage. They can also optimize dispatch in a renewable-heavy grid, learning from vast data to balance supply and demand in real-time.

Another global strategy is improving energy efficiency. Enhanced efficiency can deliver a large wedge of emission reductions by curbing energy demand growth. In New Zealand, opportunities include better insulation and heating systems in buildings, efficient industrial processes, and smart grid technologies. ML can drive gains here through intelligent control systems – e.g. ML-based building energy management that learns occupancy patterns and weather forecasts to control heating/cooling and cut waste. On a grid scale, ML can analyse consumption data to detect inefficiencies and suggest retrofits or operational changes. These innovations echo global best practices and tie into New Zealand's Emission Reduction Plan, which emphasizes demand-side measures alongside clean supply.

The "path forward" for New Zealand involves integrating these elements – renewable expansion, electrification, agriculture innovation, and efficiency – into a cohesive strategy. This must be underpinned by supportive policies (like the NZ Emissions Trading Scheme, clean energy incentives, R&D funding for CH_4 reduction) and public engagement.

Internationally, New Zealand can punch above its weight by showcasing solutions for an agricultural economy. For instance, breakthroughs in CH_4 vaccines or low-emission animal feeds could be shared globally to help other livestock-heavy nations. Collaborative research (potentially aided by Artificial Intelligence (AI) for genomic analysis or fermentation modelling) could yield scalable methods to curb CH_4 . Additionally, New Zealand participates in global climate initiatives (such as the Global CH_4 Pledge and Mission Innovation) to both learn from and contribute to the world's knowledge pool.

It's important to note that near-term actions have long-term impacts: every tonne of CO_2 or CH_4 reduced today helps limit cumulative warming. Conversely, delays lock in higher temperatures and more severe climate impacts for decades. Thus, New Zealand's pathway must balance urgent action (e.g. phasing out coal use, accelerating electric vehicle uptake) with sustained, long-range planning (e.g. transitioning to a zero-carbon agricultural model by 2050) [75]. In practical terms, this means investing in infrastructure now (such as upgrading transmission for new wind farms, building EV charging networks, and developing train/bus electrification) while also funding innovation for the harder-to-abate sectors [76]. Energy system models that include scenario analysis (e.g. "What if NZ achieves 100% EV sales by 2030?" or "What if a CH_4 inhibitor reduces dairy emissions 30%?") can inform these investments. By exploring such scenarios, researchers can identify synergistic effects (for example, widespread EV adoption could provide grid storage via vehicle-to-grid systems, enhancing renewable integration) and flag potential pitfalls (e.g. over-reliance on uncertain technology).

2.4.2 Renewable Energy Potential in New Zealand

New Zealand is richly endowed with renewable energy resources, positioning it well for deep emissions cuts through clean energy deployment. As of 2023, about 87% of electricity generation in NZ comes from renewables – primarily hydro (rivers and lakes), geothermal, and wind [77]. This renewable share has grown significantly over the past two decades, as depicted in Figure 2.17. In the early 2000s, coal and gas-fired generation played a larger role, especially during dry years when hydro lakes were low. But since then, wind power has expanded from virtually zero in 1990 to over 2,800 GWh/year in 2023 [78], and geothermal energy production nearly tripled between 2000 and 2020 [79]. Meanwhile, coal-fired generation has markedly declined (near-zero in some recent years), thanks to policies favoring renewables and the retirement or conversion of old fossil plants. Figure 2.17's long-term view also highlights the stability of hydro output – hovering around 24–26 thousand GWh annually, constrained by available sites and hydrological limits.

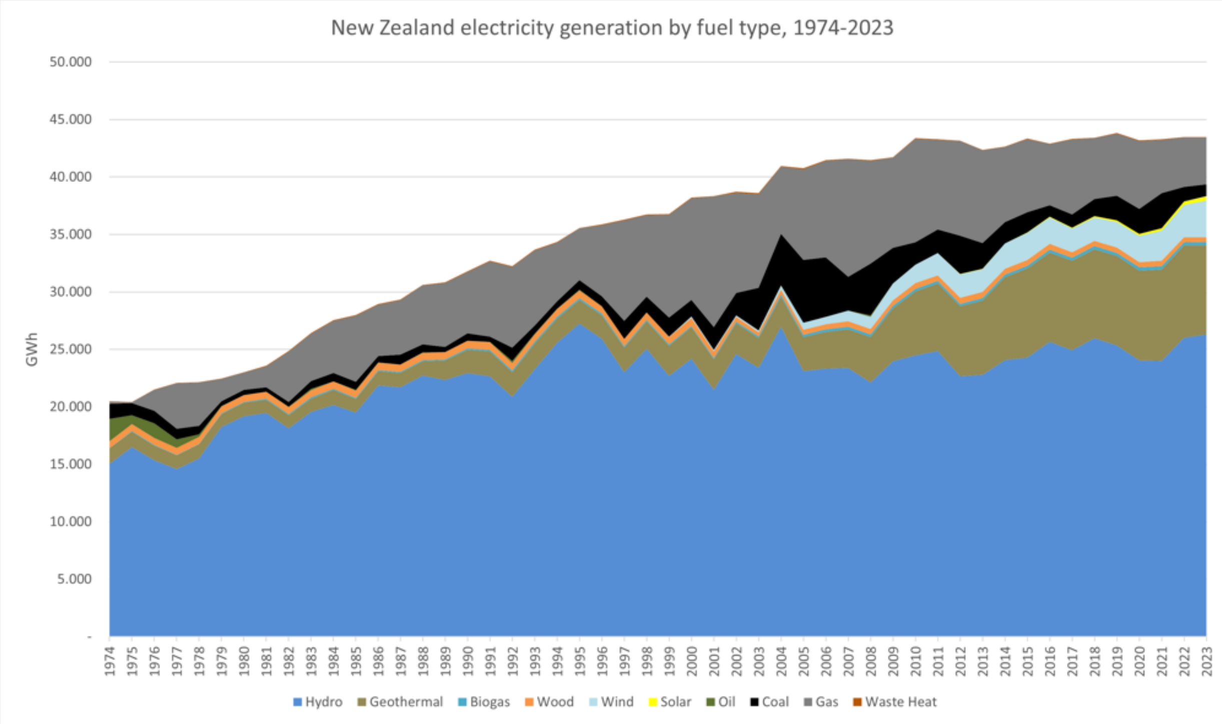


Figure 2.17: Evolution of NZ Electricity Generation by Fuel (1974–2023): Growth in Renewables, Decline in Fossil Fuels [79].

The recent growth of solar PVs, although from a small base, is notable as well: from negligible generation in 2010 to around 300 GWh in 2023 [79]. Solar uptake has been especially pronounced in the residential and commercial sectors, with over 40,000 small-scale PV installations by 2025, driven by falling panel costs [80]. Despite this progress, significant untapped potential remains across multiple renewable resources:

- Wind Energy: New Zealand's geography (long coastlines and mountainous regions) provides extensive wind resources. Average wind speeds in many areas exceed 7–8 m/s (at turbine hub height), which is considered very favourable for wind farming [81]. Most existing wind farms are in the North Island (e.g. Manawatu, Taranaki, Wellington regions) where strong, steady winds blow, and in the Canterbury plains of the South Island [82]. A 2021 study estimated that exploiting just a fraction of high-wind sites could enable New Zealand to reach its 100% renewable electricity goal with wind supplying 20–30% of generation [81]. Furthermore, offshore wind is now emerging as a frontier: the Taranaki Bight (between the North and South Islands) has consistent winds of 8–11 m/s over the ocean, with relatively shallow waters suitable for turbine installation [83]. Several companies are investigating large offshore wind projects in this area, which could each provide 1 GW or more capacity by the 2030s. For energy system modelling, the variability of wind is a key consideration – wind output can change rapidly with weather. Here, ML can assist by predicting wind power generation hours or days ahead, allowing grid operators to schedule reserves or storage. ML models can also optimize the siting of new wind farms by analysing meteorological and geographic data to identify locations with the best wind yield and minimal environmental impact [84].

- Solar Energy: New Zealand receives between ~1,100 and 1,600 kWh per square meter of solar irradiation annually, comparable to or even exceeding solar-rich European countries [85]. The North Island, especially the far north and eastern regions, enjoys the highest sunshine hours (e.g., Whangarei, Bay of Plenty) and thus the greatest PV output potential. Figure 2.18 shows a solar resource map of New Zealand, with warmer colours indicating areas of higher PV power potential (around 4 kWh per kW of solar PV per day in the best locations) [86]. This map highlights that while the South Island has good solar resource in coastal Otago and Canterbury, the North Island's subtropical latitude yields more consistent year-round sun. Importantly, New Zealand's cool ambient temperatures and high sunlight hours in summer mean solar panels can operate efficiently (solar cells often perform better in cooler conditions for the same irradiance). The government has begun to recognize Solar's promise: utility-scale solar farms are now being consented (e.g. a 150 MW solar farm in Waikato)[87], and policies are being explored to encourage rooftop solar on homes and businesses. From a grid integration perspective, Solar's daily production profile (peaking at midday) can complement wind's tendency to blow more at night in some regions. Energy system simulations show that a hybrid wind-solar mix can significantly smooth overall renewable output, reducing periods of shortfall [88]. Thus, expanding solar in NZ not only helps decarbonize the grid but also improves reliability when combined with wind – a synergy that optimization models can exploit when planning future capacity. ML can further enhance solar deployment through advanced forecasting of solar irradiance and adaptive control of distributed PV (for instance, using AI to manage smart inverters and storage so that solar-rich neighbourhoods don't destabilize local voltages) [89].

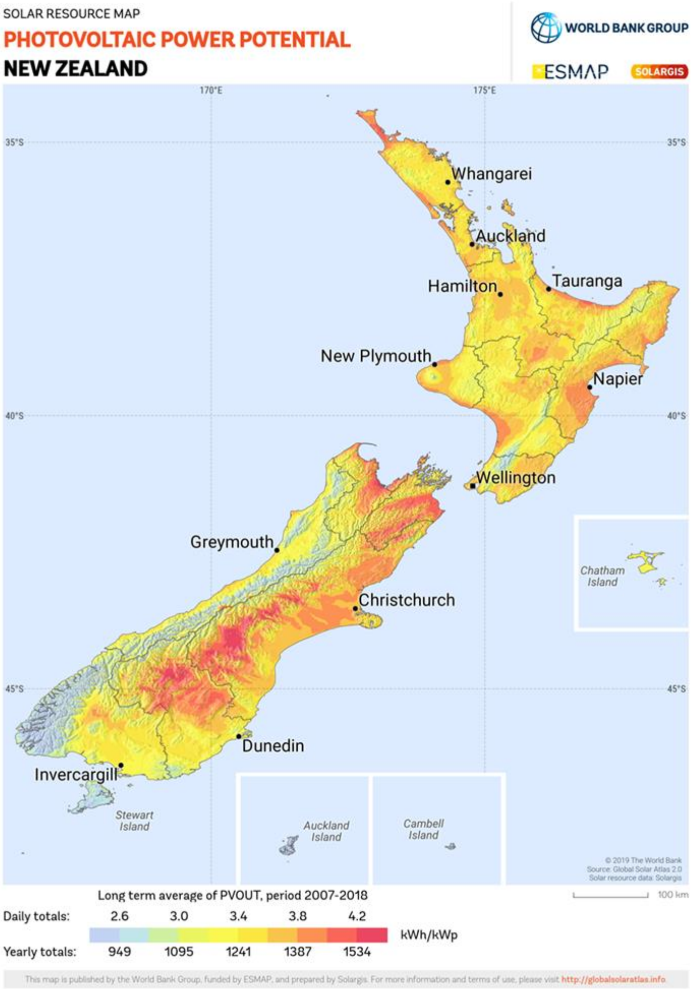


Figure 2.18: Photovoltaic Power Potential Across New Zealand (2007–2018 Average)[86].

- Geothermal Energy: Owing to its tectonic position along the Pacific Plate boundary, New Zealand possesses substantial geothermal resources concentrated primarily in the Taupō Volcanic Zone of the central North Island. High-temperature geothermal fields (150–300°C) are concentrated in this region. These fields have been harnessed for decades; as of 2023 NZ had over 1,000 MW of installed geothermal capacity, supplying around 17% of national electricity [90]. Unlike wind and solar, geothermal provides baseload generation – it can run 24/7, offering a stable backbone for the grid. Studies indicate there is still additional geothermal potential, though growth is constrained by resource consents (due to subsidence and cultural considerations of using geothermal fields) and the need to reinject fluids to sustain pressure. Nonetheless, some new geothermal projects are in development (e.g. Tauhara Stage II, ~152 MW, coming online mid-2020s) [91]. Geothermal's steady output pairs well with variable renewables in an optimized energy mix. From an ML perspective, geothermal operations can benefit from predictive maintenance algorithms (to detect equipment faults in turbines or scaling in wells) [92] and from reservoir modelling to forecast resource longevity under different extraction rates [93].
- Hydro and Other Renewables: Hydroelectricity has long been the workhorse of NZ's power system, typically supplying 50–60% of generation [94]. Most feasible large

hydro sites are already developed (e.g. Waitaki River dams, Manapouri, Clutha dams). Climate change is projected to shift patterns of rainfall and snowmelt, which could affect hydro generation seasonality. While new large dams are unlikely due to environmental constraints, pumped hydro storage is under investigation – notably a proposed scheme at Lake Onslow could add a massive 5 TWh storage reservoir to buffer dry-year risk by storing excess renewable energy and releasing it when needed [95]. Other renewables such as biomass and marine energy play smaller roles. There is potential to use more biogas from landfills and wastewater (currently ~4–5% of electricity comes from biogas and wood) [96], and research is ongoing into wave and tidal energy (New Zealand's Cook Strait has strong currents, for example)[97]. For system modelling, these diverse resources increase the robustness of a 100% renewable scenario, but they require detailed data and forecasting to integrate efficiently. ML techniques could optimize the operation of a future pumped-hydro or battery storage by learning from years of weather, load, and price data when to charge or discharge for maximum benefit.

New Zealand's expanding portfolio of renewable resources—spanning solar, wind, geothermal, and hydro—presents significant opportunities for deep decarbonization across electricity, transport, and industrial sectors. As renewable penetration rises, energy system modelling must address increasing complexity, from managing variability and seasonal shifts to ensuring resource adequacy and planning grid enhancements. Integrative modelling frameworks that incorporate both physical system dynamics and historical operational data will be critical for building a flexible and resilient energy future.

Building on this foundation, the following section explores how ML techniques are being applied across energy systems to support forecasting, optimization, and intelligent control in the context of high renewable integration.

2.5 Application of Machine Learning in Energy Systems

The integration of ML into energy systems has become increasingly significant in recent years, driven by the rising complexity of managing energy generation, distribution, and consumption under variable conditions. Energy systems are facing unique challenges: renewable sources such as solar and wind are inherently intermittent and weather-dependent, while conventional systems such as oil and gas are characterized by large-scale operations, complex regulations, and critical reliability requirements. In response to these challenges, ML offers powerful methods for modelling system behaviour, adapting to changing conditions, and providing accurate predictions that support decision-making in real-time.

This section focuses on two key applications of ML in energy systems. The first part explores oil and gas energy systems, where ML methods have been applied to enhance operational efficiency, support predictive maintenance, and optimize processes. These approaches enable operators to extract valuable insights from complex datasets, anticipate failures, and improve system resilience. Techniques such as neural networks and decision trees have shown effectiveness in modelling the intricate dynamics of these systems.

The second part examines forecasting renewable energy generation, highlighting how ML techniques can be used to predict the variability of renewable sources. Approaches such as RNNs, LSTM models, and CNNs are particularly suited for handling the spatio-temporal complexity of solar and wind generation. Accurate forecasting supports grid stability and facilitates the integration of renewable energy into power systems, ensuring more reliable and sustainable energy supply.

Together, these two parts illustrate how ML is reshaping both traditional and modern energy systems. The upcoming sections will provide detailed insights into specific ML models, their implementation, and their effectiveness in addressing the challenges faced by different energy sectors. By examining real-world applications and case studies, this section demonstrates the transformative potential of ML in enhancing energy efficiency, reliability, and sustainability.

2.5.1 Machine Learning and Oil/Gas Energy Systems

Mathematical models require many parameters to detect the output of the system. Instead of mathematical models, ML models enable fewer predictions to be made with a reduced number of inputs. The fundamental concept that underpins ML is to facilitate self-directed learning by computers via the utilization of models, as opposed to solely relying on pre-established programs designed by programmers. As new data is introduced, ML models improve in their accuracy and their comprehension of the data [98].

The previous works done in this field of research can be examined from different perspectives. First, the ML method that was used to use the input data and create the model, secondly, the goal they were looking for by using that model. As it can be seen in the Table 2.2, different objectives of the past researches have been examined, which can be categorized in the general categories of predicting environmental impacts, predicting parameters related to the energy producing sector (such as power , life, ...), predicting and estimating the parameters related to the customer sector and energy consumers (load and electricity consumption, ...) and finally environmental effects such as predicting the amount of Sox, NOx gas emissions of the class put a clause.

Table 2.2: Summary of recent learning models in energy systems.

ANN Type	LSTM	MLP	SVM
fault detection	[99]	[100]	
Load & electricity Forecasting	[101], [102], [103], [104]	[105], [106]	[106]
Improve Safety	[107]		
life prediction	[108]		
Performance Degradation and Modelling	[109]	[110], [111] ,[112]	
Heat Rate Prediction		[113]	
Environmental Impact		[114]	[115]
others		[116], [117]	

Table 2.3: Details of methods and aims about previous research about ML and Oil/Gas Energy systems.

#	Reference	Method	Aim of the research
1	Lei *et al.* (2022) [107]	LSTM	Enhance the level of safety in case of a significant failure of an operating parameter sensor.
2	Zhao and Foong (2022) [48]	Metaheuristic technique called electrostatic discharge algorithm (ESDA) is coupled with ANN	Predicting electrical power (PE) output of Combined Cycle Power Plants
3	Pachauri and Ahn (2022) [118]	Generalized Additive Model (GAM), linear regression (LR), Gaussian process regression (GPR), MLP, support vector regression (SVR), decision tree (DT), and bootstrap aggregated tree (BBT)	Electrical power prediction of a Combined Cycle Power Plants at full load
4	Arferiandi *et al.* (2021) [113]	Different ANN structures	Prediction of the Heat rate with high accuracy
5	Wood (2020) [119]	Transparent Open Box (TOB) ML method	Electrical power predictions with insight to prediction errors
6	Abbas (2020) [120]	Deep Extreme Learning Machine (DELM)	Prediction of the maximum electrical power output of a baseload power plant
7	Elfaki and AH Ahmed (2018) [121]	Feed-foreword ANN	Predicting electricity production of a combined cycle power plant with 4 inputs while considering fluctuating parameters such as hidden layer size, training dataset size, input variable inclusion, training function, and dataset normalization before network training.
8	Tüfekci (2014) [122]	ML Regression Methods, WEKA[123] toolbox	Proposing a new approach to forecast the electricity production of a continuously operating power plant at maximum capacity.
9	Kaya *et al.* (2012)[124]	Conventional Multivariate Regression, Additive Regression, KNN, feedforward ANN and K-Means clustering	Hourly prediction of electrical energy output
10	Moayedi and Mosavi (2021) [125]	Water cycle (WCA), Ant Lion (ALO) and Stain Bowerbird (SBO) optimised MLP	Assessing three effective metaheuristic techniques to precisely analyse electricity generation efficiency.
11	Siddiqui *et al.* (2021) [126]	KNN, gradient-boosted regression tree (GBRT), linear regression (LR), ANN, and Deep Neural Networks (DNN)	Using ML methods to forecast electricity production of a CCPP hourly.
12	HM Ali (2021) [127]	Regression ML with Decision Tree	Developing a model that be able to predict electricity generated from a composite cycle power plant
13	Z Chen *et al.* (2019) [128]	Multi-linear Regression, SVR, Backward propagation neural networks and CART based algorithm XGBoost	Forecasting electricity generation of Combined Cycle Power Plant (CCPP) using four diverse models.
14	Islikaye *et al.* (2018) [129]	KNN , Simple Linear, DT, Bayesian Linear, Gaussian Naive Bayes, and RANdom SAmple Consensus (RANSAC) Regression methods	A method for approximating the net energy generated by a combined cycle power plant is introduced.
15	Shuvo *et al.* (2021) [130]	DT	Assessing four ML regression methods to predict the hourly energy production of a combined cycle power plant.
16	Niu and Liu (2008)[131]	Linearization Modelling Technique	This paper utilizes linearization modelling technique to apply multivariable generalized predictive control strategy for regulating velocity and power.
17	Wankhede and Ghate (2018) [132]	ML Regression technique by using a Feed Forward Neural Networks	Forecasting electricity production per hour of a power plant solely when operating at maximum capacity.
18	Ahn and Hur (2016) [133]	ANN, Regression Tree (RT), and Multiple Linear Regression (MLR)	Proposing a more accurate prediction model for electricity generation at maximum capacity in a continuously running combined cycle power plant.
19	Raghavendran *et al.* (2021) [98]	Linear Regression, Polynomial Regression, Linear Support Vector Machine	Utilize these models to forecast electricity generation of the combined cycle power plant.

Apart from the mentioned cases, various types of energy producing systems have been investigated in the articles, among which can be independent gas turbine, system consisting of turbines and steam turbine [122], captive gas turbine [105], micro gas turbine [4, 5], … cited. Also, the software platform used in the research has been used, in some cases the use of MATLAB software [114], [135] and in a few cases the use of WEKA software [122] has been observed. In most cases, including the results in this article, the Python platform has been used. In addition to the classification mentioned earlier, Table 2.3 deals with a more detailed review of the method used in each research for the intended purpose.

2.5.2 Forecasting Renewable Energy Generation

Accurate forecasting of renewable generation (solar, wind, hydro) is essential for grid reliability and scheduling. Unlike conventional plants, renewable outputs fluctuate with weather and are inherently stochastic. ML approaches have been extensively adopted to capture these non-linear patterns and improve forecast accuracy over traditional statistical models [136]. For instance, Deep Neural Networks (DNNs), recurrent networks (LSTM/GRU), and convolutional architectures have been successfully applied to solar irradiance and wind power prediction [137] . Obiora et al. developed ML models that predict solar irradiance at multiple time horizons (minutes to days ahead), demonstrating that high-resolution input data and optimized network architectures yield precise forecasts [138]. Similarly, Gupta and Singh employed a deep learning model combining multivariate Empirical Mode Decomposition (EMD) with LSTM for hourly global horizontal irradiance forecasting, achieving RMSE improvements over baseline methods [139]. These studies underscore that careful selection of input features (meteorological variables, historical power) and data granularity strongly influence performance: finer-grained time-series data often lead to more accurate short-term predictions, whereas longer horizons may require additional contextual factors.

Recent reviews report that ensemble methods and hybrid models further enhance forecasting. For example, Pang et al. modelled combined photovoltaic-wind systems using feature-selection hybrid techniques to exploit correlations between resources, achieving superior error reduction compared to single-model approaches [140]. Deep learning has also been applied to high-dimensional sky image data and numerical weather inputs. CNNs effectively learn cloud patterns for hour-ahead solar power forecasts [141]. In all cases, literature emphasizes that time resolution and data preprocessing are critical models must be tailored to the forecast range (e.g., minutes vs days), and input data should capture physical seasonality and diurnal cycles.

Moreover, adaptive learning strategies are increasingly used to maintain accuracy as conditions evolve. SML methods update models with new data in real time, addressing concept drift in renewable patterns. For example, Xiang et al. review concept-drift adaptation methods under deep learning frameworks—such as discriminative, generative, and hybrid update modes—that enable online model updating and error-monitoring to sustain forecast performance [142].

In summary, ML-based renewable forecasting now comprises a rich toolkit (ensemble, deep, hybrid, and adaptive models) that, when combined with robust validation on historical and real-time data, significantly outperforms static statistical approaches. These improvements in forecast accuracy directly translate to better energy scheduling and reduced reliance on reserve power, facilitating higher penetration of renewables into the grid.

- Artificial Neural Networks and Wind Energy Systems

In line with the increasing adoption of ML in renewable energy forecasting, AI has also demonstrated considerable potential across broader wind energy applications. Deep learning, in particular, has emerged as a powerful tool for analysing the large and complex datasets generated by wind turbine systems. Compared to conventional ML approaches, deep learning models offer enhanced capabilities in feature extraction and pattern recognition, especially when deployed on GPU-accelerated platforms. Although these models typically involve longer training times, they consistently achieve superior performance when applied to high-dimensional operational data from wind farms [143].

ANNs, a core component of deep learning, have been widely utilized in wind energy systems for a variety of tasks. These include wind speed and power prediction, wind farm layout optimization, and fault detection and diagnosis. Accurate wind speed forecasting is particularly critical for optimizing turbine operations and improving grid integration. By learning from historical meteorological data, ANNs can provide precise short-term predictions that support dynamic control and scheduling strategies.

Additionally, ANNs have proven effective in optimizing the spatial configuration of wind turbines within a farm. Wind farm layout optimization aims to maximize total power output while minimizing wake interference and energy losses. By incorporating variables such as wind flow characteristics, terrain features, and turbulence intensity, ANN-based models can learn optimal placement strategies tailored to specific site conditions [143]. These applications highlight the versatility of deep learning techniques in enhancing the efficiency, reliability, and resilience of modern wind energy systems.

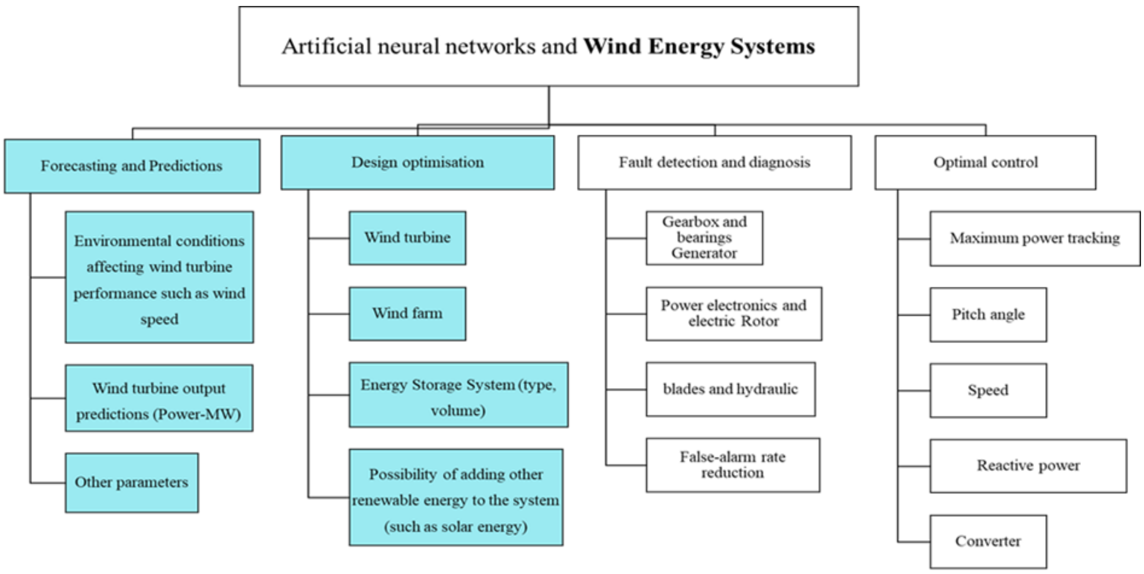


Figure 2.19: Categorising the usage of Artificial neural networks applied to Wind Energy Systems.

Fault detection and diagnosis in wind energy systems is also an important task. ANN can be trained to detect and diagnose faults in real-time, allowing for timely maintenance and repairs, which can improve the reliability and efficiency of wind energy systems.

ANNs have demonstrated strong potential in enhancing various aspects of wind energy system performance. Their ability to learn complex patterns from historical and real-time data supports a wide range of applications, including short-term forecasting, system design optimization, fault detection, and intelligent control. As illustrated in Figure 2.19, these applications contribute to improved predictive accuracy, better operational decision-making, and increased overall efficiency across different layers of wind energy system management.

- Wind Turbine Forecasting & Predictions

Energy forecasting, the energy generated from wind turbines, either individually or in the form of fields, is considered one of the important issues. This prediction and its accuracy have many benefits for the suppliers of this energy, such as preventing excessive production and production based on the estimated need [144]. Having a high-accuracy estimate of the energy production of the system, it is possible for the maintenance team to perform better, which improves the overall performance of the system [110], [112].

As previously explained, the prediction of important factors related to energy producing systems is based on the data stored in the base data. The time intervals of ten data collections for learning the model will have a high impact on the prediction of the model; So that the shorter the time intervals between the collected data, the higher the accuracy of the model resulting from learning using that data. The literature review suggests that Table 2.4 can be used to classify the forecasts based on the time intervals of data collection.

In addition to the effect of time interval on the accuracy of the prediction, the amount of data used to train the model has a significant effect on the accuracy of the prediction made in the model. In the literature review, approximately a time span of more than 6 months has been used to train the model.

In the articles related to the prediction of wind turbine parameters, two methods have been used to predict the output power of wind turbines. The first method is to predict the output power by using the output power data that is stored in memory for a long time. The energy required for your production in wind turbines is provided by the kinetic energy in the wind due to its speed [145], [146]. In the second method taking that point into consideration, wind speed is predicted using stored (long-term) information.

Table 2.4: Time horizons for prediction of wind turbine power.

Time horizon	Range
Short-term	Few minutes to 30 min
Medium-term	30 min to 1 day
Long-term	1 day to a month

And finally, they calculate the output power of the turbine by using models and mathematical equations that calculate the output power by taking the speed of the wind [121], [123]. Equation **(15)** provides a mathematical representation of the kinetic energy and power generated by a wind turbine [147]. The equation states that the kinetic energy is equal to half of the air mass multiplied by the square of the wind speed. The variables used in this equation include "E" for kinetic energy measured in Joules, "m" for air mass measured in kilograms, and "V_{wind}" for wind speed measured in meters per second. This equation is useful for understanding the amount of energy that can be generated by a wind turbine and can be applied in various wind energy applications.

$$E = \frac{1}{2} m V_{wind}^2 \tag{15}$$

Mass flow can be rewritten using air density "ρ" (kg/m^3) through an input area of wind turbine "A" as the following equation:

$$m = \rho A V_{wind} \tag{16}$$

Finally, using the replacement of the new mass flow equation in the relationship between the wind turbine's energy output and the wind speed "V_{wind}", the following equation is calculated.

$$P_{\dot{wind}} = \frac{1}{2} \rho A V_{wind}^3 \tag{17}$$

Considering the different goals of forecasting, the models offered for forecasting can be classified into two groups:

- Wind turbine prediction: In this group, the model seeks to predict the output power for a wind turbine[124], [125], [126].
- Wind farm prediction: In this case, in order to present the model, a lot of information from the wind turbines in the collection must be collected, which is more difficult than the previous case[127], [128], [129].

As explained in the section 2.2.4 (Sensitivity and correlation analysis), choosing the right number of inputs is one of the things that will have a high impact on the model's prediction accuracy, and necessarily having a high number of inputs will not lead to high accuracy in the model's output. Usually, having a large number of inputs will increase the complexity of the system [148]. The power output from wind turbine farms is one of the things that some of the researches have tried to predict and improve. In order to achieve high accuracy, this research has considered different inputs for their model. Both research conducted in [149] and [150] have used the LSTM method with the aim of generating power in wind turbine farms. The time period in which these two researchers predict the production power is classified in the short-term category. In [149], two inputs of wind speed and turbine power are considered, while in [150], in addition to Wind speed, wind direction and wind temperature are considered as input of the model. Finally [149], has achieved a lower RMSE despite the lower number of inputs, this is one of the cases where a higher number of inputs does not necessarily mean better accuracy and performance.

Emphasizing the importance of wind energy in the development of renewable energy, the random nature and uncertainty in wind turbines is considered one of the main challenges in the field of using this type of energy either individually or as a farm. One of the conventional solutions for the use of energy storage batteries, which has its own challenges, including cost and environment occupation, will be discussed in detail in section 2.62.6 (Renewable Energy Storage Systems).

A more efficient approach in this regard is to correctly estimate the energy production amount and decide on the distribution of electricity accordingly. Many methods have been used in the field of AI to estimate the production energy of wind turbines since the past. Based on the literature review, most of the recent research in this field, taking into account the characteristic of time dependence in the investigated data (wind speed, wind direction, wind angle and environmental conditions such as temperature and humidity, etc.) from the LSTM method, MLP, combining LSTM with other methods and RNN have been used.

Most of the recent research have focused on the prediction of the production power of wind turbines and have compared the LSTM method with other methods in this field. In the following, an attempt has been made to select the most complete of these articles for review. Zhang *et al.* , in [151], discussed the combination of GMM and LSTM methods for wind turbine production considering a real case study (wind farm in China). In this research, the data related to the wind speed stored in the three-month period was used as the input of the model. In this research, the authors have compared their own model with five methods of Wavelet, Elman, Radial Basis Function (RBF), Deep Belief Networks (DBNs) and Back Propagation (BP). Finally, by using the RMSE factor, they have come to the conclusion that the LSTM model is more accurate.

Zhou *et al* [148] focused on the importance of the country's grid using the LSTM model to predict the amount of energy production in wind turbines. This research has benefited from the idea of using the Autoencoder method [152] to reduce the dimensions of the input and reduce the execution time of the model. One year's collected data consisting of wind speed, temperature, humidity and ambient pressure is considered as the input of the model. The time interval between each data line in this model is one hour (Medium-term). Finally, this research compares the LSTM method with the SVM method and by using the RMSE factor, it has come to the conclusion that the LSTM model is more accurate.

Examining the performance of the proposed model for prediction in terms of Overfitting and Underfitting is a method that some researchers have used in this field. As explained earlier, model training for prediction is done by part of the data called training data. Overfitting occurs when the model makes a very accurate prediction for the training data (low RMSE), but in the case of the testing data, the prediction accuracy is visibly low (high RMSE). Usually, this problem occurs when the training data volume is too small or too large [153]. In order to avoid this error in model evaluation, most of the existing research try to use different inputs to evaluate the accuracy of the model. This input data is provided from two different directions. The first category of data is provided by different wind farm in different areas. If it is not possible to reach the first group, the data available in different time periods are used to evaluate the model.

Table 2.5: Comparison of the research done on the prediction of wind turbine parameters.

#	Reference	Method	Input Parameters	Time horizon	Contribution
1	Ahmad and Zhang (2022)[154]	sequence-to-sequence LSTM regression model	Wind speed	Long-term	Wind speed prediction
2	Chen *et al.* (2021)[155]	CNN + LSTM	Wind speed	Medium-term	Wind speed prediction
3	Lin *et al.* (2021)[156]	RNN+LSTM	Wind speed, wind power	-	Capturing the prediction fault into the output power of wind turbine
4	Dong *et al.* (2021)[157]	LSTM+ Gaussian mixture mode (GMM)	Wind speed, wind direction,	Long-term	Wind power prediction
5	Zhang *et al.* (2021)[150]	CNN + LSTM	Wind speed, wind direction, wind temperature	Short-term	Wind farm power prediction
6	Zhou *et al.* (2020)[158]	Lower and upper bound estimation (LUBE)+LSTM	Wind power dataset	Short-term	Wind power prediction
7	Yan *et al.* (2020)[159]	Deep Belief Networks (DBNs)+LSTM	Wind speed	Short-term	Wind speed prediction
8	Lan and Chen (2018)[160]	Extreme Learning Machine (ELM)+ LSTM	Wind speed	Short-term	Wind speed prediction
9	Zhang *et al.* (2019)[151]	Gaussian Mixture Model (GMM)+LSTM	Wind speed	Short-term	Wind power prediction
10	Chen *et al.* (2020)[161]	LSTM	lidar speed and direction signals (simulation)	Short-term	Yaw angle control
11	Pourmousavi and Ardehalib (2011)[162]	MLP	Wind speed	Short-term	Wind direction prediction
12	Dzulfikri *et al.* (2020)[163]	MLP	Wind speed, temperature, humidity, pressure, and altitude	Medium-term	Wind direction prediction
13	Delgado and Fahim (2021)[164]	LSTM	The theoretical power, active power, wind speed, and wind direction of SCADA system (Turkey)	Short-term	Wind speed and direction prediction
14	Zhou *et al.* (2019)[165]	LSTM	Wind speed, temperature, humidity and ambient pressure	Medium-term	Wind power generation prediction
15	Wu *et al.* (2016)[149]	LSTM	Wind speed and wind power	Short-term	Wind farm power prediction
16	Wang *et al.* (2015)[166]	Elman recurrent neural networks (ERNN)	wind speed series	Short-term	Wind speeds forecasting
17	Araya *et al.* (2018)[167]	LSTM	Wind speed	Medium-term	Wind Speed Forecasting
18	Cao and Gui (2018)[168]	LSTM	Wind speed, wind direction, atmospheric pressure and temperature	Medium-term	Wind power forecasting
19	Lu *et al.* (2018)[169]	LSTM network-based encoder-decoder (ED) model	Wind speed, air temperature, air pressure, air density and wind direction	Short-term	Wind direction prediction

Araya *et al.* [167] has used the wind speed data extracted from the wind turbine farms in the northern part of the Chile in different time intervals to evaluate its model to avoid overfitting/underfitting problem. In the Table 2.5, we have tried to classify the recently published articles from the different perspectives explained, considering the concept and method used.

- Wind Turbine Design Optimisation

The use of ANN in wind turbine and wind farm design optimization is possible because of their capability to model complex non-linear relationships and predict output from input data. In the context of wind turbine design, ANN can predict the power or efficiency of a turbine based on its design features, like blade shape and rotor diameter. By training the ANN on a large dataset of wind turbine designs, the relationship between design and performance can be learned and used to optimize the turbine design for specific wind conditions. Additionally, ANN can also be used to optimize the layout of a wind farm by predicting the power output of multiple turbines under varying wind conditions. The utilization of ANN in the optimization process results in faster and more precise design, leading to improved wind turbine and wind farm designs.

As a sample the research presented in [170] describes a method for designing a wind turbine using evolutionary techniques, which are computer-based methods that mimic the process of natural selection to find optimal solutions. The goal is to automatically generate an aerodynamic design for a wind turbine that is both efficient and cost-effective. The process involves using algorithms to evaluate and optimize various design parameters, such as blade shape, rotor diameter, and tower height. The best designs are selected based on their performance, and then used as the basis for further iterations of optimization. The final result is a wind turbine design that is optimized for specific wind conditions and can produce more energy with lower costs.

- Artificial Neural Networks and other Renewable Energy Systems

Of all the renewable energy sources available, the quickest way to turn solar radiation into electricity is through PV technology. Despite this, the output from PV systems can be impacted by various factors such as location and cloud coverage. With an increase in the number of PV plants and their significant contributions to the electricity grid, it is crucial to manage their variability effectively. Thus, accurate solar PV forecasting is essential in ensuring a dependable and cost-efficient grid operation and control [171]. LSTM networks, as previously mentioned, are structured to analyse the changing nature of data while also addressing the challenges of long-term dependencies. Because of this capability, a number of researchers have suggested different LSTM models to predict the generation of solar PV power, as demonstrated in Table 2.6. The research presented in [172] is an example that has been used in a practical way. The research paper delves into the utilization of a Long Short-Term Memory Recurrent Neural Networks (LSTM-RNN) to predict the future output of solar PV installations in South Korea.

By using this method, the study's authors aim to determine its efficiency in forecasting solar energy production in the country. The results of the evaluation indicate that the LSTM-RNN is proficient in capturing the temporal patterns present in the data, leading to a more accurate prediction of solar power generation. This case study presents a new, cutting-edge ML strategy for enhancing the precision of solar energy forecasting in South Korea.

Apart from renewable energies such as solar and wind, other renewable energies such as biogas and water (Hydroelectric Power Plant) are among the cases where AI methods have been used. In the following, in relation to each of them, sample research is examined. Rayess *et al.* [173] looks at using advanced ML, including deep learning models, to predict the amount of power that can be generated by the Almus Dam in Turkey.

The results show that deep learning models are effective in this prediction and outperform other ML techniques. This study is useful for understanding how to improve the efficiency of hydroelectric power generation in similar dams.

Rossi *et al.* [174] present a study on predicting biogas production from the dry anaerobic digestion of organic fraction of municipal solid waste (OFMSW) using a multilinear regression model. The authors develop a multilinear regression model to predict biogas production from the dry anaerobic digestion of OFMSW and evaluate its performance. The results of the study suggest that the multilinear regression model is effective in predicting biogas production and provides valuable insights into the factors that impact biogas production in dry anaerobic digestion systems.

Table 2.6: Comparison of the research done on the prediction of generation of solar photovoltaic (PV) power.

#	Reference	Method	Input Parameters	Time horizon
1	Jung *et al.* (2019)[172]	LSTM	Mean monthly temperature month of operation, relative humidity, cloud amount, duration of sunshine, wind speed, precipitation	Long-term
2	Yongsheng *et al.* (2020)[175]	Extreme Learning Machine (ELM) + LSTM	Factors that affect the performance of PV systems include solar radiation intensity, wind speed and direction, temperature, atmospheric pressure, and humidity levels.	Medium-term
3	Wang *et al.* (2020)[176]	Time correlation modification + LSTM	solar radiation, temperature	Medium-term
4	Chen *et al.* (2020)[177]	LSTM	PV output power, temperature, humidity, solar radiation	Short-term
5	Gao *et al.* (2019)[178]	LSTM	Solar radiation intensity, ambient temperature, humidity levels, wind velocity and direction, cloud coverage, as well as air pressure	Short-term
6	Mei *et al.* (2020)[179]	LSTM	PV output power, temperature	Medium-term

2.6 Renewable Energy Storage Systems

Energy storage is becoming increasingly important for ensuring the necessary reliability of electricity supply to industries and urban areas, due to the unpredictable nature of renewable energy sources. The purpose of using an energy storage device is to store this energy during the peak hours of energy production when there may not be much demand from the consumer. Also, during peak hours of energy consumption, when there may not be much energy production due to environmental conditions, this stored energy will be available to consumers. The above explanation is another expression of power system stability. Apart from these cases, frequency regulation and voltage stability are important factors in the network [180], for which the appropriate size and type of energy storage should be used. Energy storage devices can be classified as follows [181]:

1. Electrical energy storage: (I) Capacitors and supercapacitors fall under electrostatic energy storage while SMES comes under magnetic/current energy storage.
2. Mechanical energy storage: (I) Flywheels belong to kinetic energy storage while potential energy storage comprises of PHS and CAES.
3. Chemical energy storage: (I) Conventional batteries such as lead-acid, nickel metal hydride, lithium ion and flow-cell batteries such as zinc bromine and vanadium redox are included in electrochemical energy storage. (II) Fuel cells, Molten-Carbonate Fuel Cells (MCFCs), and Metal-Air batteries are part of chemical energy storage. (III) Thermochemical energy storage consists of solar hydrogen, solar metal, solar ammonia dissociation-recombination, and solar CH_4 dissociation-recombination.
4. Thermal energy storage: (I) Aquiferous cold energy storage and cryogenic energy storage fall under low temperature energy storage. (II) Sensible heat systems such as steam or hot water accumulators, graphite, hot rocks, and concrete, as well as latent heat systems such as phase change materials, are classified under high temperature energy storage.

In comparison to other energy storage methods, batteries (which store electrical energy) are the most economical choice for integrating wind farms. These batteries can not only store energy but also regulate active and reactive power at the point where the wind farm connects to the grid. As a result, they can be utilized to maintain the stability of the system. [180].

The structure used to model electric batteries for energy storage is shown in Figure 2.20. This electric circuit consists of a variable resistance and an independent current source.

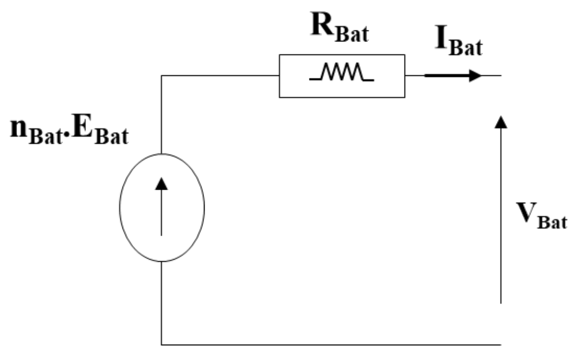


Figure 2.20: Modelling of electric batteries.

The amount of current and the amount of variable resistance and finally the amount of voltage obtained from the energy storage depend on the number of batteries connected in series with each other. The displayed mathematical relationship is used as one of the objective functions in calculating the optimal number of batteries and its appropriate volume in future works [182]. The voltage supplied by the battery used in wind farms is calculated by the following equation:

$$V_{Bat} = n_{Bat}.E_{Bat} \pm n_{Bat}.R_{Bat}.I_{Bat} \tag{18}$$

In this equation V_{Bat} is the terminal battery voltage, E_{Bat} Open circuit voltage, R_{Bat} is internal resistance, I_{Bat} is the battery current and n_{Bat} is number of series cells.

The common structure for using battery next to a renewable energy system (wind energy) is shown in Figure 2.21. A network has been provided by Bus AC/DC to supply energy to consumers. This global connection has caused all the connected components to use voltage especially for use or energy production, that's why in the above structure, Converter is used to connect each of the components.

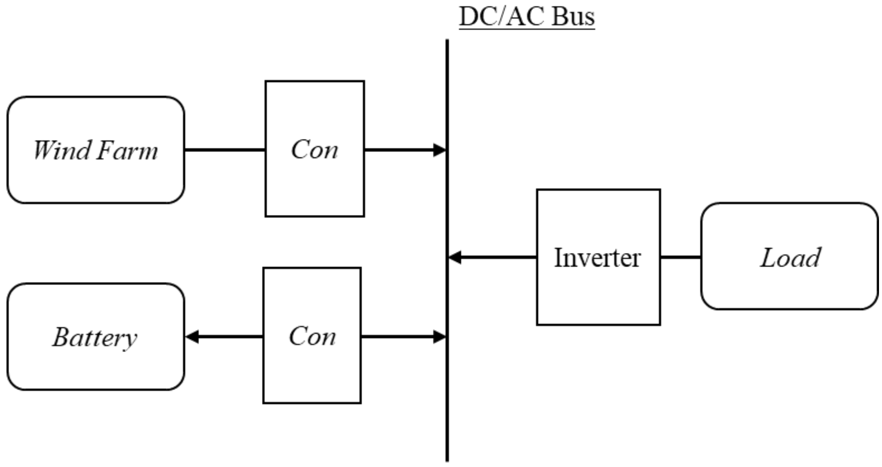


Figure 2.21: Structure of connection of battery and the grid.

In the presented structure, in order to supply the energy needed by the consumers, the energy produced by the battery and the wind farm must always have a greater amount equal to the load. The equation that is given explains this.

The above equation **(19)** represents an energy balance constraint in the context of Battery Energy Storage Systems (BESS). It states that the sum of the active power supplied by the wind turbines (N_{WT}) and the batteries (N_{BAT}) must be equal to or greater than the active power demand (P_L) at any given time "t". The equation ensures that the energy demand can be met at all times by either the wind turbine or the battery, or a combination of both. The active power balance must be maintained to ensure reliable and stable operation of the system.

$$N_{WT}.P_{WT}(t) + N_{BAT}.P_{BAT}(t) \geq P_L(t) \tag{19}$$

Apart from the limitations expressed in the form of mathematical equations to determine the optimal storage size, the following constraints can be mentioned on how to use the battery when the system needs to use the stored electricity [183]:

- Environmental: Environmental restrictions are defined as the provision of a limit on GHG gas emissions, a fine for CO emission, the factors of recycling used batteries, and the dangerous materials from battery recycling or dumping [184], [185], [186].
- Capacity: The term "capacity constraints" refers to a cap on the amount of energy that may be extracted from a battery in order to slow down battery deterioration. The capacity limitations must be established before the battery health analysis or remaining capacity can be calculated [187], [188], [189].
- Charging and discharging: The charging and discharging limitations refer to the ideal minimum and maximum limits established in an ESS model, with lesser SOC changes considering prolonging the battery's calendric life [190].

Recent studies have shown that there are several ways to determine the size of a BESS. These include load analysis, energy management, and economic analysis. Load analysis determines the capacity and power rating needed for the BESS by analysing the daily energy demand and patterns of renewable energy generation. Energy management optimizes the use of renewable energy sources and the BESS to minimize costs, while economic analysis looks at factors such as energy costs, battery degradation, and maintenance to determine the best size for the BESS.

It is also important to consider the needs of the end-user when figuring out the size of a System (BESS). For instance, the size of a BESS in a microgrid may be determined based on peak load, while in a larger utility-scale system, it may be based on daily energy consumption. Furthermore, the type of battery technology and its performance capabilities can also influence the size of the BESS.

In conclusion, determining the correct size of a BESS is crucial for achieving optimal performance and cost-effectiveness in renewable energy systems. A comprehensive approach considering various parameters, including load analysis, energy management, economic analysis, and end-user requirements, is necessary for determining the optimal BESS size [191], [192], [193].

3. Design and Implementation of a Simple Low-Cost and Real-Time Solar Power Monitoring System

3.1 Introduction

The growing interest in renewable energy systems has led to an emphasis on technologies that optimize their performance. Solar energy, being one of the most accessible and sustainable energy sources, has garnered significant attention. Its cost-effective nature, zero GHG emissions during operation, and widespread availability make it a suitable alternative to fossil fuels for electricity generation. However, the intermittent and site-specific characteristics of solar energy present challenges in system design and energy dispatch, necessitating advanced monitoring and forecasting systems [9].

Solar power monitoring systems play a critical role in addressing these challenges. They enable the real-time assessment of PV system performance, provide insights into site-specific energy production capacity, and facilitate predictive analytics for energy management. These systems also serve as foundational tools for implementing digital twins—cyber-physical replicas of PV systems—that enhance operational efficiency and support the transition to Industry 4.0 [194].

Existing research has explored various approaches to solar power monitoring. For example, Arduino Uno-based systems have been widely used for real-time voltage and current measurements in PV systems [195], [196], [197]. Some designs integrate shunt resistors for current sensing, while others incorporate Wi-Fi modules such as ESP8266 for data transmission [198]. Advanced systems employ GSM modules [198], [199] or microcontrollers such as PIC [200], enabling remote monitoring and control. More sophisticated solutions utilize Raspberry Pi platforms[201], [202], leveraging their computational capabilities for data processing and integration with cloud-based platforms such as ThingSpeak® [203].

While many reported systems focus on performance monitoring or health assessment of large-scale PV installations, there is a growing need for low-cost, portable solutions tailored to assess site-specific energy generation potential. Such systems should measure key parameters — open circuit voltage (V_{OC}), short-circuit current (I_{SC}), and ambient conditions — while addressing limitations in current sensor sensitivity for low-powered PV modules [204], [205].

This chapter introduces a novel, low-cost, real-time solar power monitoring system designed to measure V_{OC} and I_{SC} of PV modules accurately. The system integrates edge computing using Raspberry Pi and cloud-based monitoring through the ThingSpeak® platform. By addressing gaps in existing solutions, this system enables realistic assessment of energy harnessing capacity, supports the development of digital twins, and provides a preliminary step and supportive groundwork for enhanced renewable energy management.

3.2 System Design and Architecture

The solar power monitoring system is implemented using two identical 5W PV modules, the characteristics and specifications of which are shown in Table 3.1.

Table 3.1: Solar Photovoltaic (PV) Module Characteristics

Characteristics	Rating
Maximum Power	5W
Voltage at Maximum power	17.6V
Current at Maximum Power	290mA
Open Circuit Voltage	22V
Short-Circuit Current	300mA

The output terminals of one of the modules are directly connected to a voltage sensor for measuring its V_{OC}, and the other module is shunted through a 0.5Ω resistor of 5W for measuring its I_{SC}. The voltage across the shunted resistor is amplified using a non-inverting amplifier which employs an industry grade instrumentation operational amplifier (Op-Amp). The system diagram is shown in Figure 3.1.

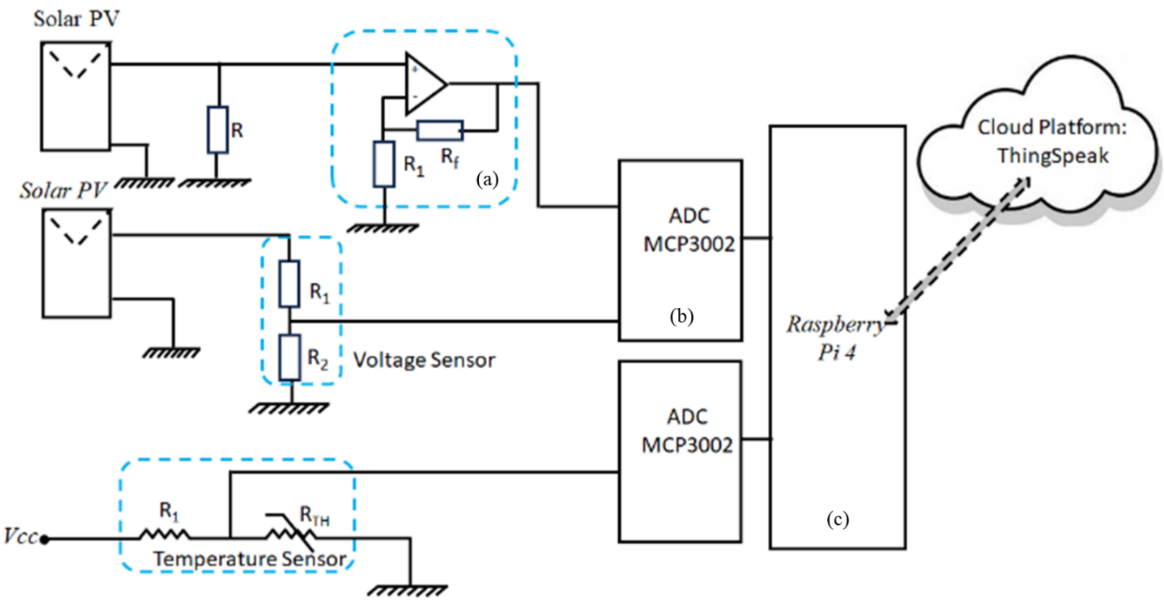


Figure 3.1: The complete solar power monitoring system, exploring an operational amplifier (a), two analogue to digital converters (b), and a Raspberry Pi processor (c).

Two channels of an Analogue to Digital (A2D) converter Integrated Circuit (IC) MCP3002 are used to digitize the sensed analogue current and voltage. One channel of a second A2D converter is utilized for the distinct purpose of ambient temperature data, adding a valuable layer to the environmental insight.

The Maximum Power Point (MPP) of a PV module can be estimated from the measured I_{SC} and V_{OC} using equation (20),

$$P_{MPP}(t) = FF \times V_{OC}(t) \times I_{SC}(t) \quad \textbf{(20)}$$

Here, FF (the Fill Factor) is a dimensionless measure of a solar module's electrical quality [206]. It is defined as the ratio of the actual maximum power output to the theoretical maximum ($V_{OC} \times I_{SC}$). Higher FF values—typically between 0.70 and 0.80—indicate lower internal losses and a more "square" I-V curve, reflecting better efficiency [207], [208]. In this system, FF is approximated as 0.75, a typical value for commercial PV panels. This relation is depicted in Figure 3.2 [209].

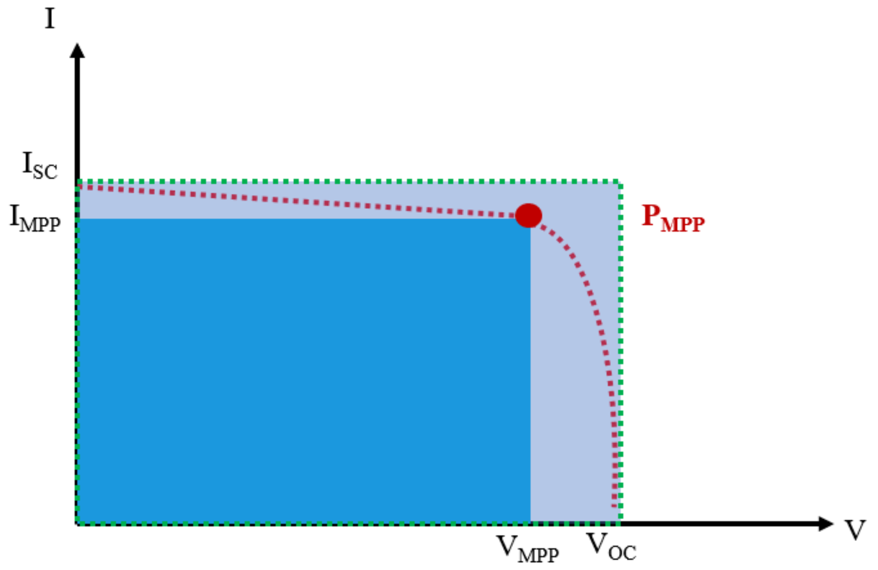


Figure 3.2: Power characteristics curve of a PV module.

This process of calculating the output power using the digital data from the A2D converter is performed using Python code on the Raspberry Pi. This code orchestrates the alternation between gathering voltage and current readings, ensuring precise and reliable computations. Moreover, as previously explained, the ThingSpeak® cloud platform [196] plays a pivotal role in this setup. It acts as the conduit through which the data seamlessly flows to the cloud, primed for further analysis and interpretation, underscoring the seamless integration of cutting-edge technology in our research methodology.

3.3 Components and Implementation Details

As depicted in Figure 3.1, this system consists of four major blocks – (i) data acquisition, (ii) A2D conversion, (iii) processing, and (iv) the cloud monitoring platform. Each of the blocks is briefly described below.

- Data Acquisition:

Three different parameters, open-circuit voltage, short-circuit current, and temperature, are to be sensed and measured. Two separate simple voltage divider circuits effectively employed to acquire voltage and temperature data. The voltage divider circuit steps down the open-circuit voltage of the solar PV module so that it can be sensed and measured by the Raspberry Pi. Figure 3.3 depicts a simple voltage divider circuit.

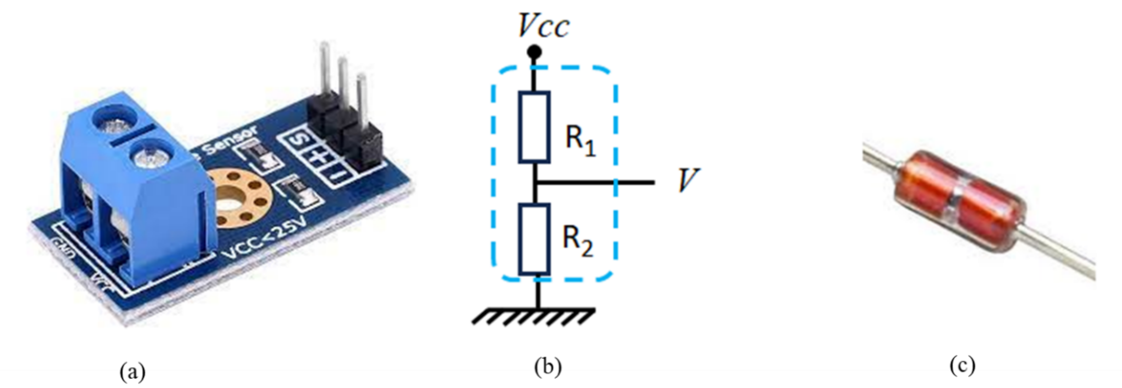


Figure 3.3: Circuit elements for voltage divider (b) with thermistor (c), standalone resistor, and amplifier interface (a).

The output voltage across R_2 is sensed in order to estimate the open-circuit voltage of a PV module. The relation of open circuit voltage, Vcc and the sensed voltage, V is expressed in equation (21).

$$V = \frac{R_2}{R_1+R_2} \times V_{CC} \tag{21}$$

In the case of temperature, the voltage divider circuit includes a thermistor in place of R_2 as shown in Figure 3.3 for temperature sensing. The thermistor resistance, R_2 is estimated for a fixed supply voltage, V_{CC} and the sensed voltage, V. Temperature at any time is given by equation (22),

$$\frac{1}{T} = \frac{1}{T_o} + \frac{1}{\beta} \ln\frac{R_2}{R_o} \tag{22}$$

where T is the temperature in °K, T_o is the test temperature which is 298°K (equivalent to 25°C), R_o is the nominal resistance of the thermistor at test temperature and β is the beta parameter, which is considered in the context of thermistor behavior and electrical conduction [210].

The short-circuit current of a solar PV module is sensed and measured by connecting it to a high power rated and very low value (0.5Ω and 5W) resistor and thus converting the current to a corresponding voltage. As the voltage corresponding to short-circuit current is very low, it is amplified using a non-inverting amplifier [211], as shown in Figure 3.4. The feedback and shunt resistors of the amplifier are selected such that the output voltage remains below the reference voltage of the A2D module i.e., 3.3V.

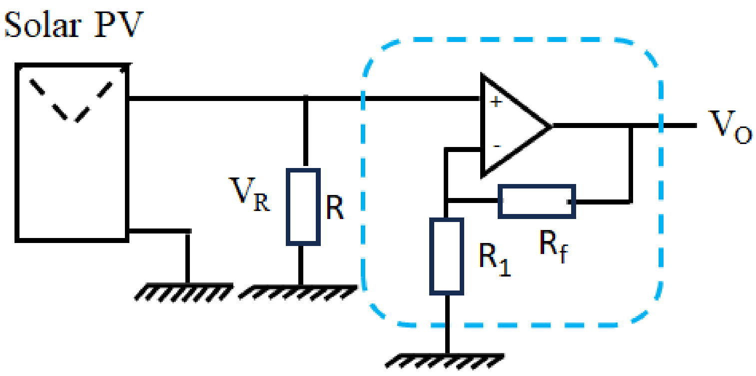


Figure 3.4: Circuit diagram of the current sensing unit

The relation of input and output voltage is expressed by the equation (23).

$$V_o = (1 + \frac{R_f}{R_1}) \times V_R \tag{23}$$

The value of the resistors, R_f and R_1 for the non-inverting amplifier are 19kΩ and 1kΩ respectively. With this configuration the maximum voltage (0.15V) across the shunt resistor, R due to maximum short circuit current (0.3A) from the PV module, results in a voltage V_O of 3V. This is below the supply voltage (5V) of the Op-Amp as well as the reference voltage (3.3V) of the A2D converter.

- A2D Conversion block:

A two-channel IC, MCP3002 is employed for the purpose of converting the analogue signal into a digital signal. It produces a 10-bit digital code and supports Serial Peripheral Interface (SPI) communication. Two of these ICs are integrated with the single SPI module of the Raspberry Pi.

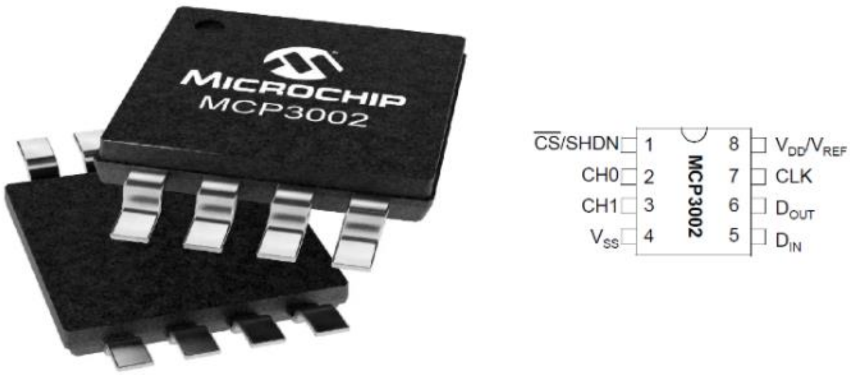


Figure 3.5: Analogue to Digital Converting Unit and its Pin-out diagram

- Processing Block:

A powerful low-cost processor, Raspberry Pi, is used as the processing unit of the system. This central processing unit forms the edge computing platform for the cloud monitoring platform. A basic Raspberry Pi board is shown in Figure 3.6.

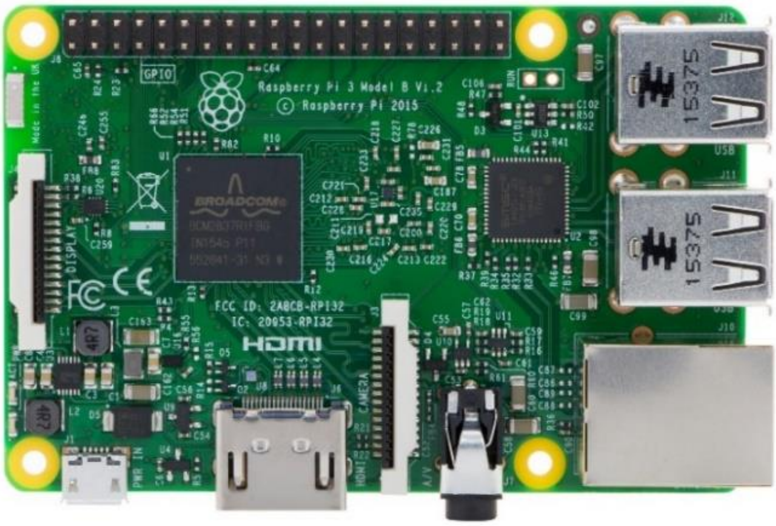


Figure 3.6: A Raspberry Pi single board computer.

The system is implemented by interconnecting all of these components on a breadboard for the purpose of testing. The implemented system is shown in Figure 3.7.

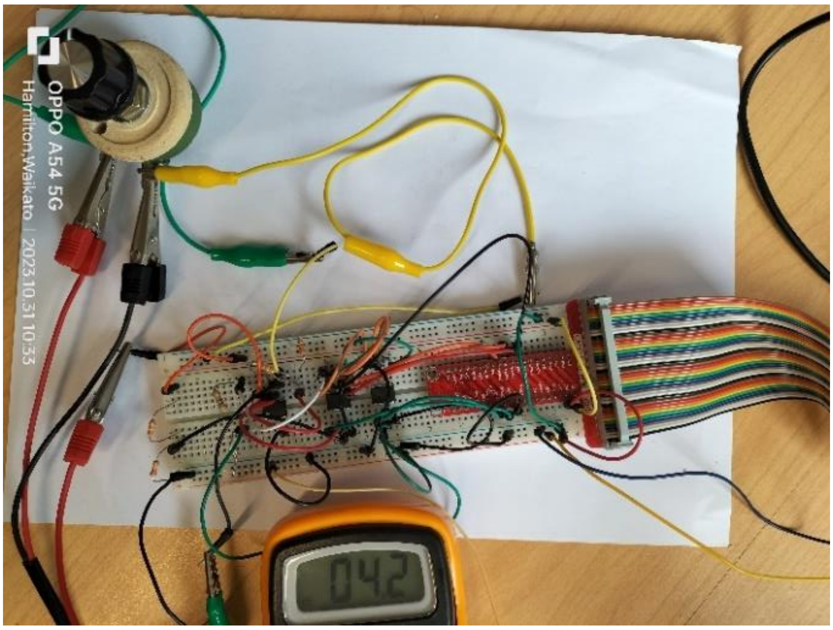


Figure 3.7: Prototype system implementation on a breadboard.

- Monitoring Platform:

The cloud platform enables further analysis, visualization and actuation of appliances, in addition to remote monitoring. ThingSpeak® is an open access platform within a certain level developed and maintained by MathWorks [212]. Figure 3.8 depicts the channel created for data logging and remote monitoring purposes.

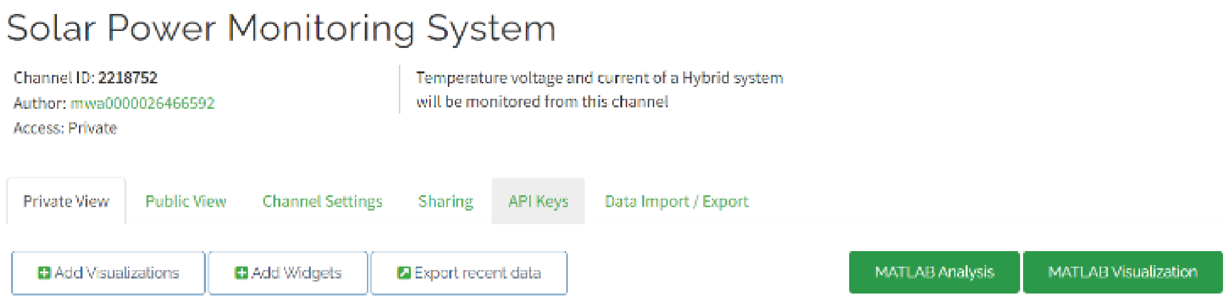


Figure 3.8: Remote monitoring system implemented on ThingSpeak® cloud platform.

3.4 Performance Analysis and Future Enhancements

The performance of the prototype system has been tested in a test environment. In this test environment, the solar PV modules were placed inclined to a window towards west direction. The test was performed on a moderately sunny day. The result suggests that the system responds well to varying conditions. Figure 3.9 depicts the monitored data on the ThingSpeak® platform.

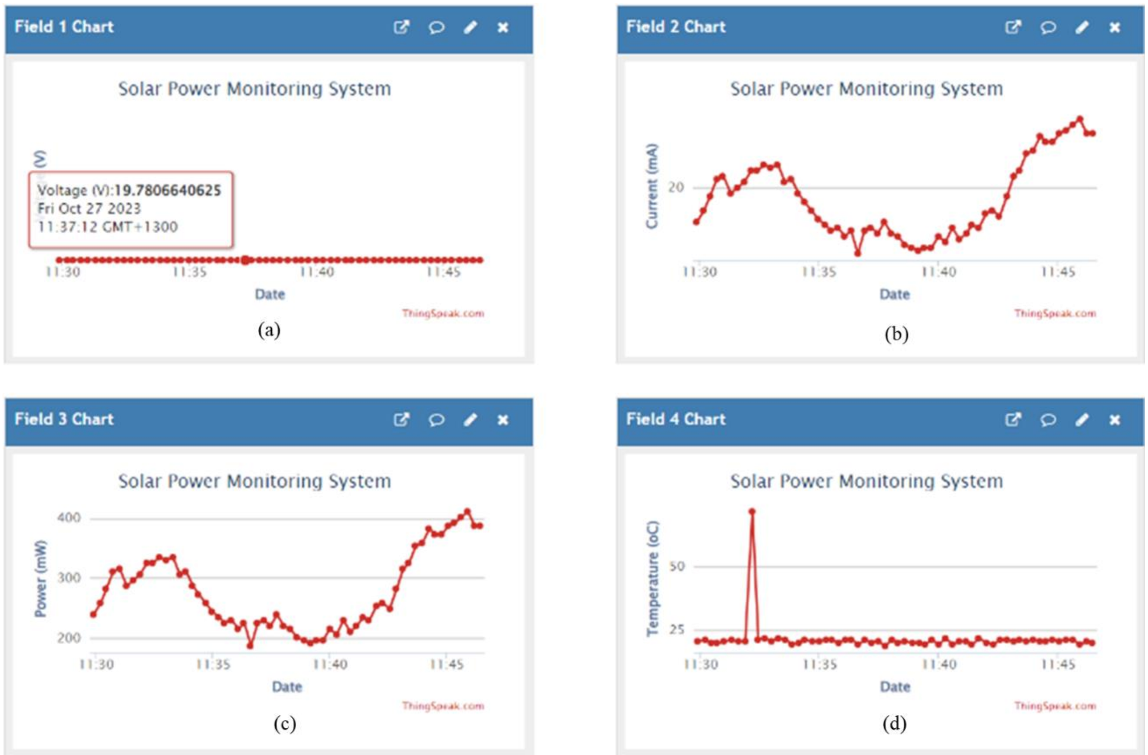


Figure 3.9: (a) Monitored Voltage, (b) Current, (c) Power and (d) Temperature on ThingSpeak® platform from a test environment.

In addition, the performance of the system was investigated under varying intensity conditions created by three light bulbs. Two of the light bulbs were of full intensity and the third one was differently coloured and with low intensity. The monitored data has clearly shown effective performance of the designed system and it is shown in Figure 3.10.

Figure 3.10: Performance Evaluation of the system under light bulb illumination, (a) voltage and (b) current.

The short-circuit current of a PV module is calculated from the direct measurement of voltage across the shunt resistor for various sunlight throughout a day. The comparison with sensed current is presented in Table 3.2. Here V_R is the measured voltage across the shunt resistor, I_R is the corresponding current through the shunt resistor and I_S is the current sensed by the Raspberry P_i.

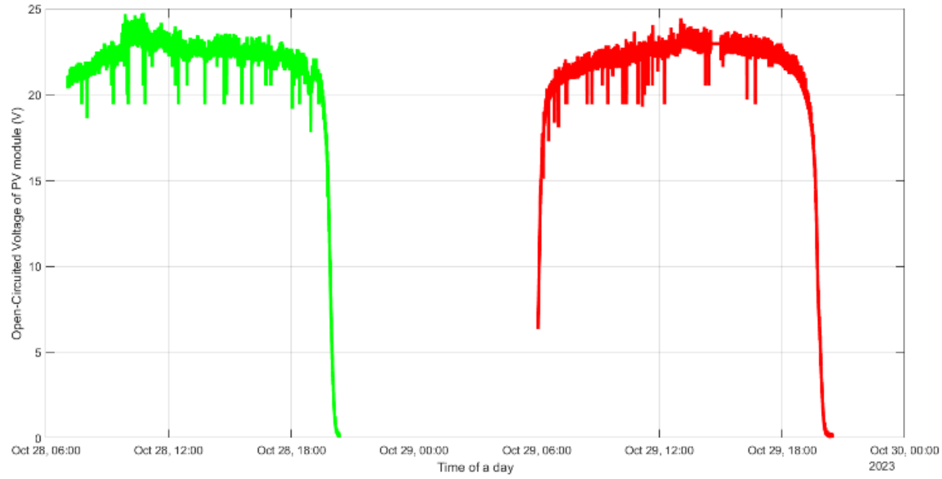
Table 3.2: Comparison of Measured and Sensed Current

Sl. No.	V_R(mV)	I_R(mA)	I_S(mA)	%Error
01.	4.0	8.0	8.70	8.75
02.	4.6	9.2	11.6	26.09
03.	5.3	10.6	12.89	21.6
04	6.6	13.2	15.79	19.62
05.	13.0	26	30.29	16.5
06.	16.6	33.2	37.38	12.6

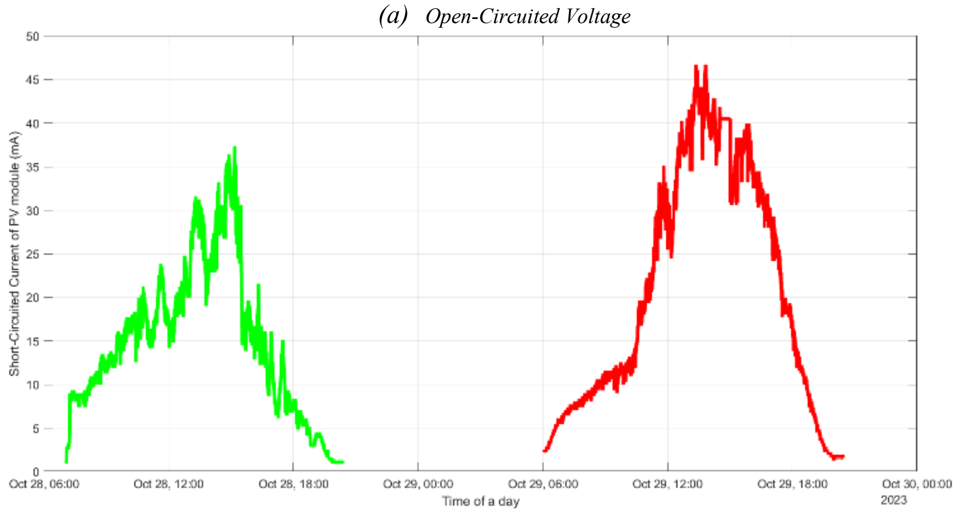
It is also noticed that there is a consistent change of voltage and current under the varying weather conditions. Figure 3.11 shows evidence of these consistent changes of voltage and current under the varying weather conditions.

Figure 3.11: Consistent change of (a) voltage and (b) current under varying weather conditions.

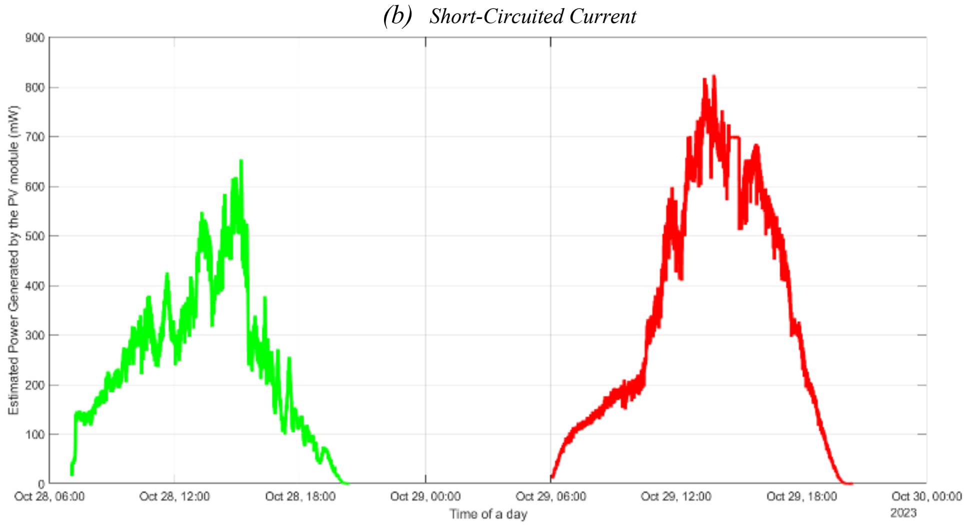
The system was tested for its performance over two days (28th October and 29th October 2023). The variation of the monitored estimated power for the sensed voltage and current of those days are shown in Figure 3.12.



(a) Open-Circuited Voltage



(b) Short-Circuited Current



(c) Real-Time Power Generation from the PV module

Figure 3.12: Monitored Estimated Power from the sensed voltage and current.

The peak power output of the system was 824mW and this was recorded on 29th October 2023. Both of the test-days were moderately sunny with intermittent clouds over the sky. However, it was a bit colder on the 29th than on 28th. This could play a role in the higher energy production on that day.

The advantage of this solar power monitoring system is evident when the monitored data is compared with an existing solar PV-battery system. In this hybrid solar PV system, the battery is the only connected load. Therefore, the monitored voltage and current data do not reflect the total power the PV panel could generate under given weather conditions. Instead, they show the actual power delivered to the battery. Since a Maximum Power Point Tracking (MPPT) system is used, the PV panel operates to extract the maximum available power based on the battery's demand and state of charge. Figure 3.13 depicts the monitored current and voltage of a solar PV hybrid system.

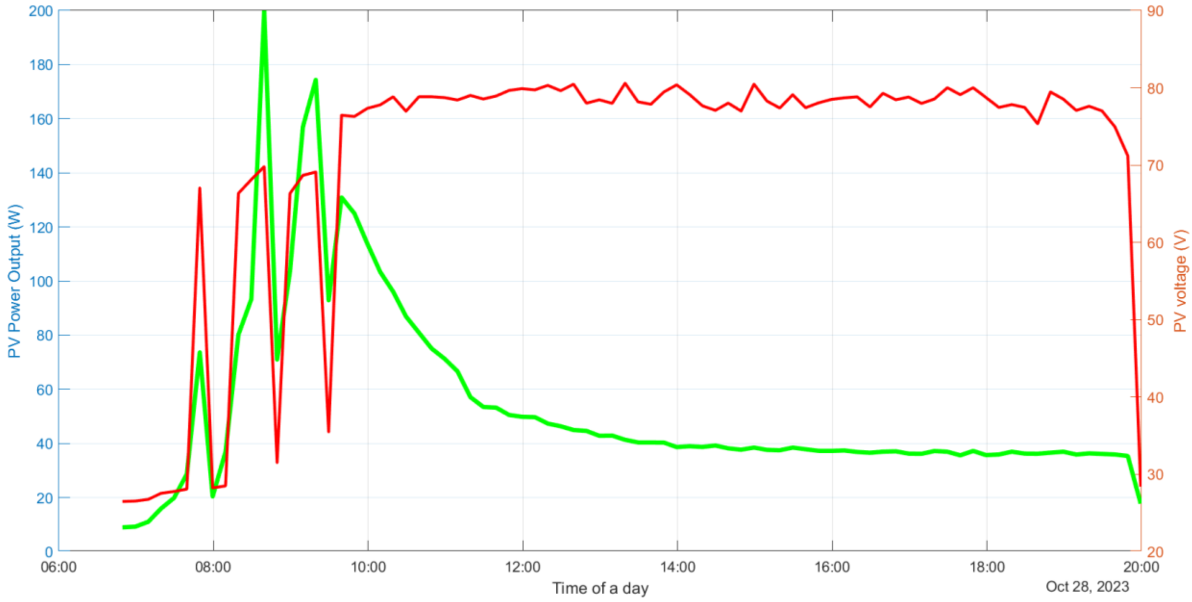


Figure 3.13: Monitored voltage and current data from the PV-battery hybrid system displayed on the ThingSpeak® platform.

3.5 Improvements and Prospects

This solar power monitoring system shows satisfactory performance in the test environment. Further testing in the field, and fine tuning of the design, could enhance performance. The performance of the system with various amplification levels for different PV modules can also be tested. As depicted in Figure 3.14 , an instrumentation amplifier and a current-to-voltage converter could be evaluated as alternatives or additions to the non-inverting amplifier to potentially improve signal quality and measurement accuracy. The figure also includes a voltage divider circuit, which is used to step down the amplified voltage to levels compatible with the A2D input of the Raspberry Pi.

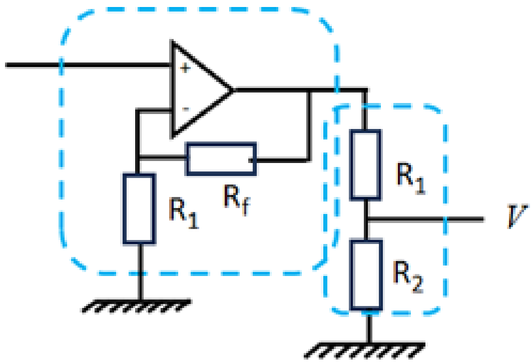


Figure 3.14: Voltage divider circuit added after the non-inverting amplifier.

Additionally, the system can be improved by using a single PV module instead of two, by enabling sequential measurements of open-circuit voltage and short-circuit current via switching mechanisms. Achieving greater sophistication in the temperature sensor module and integrating additional weather data, such as humidity and wind speed, would enable more comprehensive environmental monitoring. Based on this prototype, developing a fully independent unit with its power system would also facilitate its use in assessing solar and wind power generation for a wide range of sites. Beyond these applications, the designed system has the potential to serve as a foundational tool for creating digital twins virtual models that replicate the behavior and performance of solar PV and wind generation systems.

In this part of chapter, the primary objective was to design and implement a cost-effective, low-power (5W) solar power monitoring system. By utilizing a low-powered PV panel in conjunction with a current-to-voltage converter, the system avoided the complexity of MPPT, instead estimating the MPP using the FF and measured voltage and current. This approach enabled efficient monitoring without compromising accuracy, using two identical 5W PV modules to measure short-circuit current and open-circuit voltage.

The system developed in this case study provides real-time data-driven insights to support improvements in renewable energy management. Its role could potentially be extended by integrating AI and predictive modelling, enabling more dynamic analysis and optimisation. Looking ahead, these findings suggest a pathway for future research aimed at enhancing sustainability in industrial energy systems. By combining digital monitoring with renewable energy solutions, this system illustrates a potential step towards cleaner energy practices within the tested context.

4. Data-Driven Predictive Modelling of Steam Turbines: Insights from Monitoring Systems

4.1 Introduction

Building on the supportive groundwork of Chapter 3, which illustrated data collection and real-time monitoring systems, this chapter explores how such data can be utilized for predictive analysis. Data monitoring systems not only provide critical insights into real-time performance but also enable the development of predictive models that enhance decision-making and optimize energy systems. This transition from monitoring to prediction highlights the importance of leveraging high-quality data for advanced forecasting techniques, applicable across various case studies and energy systems.

In this chapter, the focus shifts to the prediction of power generation in a two-stage back pressure steam turbine used in the paper production industry. Accurate prediction of power production is critical in industrial applications to optimize energy utilization and reduce reliance on traditional energy systems. Unlike solar energy systems discussed in Chapter 3, steam turbines operate within a thermodynamically complex framework, making prediction a challenging task. Traditional mathematical and thermodynamic models, while effective in certain scenarios, often lack the adaptability and accuracy required in dynamic industrial environments. As part of this research, the potential of advanced ML techniques, specifically LSTM, is explored to address these challenges.

Advancements in AI have revolutionized predictive modelling, providing tools that can effectively handle time-dependent data, such as the inputs required for steam turbine power generation. ML methods, particularly supervised learning, offer significant advantages for this purpose. When dealing with temporal datasets where sequential order is crucial, specialized techniques such as RNNs, CNNs and LSTM models outperform conventional methods [20]. These models are capable of capturing intricate relationships within time-dependent data, making them particularly well-suited for predicting the output of steam turbines.

In this study, an LSTM-based predictive model is developed to forecast the power production of a two-stage back pressure steam turbine. The research begins with the collection and refinement of input data, ensuring only high-quality, relevant parameters are considered. Correlation analysis is performed to identify the most significant inputs, reducing the initial dataset from nine parameters to four. This refinement not only improves model performance but also emphasizes the importance of input selection in predictive analytics. The performance of the LSTM model is compared against the Willans Line Model — an empirical method that relates turbine power output to steam consumption under idealized isentropic conditions [12] — a traditional thermodynamic approach. By utilizing the RMSE as the evaluation metric — previously discussed in Section 2.2.3 — the LSTM model demonstrates superior predictive accuracy, underscoring the potential of AI-driven methods in the energy industry.

The integration of predictive analytics into industrial energy systems offers transformative potential. While Chapter 3 emphasized the importance of data collection and monitoring as foundational steps, this chapter illustrates how such data can be leveraged for advanced prediction and optimization. By demonstrating the applicability of AI techniques across different energy systems, this research bridges the gap between real-time monitoring and future-oriented predictive analysis. Furthermore, the comparison with thermodynamic models provides valuable insights into the evolving landscape of energy prediction methodologies, highlighting the adaptability and precision of ML approaches.

Through the exploration of predictive models and their applications, this chapter contributes to a comprehensive understanding of how data-driven techniques can enhance the efficiency and reliability of energy systems. By extending the methodologies introduced in Chapter 3 to predictive analytics, this research underscores the versatility and value of data in advancing energy solutions across diverse contexts.

4.2 Description of an Industrial Process

For this research, a two-stage steam turbine at a pulp and paper mill using the Kraft process has been used as a case study (see Figure 4.1). The power production of this steam turbine is approximately 40 MW. An hourly time series dataset containing 22 months has been obtained from industry, in order to estimate the output power with high accuracy, low computational complexity, and low execution overhead, compared to Willans line model. Figure 4.1 shows inputs and output of the model.

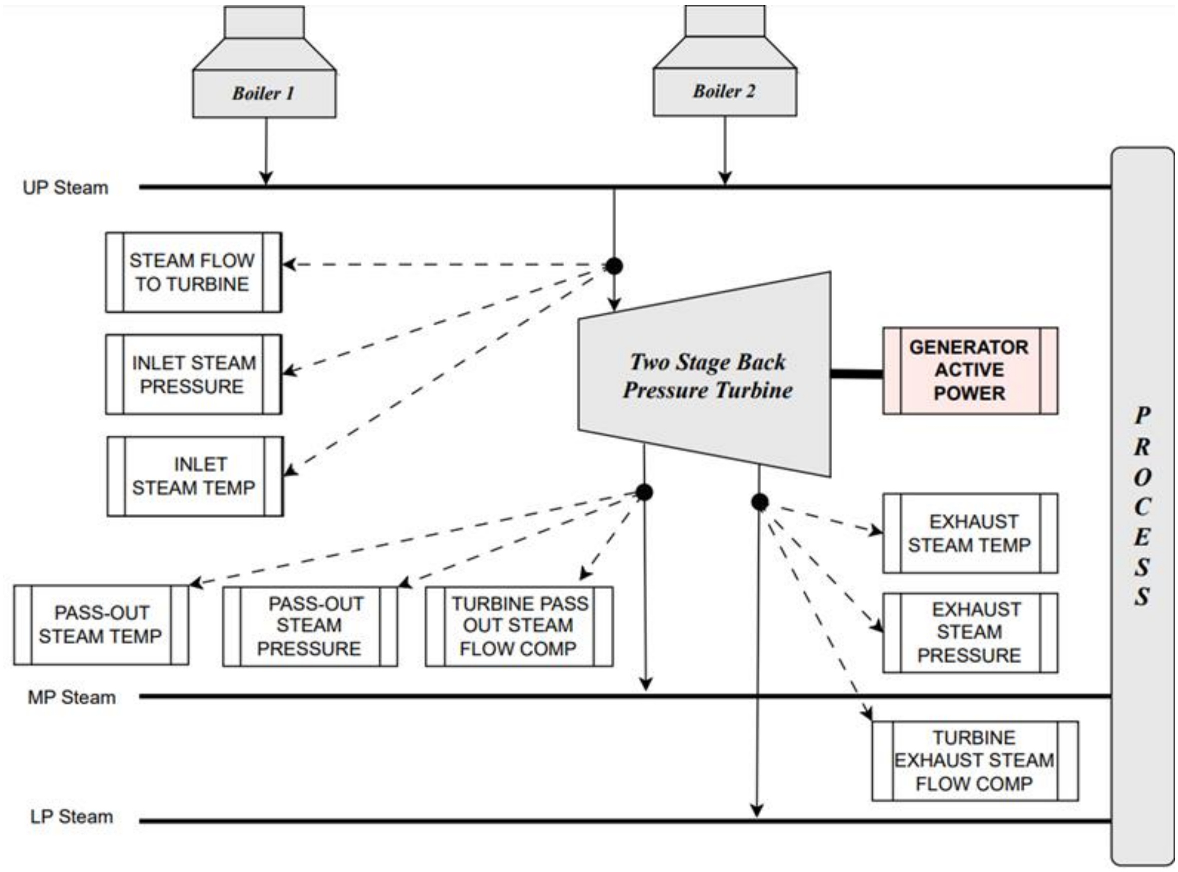


Figure 4.1: An overview of the two-stage steam turbine process and its inputs and outputs.

Input data is recorded at ten-minute intervals, although some portions of the dataset are incomplete. To address this issue, linear interpolation was utilized to fill in the missing data. To improve the accuracy of predicting generator active power, this software framework employs an LSTM model. The research described here considers different inputs as shown in Figure 4.2, with an architecture comprising 50 neurons in LSTM layer having an activation Relu (Rectified Linear Unit) activation function, which introduces non-linearity and helps mitigate the vanishing gradient problem during training [213]. The ADAM (Adaptive Moment Estimation) optimizer [214] has been included in the model to minimize errors, as it efficiently combines the benefits of momentum and adaptive learning rates to accelerate convergence during training. The model was trained using 40% of the available data, and the remaining 60% was utilized to test the model.

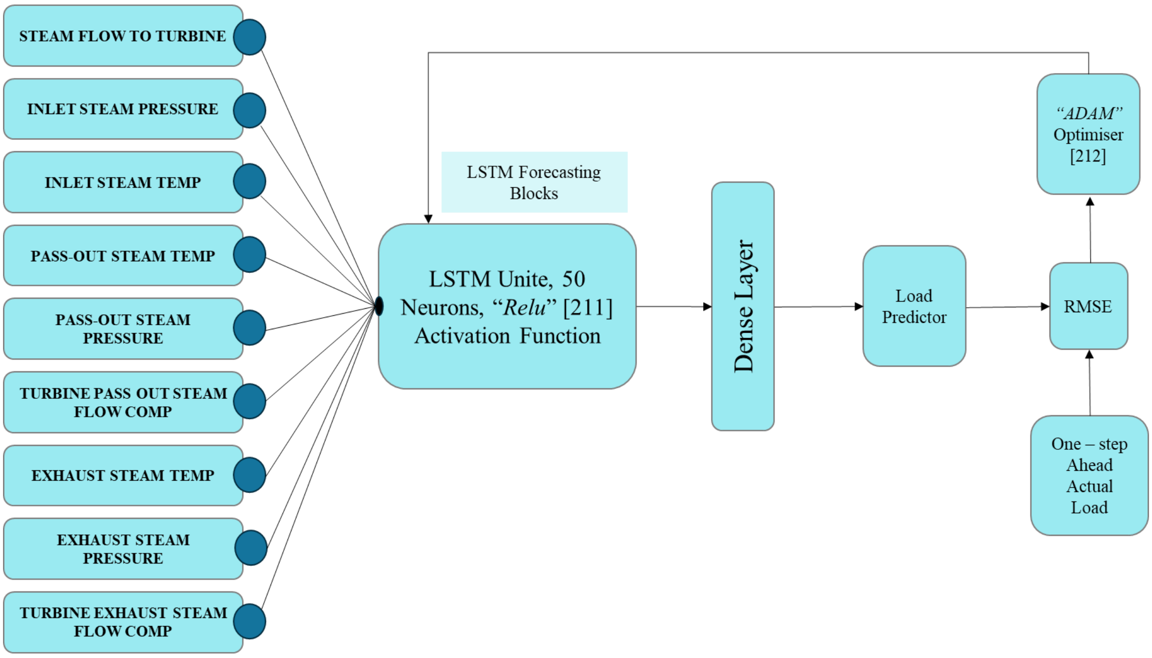


Figure 4.2: Steam turbine power prediction model architecture.

Correlation analysis plays a vital role in understanding the relationships between input parameters and their impact on the output, offering valuable insights into the dynamics of the system. This analysis evaluates the degree of interdependence among variables, helping to identify which inputs have the most significant influence on the target variable. A detailed explanation of the correlation methodology and its equation have already been provided in Chapter 2. As depicted in Figure 4.3, the correlation matrix of the input data for the steam turbine highlights these relationships, providing a visual representation of the varying degrees of correlation among the parameters. This figure emphasizes the importance of selecting highly correlated inputs to enhance model performance.

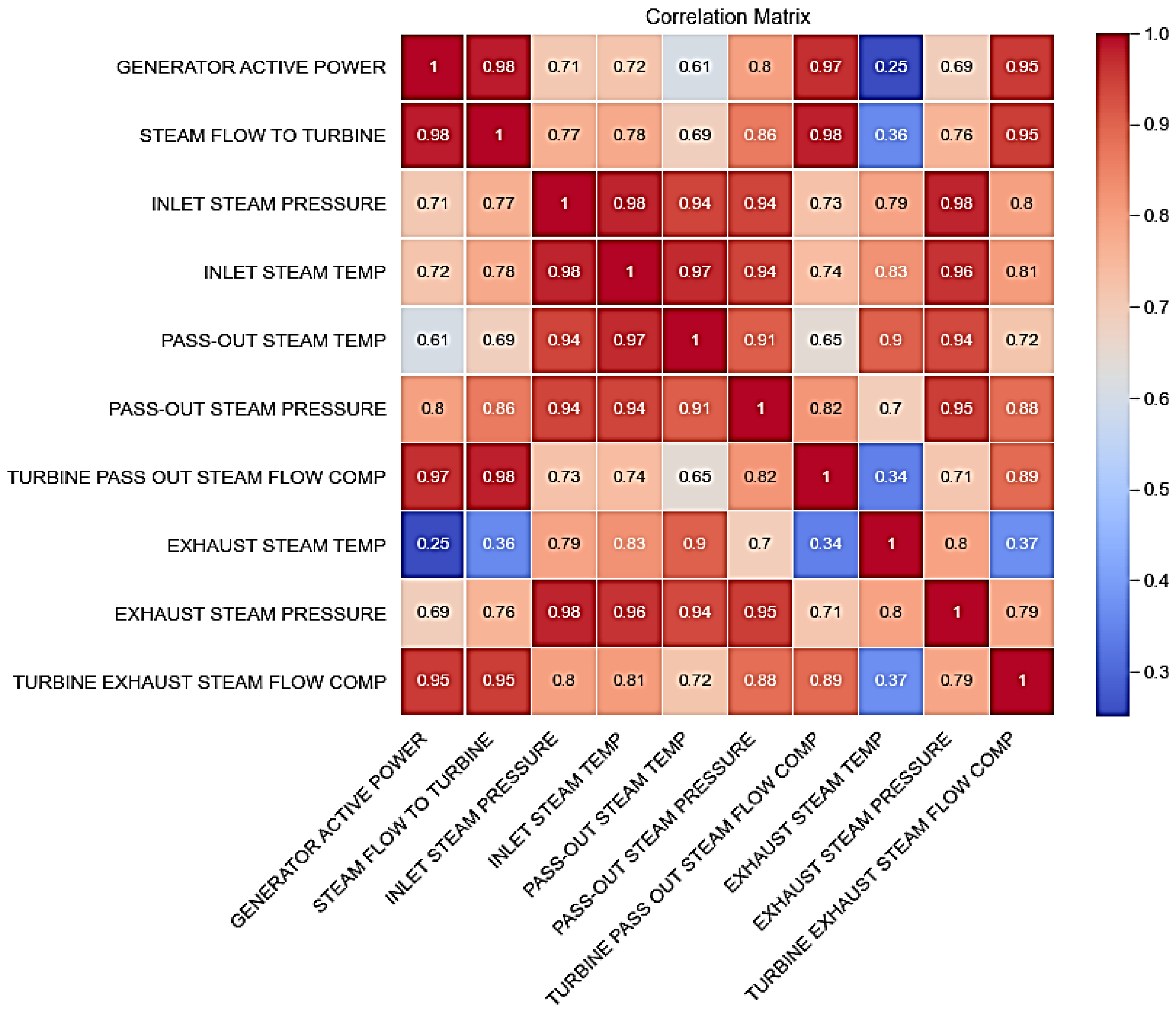


Figure 4.3: Correlation matrix of input data from the steam turbine.

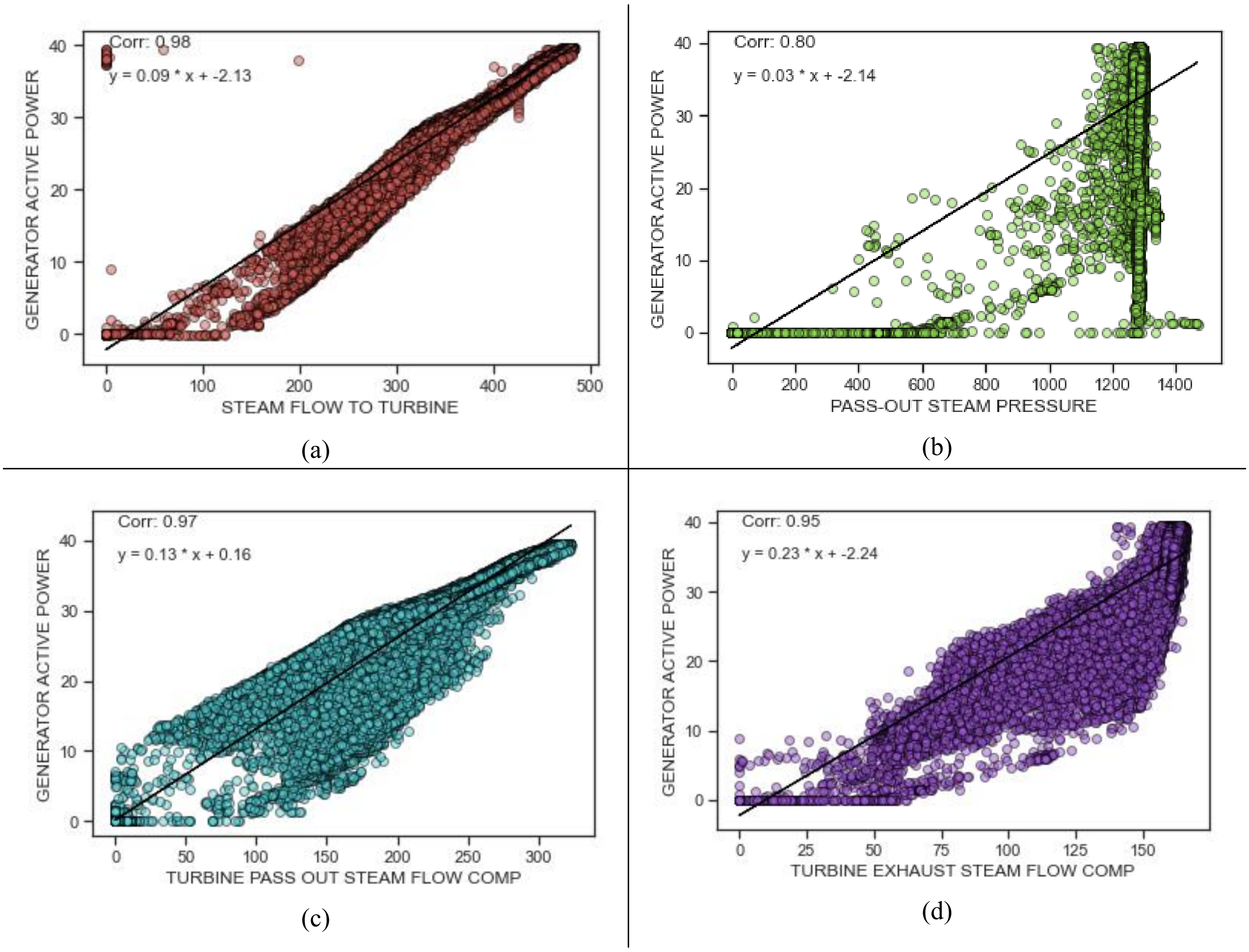


Figure 4.4: Correlation Analysis of (a) Steam Flow to Turbine, (b) Pass-out Steam Pressure, (c) Turbine Pass Out Steam Flow Comp, and (d) Turbine Exhaust Steam Flow Comp with Output Parameter, with correlation factor bigger than 0.8.

To provide a better understanding of the intricate relationships among various parameters, Figure 4.4 depicts the magnitude of influence that each parameter exerts on the others if the correlation factor between them is bigger than 0.8. Upon careful examination of these graphs, it becomes apparent that when the correlation parameter surpasses a value of 0.8, the parameters exhibit a strong association. In such cases, a line of best fit is provided as a first-order approximation to illustrate the trend between them. However, this does not imply that the relationship is strictly linear, as in Figure 4.4 (b), where the scatter pattern indicates a more complex underlying behavior despite the high correlation value.

Furthermore, in the scope of this research, the focal point revolves around the generator's production power, which is considered the ultimate output. Considering this, a careful selection process was carried out to identify pertinent inputs for the second model (main model after applying correlation analysis) under evaluation. Specifically, as shown originally in Figure 4.4, four parameters, Steam Flow to Turbine, Pass-out Steam Pressure, Turbine Pass Out Steam Flow Comp, and Turbine Exhaust Steam Flow Comp, were found to exhibit a correlation parameter exceeding 0.8 when compared with the output parameter. Consequently, these four parameters were designated as inputs for the second model in this study, as their influential impact on the output power warranted further investigation and analysis. Importantly, excluding weak or irrelevant inputs not only simplifies the model but also improves generalization by reducing noise and mitigating the risk of overfitting, a finding well established in the feature-selection literature [215],[216].

4.3 Willans Line Model and Equations for Power Generation Estimation

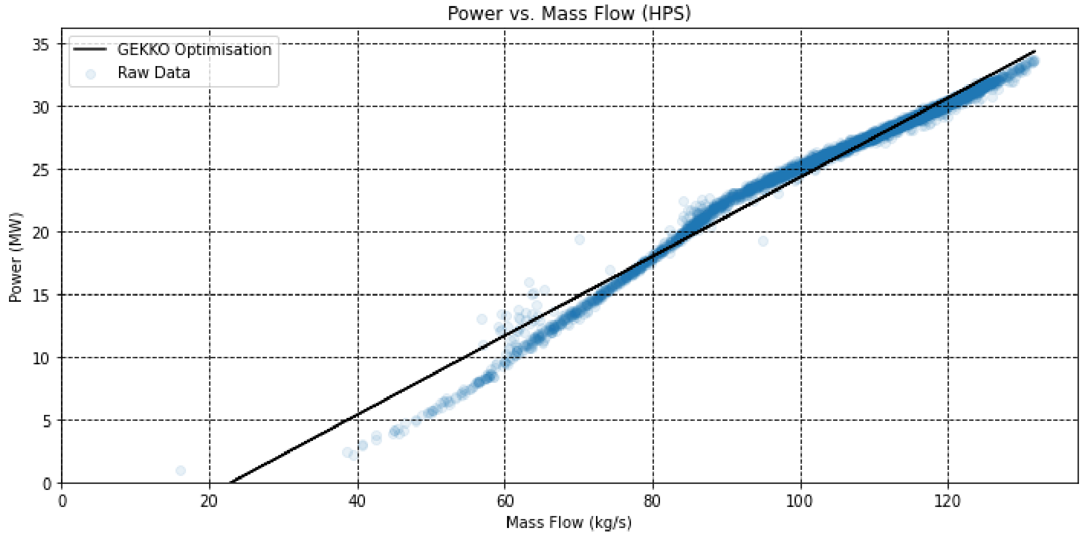
The Willans Line [12], a concept in steam turbine engineering, serves as a guiding principle in the evaluation of the two-stage back pressure steam turbine. Developed by Robert Willans, this graphical representation provides essential insights into the efficiency and performance of the turbine across varying operating conditions. The line establishes an idealized correlation between power output and steam consumption for an isentropic turbine. By comparing the actual power generated with mass flow to this theoretical line, a clear understanding of the turbine's real-world efficiency is gained [12]. The Willans Line also aids in identifying the most efficient operating range, where the turbine achieves optimal power output while conserving steam. It is a critical tool for ensuring our steam turbine operates at its highest efficiency, minimizing resource consumption and maximizing power generation.

The thermodynamic model employed in this study utilizes the Willans Line to estimate the generated power of the generator. The calculations are performed using the "*CoolProp*" software library [217], which provides functions to calculate properties such as enthalpy and entropy of the working fluid. The model considers multiple stages of the turbine, including the High-Pressure Stage (HPS) and the low-pressure stage (LPS).

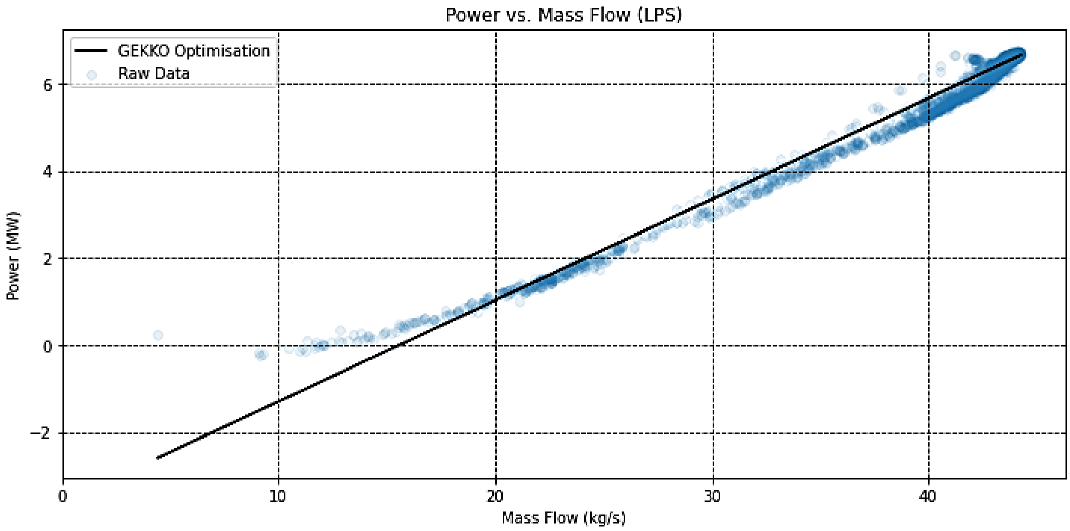
In the implemented approach, the PropsSI function from the "*CoolProp*" library is used to calculate the enthalpy and entropy of the working fluid based on the given temperature and pressure values. These calculations are performed for the HPS stage, and the resulting values are stored. Similarly, the enthalpy of the medium-pressure stage is derived from the HPS entropy using the entropy-enthalpy relationship, and the entropy of the MPS stage is calculated directly. The work transfer between stages is then computed by multiplying the mass flow rates with the differences in enthalpy. This results in the calculation of the work done in the HPS stage and the LPS stage. The total generated power is obtained by summing up the individual work transfers.

To optimize the regression of the raw data, the "*GEKKO*" software library [218] is utilized. It provides capabilities for regression analysis and optimization. The implemented model performs regression analysis using the "*GEKKO*" software library to find the best-fit parameters for the power-mass flow relationship. The resulting parameters are used to plot the regression lines.

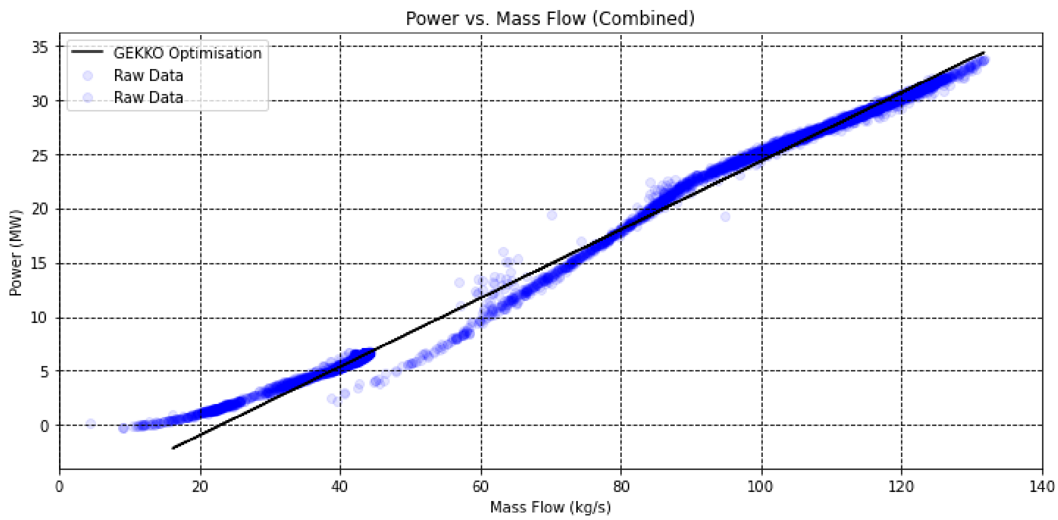
Figure 4.5 shows the different turbine stages, specifically comparing the modelled Willans Line with the raw data. These plots show the performance of the turbine over varying operating conditions, as well as how closely the model matches the raw data. For both the higher-pressure stage (a) and the lower-pressure stage (b), there is a close correlation at higher mass flows, but begins to deviate away at lower pressures. While this type of Willans Line can be useful for estimating the work generation under various loads, it is difficult to validate the outputs of these individual stages as often, multistage turbines operate with single, common generators, meaning that individual power values for stages are not available, and instead only the overall power.



(a)



(b)



(c)

Figure 4.5: (a) High-Pressure Stage (HPS) Steam Flow vs. Power Output, (b) Low-Pressure Stage (LPS) Steam Flow vs. Power, (c) Combined HPS and LPS Steam Flow vs. Power.

4.4 Prediction and Performance Analysis

This section presents the results obtained from the implementation of the model, providing a comprehensive understanding of its performance. Figure 4.6 depicts the outcomes, revealing that a significant proportion of the predicted points generated by the model, whether from the testing data or the training data, coincide with the actual data points. This alignment suggests relatively low error between the model's predictions and the real data, thus indicating that the model performed well under the tested dataset.

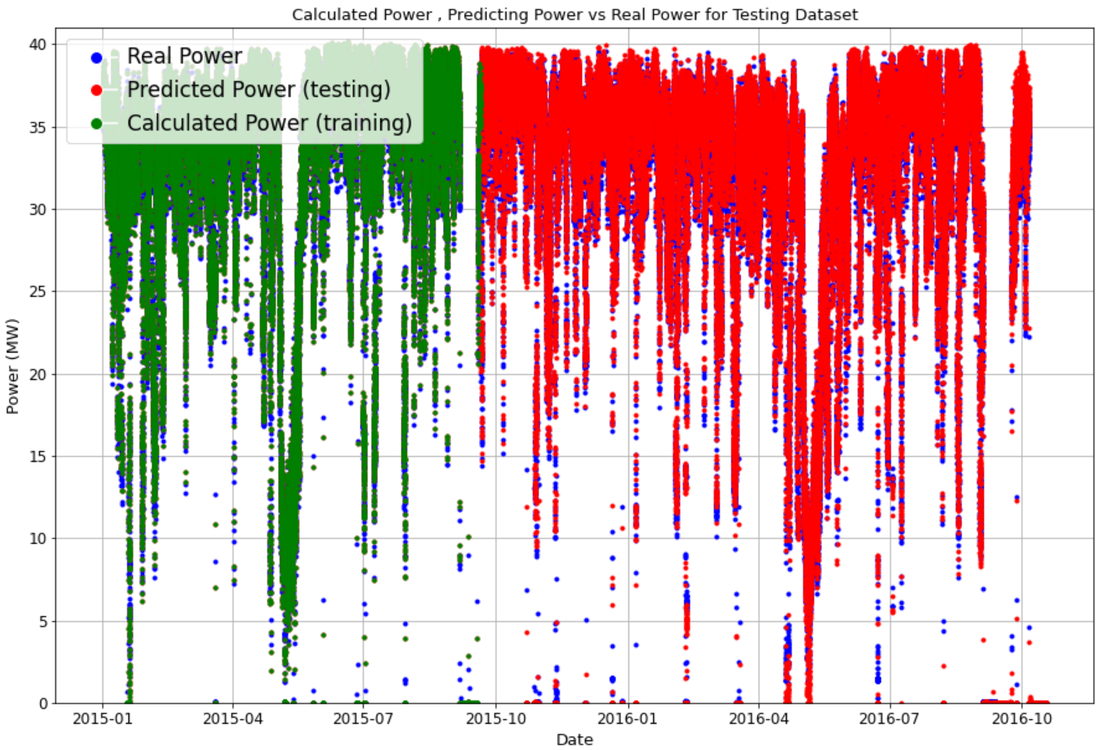


Figure 4.6: A comparison of real and predicted points of the steam turbine power prediction model.

Prior to arriving at this stage, an extensive investigation was conducted using sensitivity analysis to develop two models, each comprising four inputs. To facilitate comparative analysis, an additional model was constructed based on Willans line equations.

Figure 4.6 provides insight into the construction process of these models, shedding light on the various steps involved. Remarkably, the measured real power points, symbolized by blue dots, exhibit a striking alignment with both the green points of the training data and the red points represent the testing data. This alignment serves as a demonstration of the performance of an initial model, which consists of nine inputs. Supporting this observation, Table 4.1 presents the evaluation parameter values for both the training and testing data, with the training data showing a lower RMSE value, aligning with the anticipated outcomes. RMSE has long been recognized as one of the most reliable and widely used error metrics for evaluating the predictive performance of machine learning models in energy forecasting, making it an appropriate choice for this study [219].

Table 4.1: Root Mean Square Error (RMSE) factor for output of the model for testing and training data.

DATA	**RMSE**	**Data set size**
Training	0.36	37817
Testing	**0.47**	56725

The following section provides a comparative analysis of the results obtained from the three models, which are presented in Figure 4.7. This graph illustrates the predicted power in relation to the actual power. An ideal model would exhibit a linear relationship with a slope of one and a width of zero, signifying a perfect correspondence between the predicted and actual power values.

Considering Figure 4.7, it becomes evident that both the four-input and nine-input models surpass the performance of the Willans line model. Furthermore, a detailed examination comparing the four-input and nine-input models reveals that the former displays a narrower dispersion, indicative of its superior performance and precision. As illustrated in Figure 4.7, the line produced by the Willans Line model shows a marginal variance from the lines generated by the Willans line model. This discrepancy arises from the fact that approximately 80% of the input data exhibits a power output surpassing 22 MW, exerting a substantial impact on the accuracy of the model derived from the Willans Line. Hence, the resultant line is represented as a dashed line in regions where the power output is less than 22 MW.

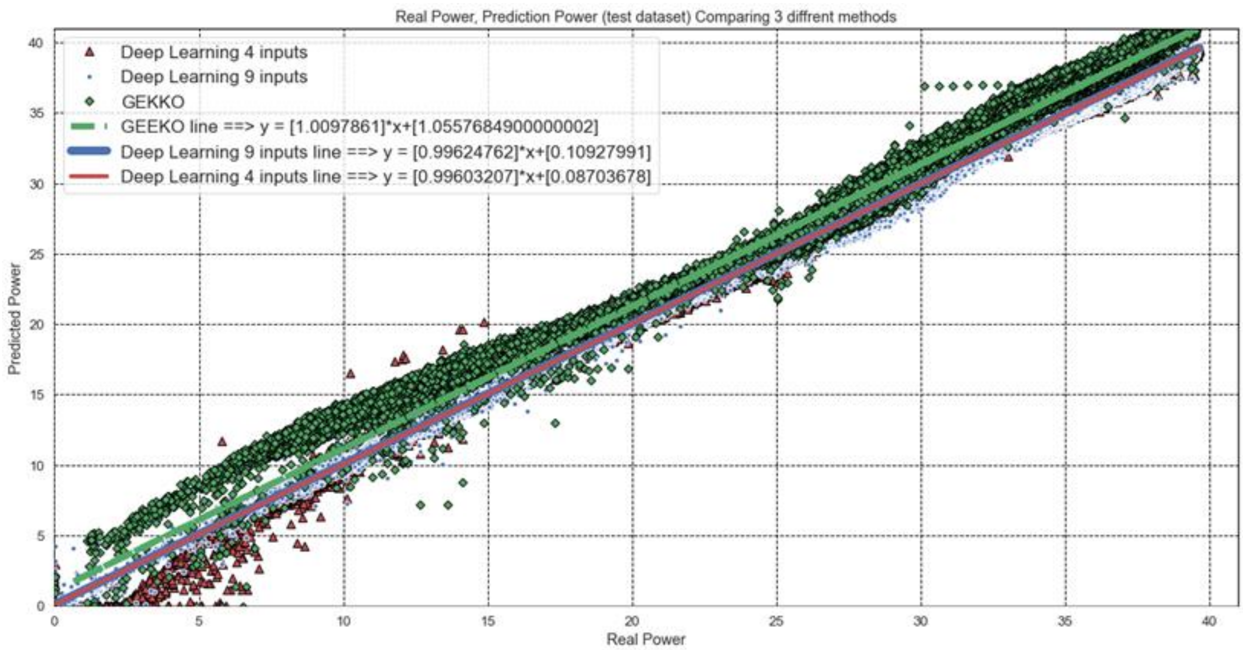


Figure 4.7: Comparison of actual power and predicted power for three models consisting of: Willans line model, 9-input prediction model and four-input prediction model.

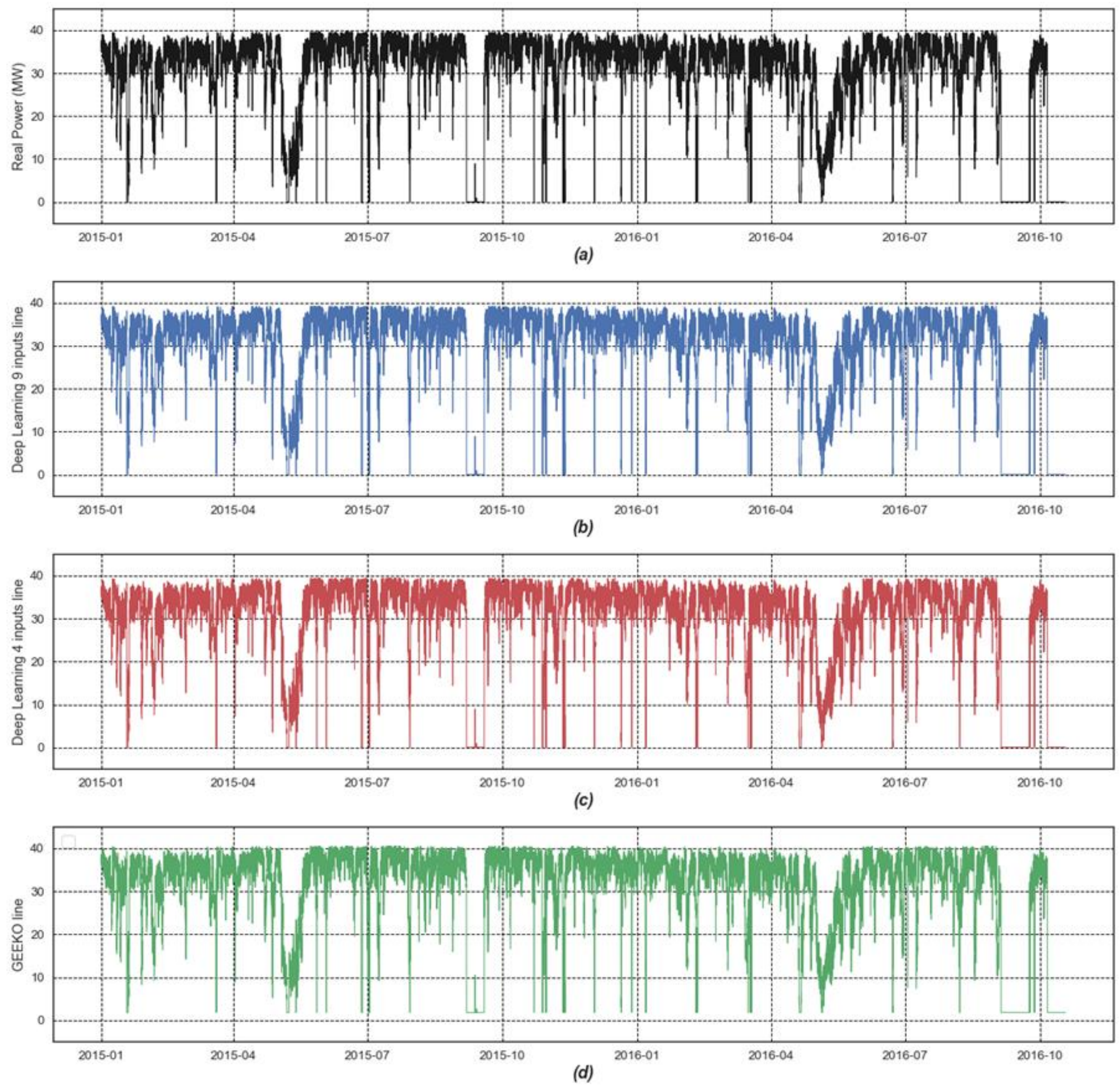


Figure 4.8: Comparison of (a) actual power and predicted power for three models consisting of: (b) 9-input prediction model, (c) four-input prediction model, and (d) Willans line model according to data collection time.

Further insights into the predicted power levels, taking into consideration the dates of data collection, are provided in Figure 4.8. Notably, the predicted powers depicted in the three charts demonstrate a high degree of consistency, with minimal disparities observed among them. To augment these findings, Figure 4.9 portrays the individual error magnitudes for each model, serving as additional evidence to reinforce the earlier explanations. This diagram further substantiates the notion that the four-input model outperforms both the nine-input model and the Willans line model, pointing to improved performance under the study conditions.

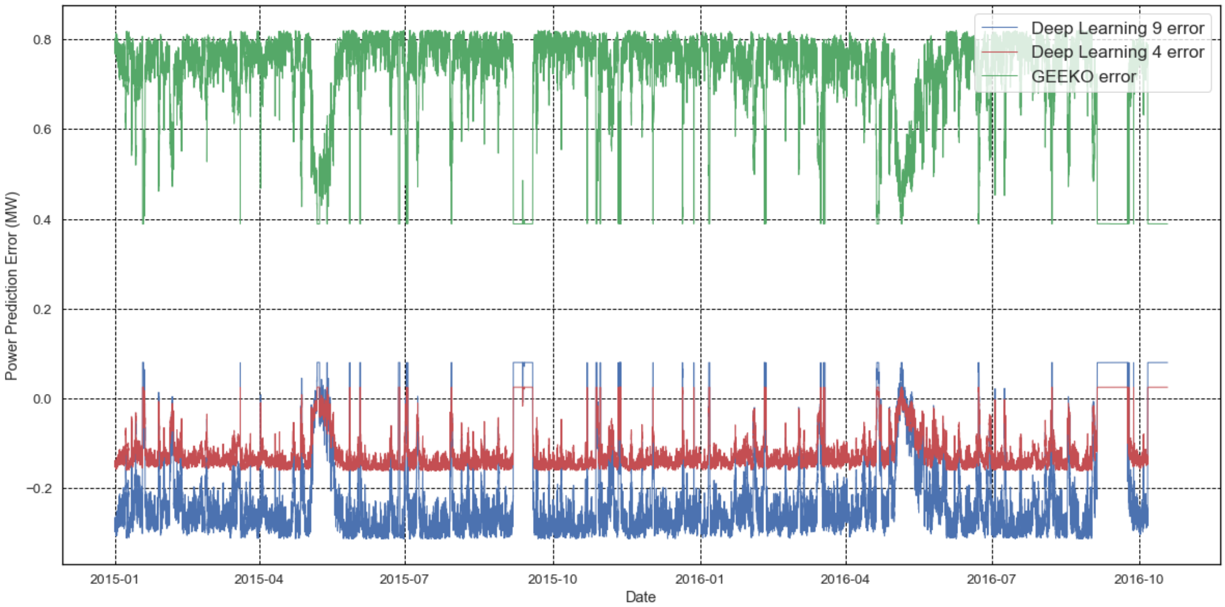


Figure 4.9: Individual error magnitudes for each of the three models.

The results obtained from the comparison of LSTM and Willans line models, evaluated using the RMSE metric, indicate that the LSTM approach achieved lower RMSE values compared to the Willans Line model in this study. Initially employing 9 inputs, the LSTM method achieved an RMSE of 0.47; however, by leveraging correlation analysis and refining the input set to 4 inputs, an even lower RMSE of 0.39 was attained. This study suggests the usefulness of using ML algorithms and AI in accurately predicting desired system performance in various industries.

Moreover, the study highlights the importance of choosing the appropriate ML technique based on the type of data examined. In cases where there is time dependence between data lines, techniques such as LSTM that consider the time dependence of input data should be utilized for accurate predictions. The study also highlights the significance of the quality of input data used in ML models, as the time interval of input data can significantly impact the performance of the system.

5. Enhancing Industrial Energy Efficiency with Predictive Analytics and Fuzzy Logic

5.1 Introduction

Building on the predictive analytics methodologies introduced in Chapter 4, which focused on the application of advanced ML techniques for forecasting steam turbine power generation, this chapter explores the integration of predictive models with energy management strategies in industrial settings. The emphasis shifts from power generation prediction to the efficient management of energy consumption, particularly in systems dependent on renewable energy sources such as solar power.

The urgency of addressing GHG emissions and mitigating global warming has led to a global shift toward renewable energy sources, including solar and wind power. New Zealand, with its ambitious goal of achieving net zero CO_2 emissions by 2050 [220], [221], [222], exemplifies this commitment. However, these new renewable energy systems face inherent challenges, such as dependency on weather conditions, and typically with an operational effectiveness much less than 100% of their capacity [223]. These challenges necessitate not only accurate prediction of energy production, as detailed in Chapter 4, but also effective energy management strategies to maximize the utility of renewable resources.

This chapter adopts a two-fold approach to address these issues, combining the predictive power of ANNs with the decision-making capabilities of FL. The first phase employs MLP to analyse open-source weather data, forecasting solar irradiance and energy availability. This phase builds on the principles of data-driven modelling introduced in Chapter 4, underscoring the critical role of high-quality input data and feature selection in achieving accurate predictions. By leveraging insights from correlation and sensitivity analyses discussed in Chapter 4, this chapter further refines the predictive model for optimal performance.

In the second phase, a FL system is implemented to manage electricity consumption within the meat processing industry, which serves as the case study for this research. FL, known for its robustness in handling uncertainty, offers a practical solution for adapting energy consumption to the variable availability of renewable resources. Unlike traditional management systems, FL excels in scenarios where precise mathematical models are impractical, enabling informed decision-making under ambiguous conditions [224]. This dual strategy not only improves prediction accuracy but also addresses the variability and uncertainty inherent in renewable energy systems.

The integration of ANN and FL provides a comprehensive framework for enhancing energy efficiency in industrial applications. While ANN models complex, non-linear relationships in weather data to forecast renewable energy production[225], FL establishes adaptive control mechanisms to optimize energy consumption based on real-time availability and demand. This approach bridges the gap between prediction and management, offering a unified solution for industrial energy systems.

Focusing on the meat processing industry, this research demonstrates the practical application of these techniques to manage electricity consumption. The methodology involves predicting solar energy radiation at the designated site using ANN, followed by the formulation of adaptive rules within the FL framework to address constraints and optimize resource allocation. These techniques aim to reduce reliance on traditional energy grids while enhancing the sustainability and efficiency of industrial operations.

By integrating predictive analytics with strategic management, this chapter extends the methodologies introduced in Chapter 4 into a broader context of renewable energy management. The combined use of ANN and FL highlights the transformative potential of ML and AI-driven techniques in addressing the dual challenges of variability and uncertainty in renewable energy systems. In doing so, this research contributes to the development of sustainable industrial energy practices that align with global efforts to mitigate climate change and achieve net zero emissions.

5.2 Description of an Industrial Process

This section explores the industry selected as the case study for this research, focusing on its electricity and energy consumption patterns. It also reviews the input data for the system and the model employed for managing energy consumption.

The case study focuses on the economic integration of renewable solar energy into a meat processing facility, which operates as a centralised factory microgrid. The electricity produced by this system is crucial for powering a variety of equipment including machines, chillers, freezers, and boilers. This setup features a significant solar installation with a capacity of 64 megawatts, supported by an energy storage system with a capacity of 16 gigawatt-hours. The configuration and operational framework of the case study are comprehensively detailed in Figure 5.1, which also shows the potential for community integration in the microgrid. This section aims to explore the practical aspects of employing solar energy in industrial applications.

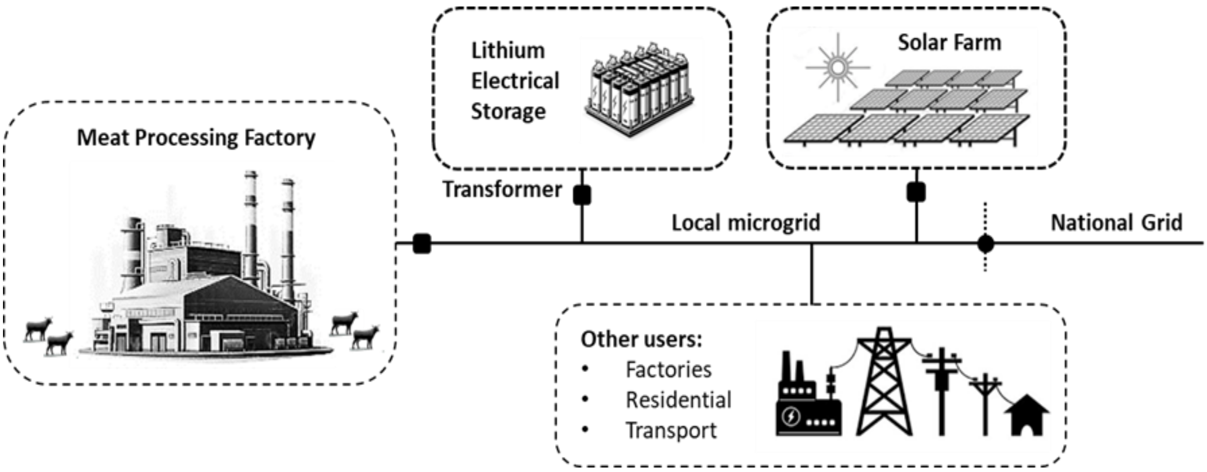


Figure 5.1: Integrating a low-carbon factory microgrid with local and national energy networks.

5.3 Description of Data and Models

In this study, two distinct data series have been employed to develop a comprehensive model for managing the electricity consumption within the system. The first dataset comprises weather-related information sourced from the New Zealand open-source National Institute of Water and Atmospheric Research database (NIWA) [226], which will inform predictions of the solar farm's production capacity.

The second dataset pertains to the power usage across various sectors of the meat processing facility, which, when combined with the forecasted data, is utilized to construct a FL model aimed at optimizing the utilization of renewable energy resources. While the two datasets differ in their time scales, necessary adjustments have been made to ensure cohesion within the data model. As part of the preprocessing, the time scales of the two datasets were aligned to ensure consistency in the analysis. The collated data encompasses a full-year cycle for 2022, segmented into 30-minute intervals.

Numerous options are available for selecting the input data for the initial model aimed at predicting solar global irradiance. These data undergo two distinct filtering processes before the selection of the most relevant variables for the model's input. Initially, based on prior research, variables such as Temperature, Relative Humidity, Wind Speed, and Azimuth are identified as influencing factors on solar radiation. In the subsequent stage, correlation analysis is employed as a sensitivity analysis tool [227], [228]. Ultimately, variables that exhibit an absolute correlation value greater than 0.4 with the ratio of solar global irradiance are selected as inputs for the MLP model. As illustrated in Figure 5.2, Temperature, Wind Speed, and Relative Humidity possess these characteristics and are thus chosen as model inputs.

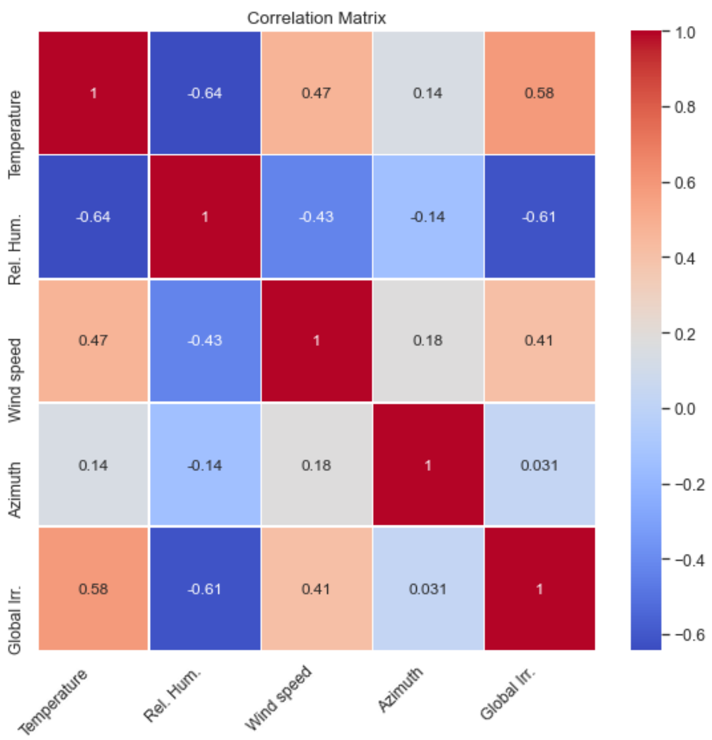


Figure 5.2: Correlation matrix of input data for the prediction model.

5.3.1 MLP Model and Solar Radiation Prediction

Once the inputs for the model are established, the next step involves selecting the appropriate method and architecture for predicting solar radiation. ANN is capable of modeling both linear and nonlinear relationships within systems [228]. Among various ANN approaches, MLP utilizing feed-forward back-propagation has been employed extensively in previous research to estimate solar radiation [229], [230], [231].

A challenge in this stage involves specifying the optimal number of hidden layers and the number of neurons within each layer, collectively referred to as hyperparameters. To address this issue, various algorithms have been explored. Notably, *Random Search* has demonstrated superior efficacy in optimizing these parameters when compared to Manual Search and Grid Search [232], owing to its ability to explore a broader range of potential solutions more efficiently. This methodology has been adopted in this research to fine-tune the MLP model.

The outcomes of this modelling approach, which significantly enhance the predictive accuracy of solar radiation estimations, are illustrated in Figure 5.3. In the context of neural networks, "dense" refers to fully connected layers where each neuron is connected to every neuron in the previous layer, enabling complex feature extraction and pattern recognition.

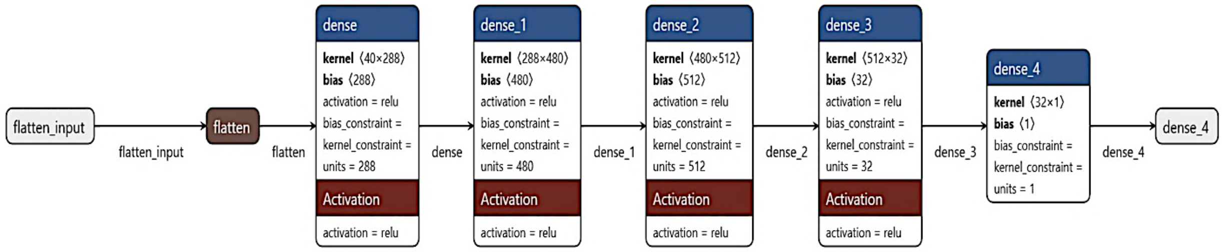


Figure 5.3: MLP architecture for prediction of solar radiation.

The optimal configuration identified comprises four hidden layers with the following neuron allocations: 288 neurons in the first layer, 480 in the second, 512 in the third, and 32 in the fourth. This structure emerged as the most suitable for handling the complexities of the dataset used. The efficacy of this model configuration is substantiated by the Mean Squared Error (MSE), recorded at 0.00439, which indicates a high level of precision in solar radiation prediction.

It should be noted that the structure and prediction error of the MLP model depend on the availability of input variables. In more realistic scenarios, where one or more inputs may be missing, the architecture would need to be reconfigured, and different levels of error would be expected. In the present study, NIWA's environmental data provided consistent availability of the selected variables, which enabled the development of the reported high-accuracy model. Nevertheless, for completeness and transparency, additional comparative results can be incorporated, including reduced-input scenarios and baseline methods (e.g., linear regression or ARIMA), to demonstrate model behaviour under alternative conditions.

5.3.2 Rule-based Fuzzy Logic and Electricity Consumption Management

In the FL system described membership functions for solar generation, electrical demand, and storage difference are established based on the maximum observed values from a dataset. Specifically, the maximum solar generation observed is 66,107 kW, and the maximum electrical demand is 19,198 kW. For each of these parameters, an Antecedent variable is created with a range extending from zero to the maximum value, or in the case of storage difference, from a negative to a positive range encompassing possible fluctuations. As is shown in Figure 5.4, solar generation and electrical demand variables are automatically categorized into three fuzzy sets ('Low', ‘Medium’, and 'High') using triangular membership functions, which are simple and effective for partitioning the input space. In contrast, the storage difference employs trapezoidal membership functions to define 'negative' and 'positive' states, designed to capture the extremes of surplus or deficit in storage relative to production and usage, thus facilitating more precise control actions in response to varying storage conditions.

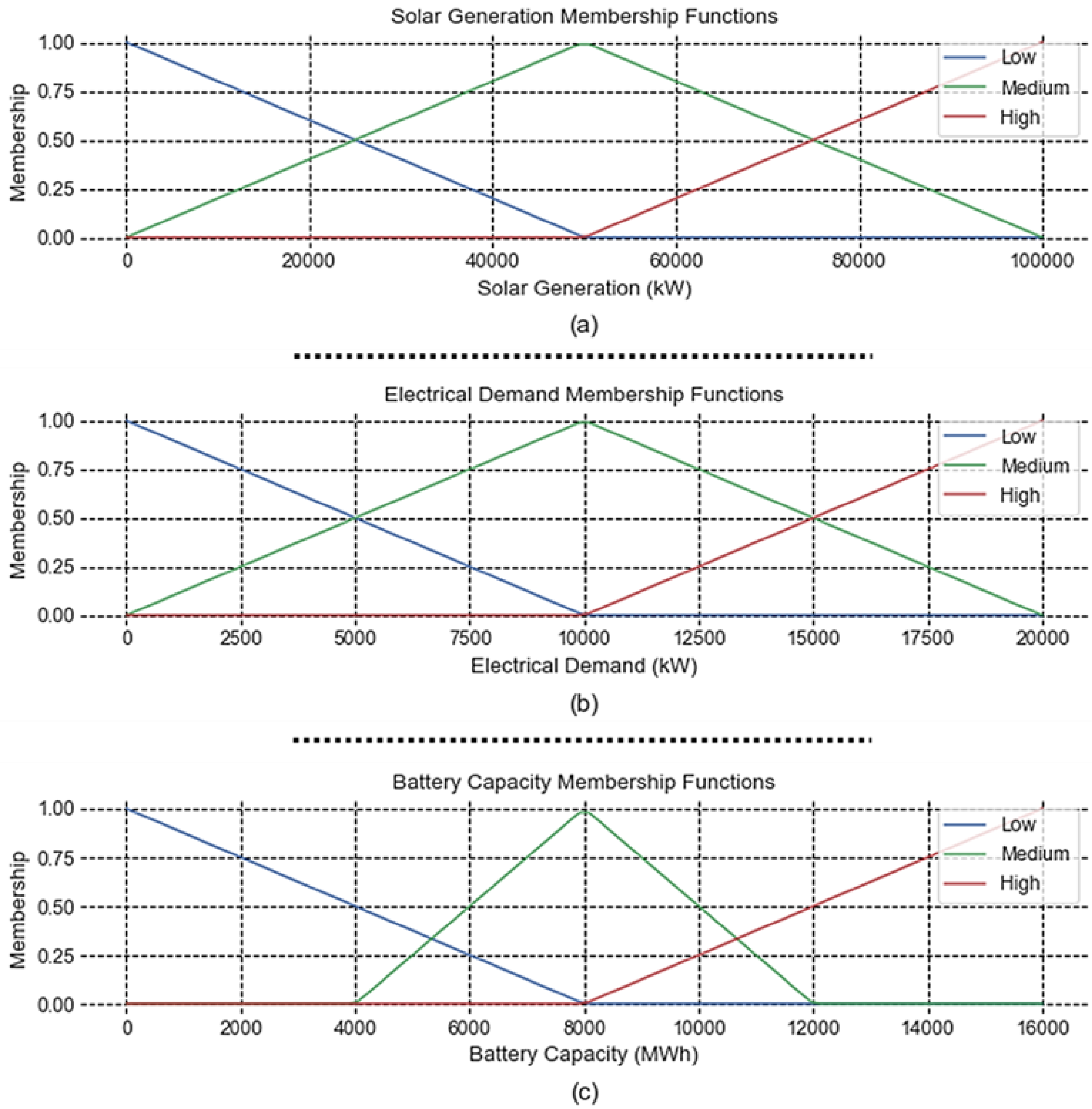


Figure 5.4: Membership functions for solar generation, electrical demand, and battery capacity.

After definition of membership functions, fuzzy rules are first established based on predicted solar radiation values, which are then used to calculate the output electricity of the solar farm. The electricity generated, alongside electricity demand from the grid and stored electricity, serves as inputs to the fuzzy system.

These inputs determine the selection of power sources and the potential electrical storage of solar energy in batteries. Initially identifying 27 decision states based on three potential input values—high, medium, or low—these are then consolidated into 12 distinct scenarios through the merging of similar states, as detailed in Table 5.1. This system prioritizes solar electricity utilization, followed by storage and grid usage, with battery charging occurring predominantly when solar output is high, and electricity demand is low. The outcomes of this approach will be further explored in the following section.

To better understand Table 5.1 and the system's decisions illustrated in the Figure 5.5, a visual representation is utilized. In Figure 5.5, each of the three rows in the chart initially indicates the level of solar energy production, followed by detailed information on the rule number, the chosen energy source for that situation, and the final decision on whether to charge the battery. Vertically, each of the three columns represents the electrical load level. Additionally, three different colours are used to depict the amount of electricity stored in the battery.

Table 5.1: Fuzzy electrical management systems rules.

	Inputs: Status of solar farm, electrical load, and electrical storage			**Output: Decision about power source & battery state**	
Rule	Solar Generation	Electrical Demand	Battery Capacity	Power Source	Charging Battery
1	Low	Low/Medium	Low	Grid	No
2	Low	Low/Medium	Medium/High	Battery	No
3	Low	High	Low/Medium	Grid	No
4	Low	High	High	Battery	No
5	Medium	Low	Low	Solar	Yes
6	Medium	Low	Medium	Battery	Yes
7	Medium	Low	High	Battery	No
8	Medium	Medium/High	Low/Medium	Solar	No
9	Medium	Medium/High	High	Battery	No
10	High	Low	Low/Medium /High	Solar	Yes
11	High	Low	High	Battery	Yes
12	High	Medium/High	Low/Medium /High	Solar	No

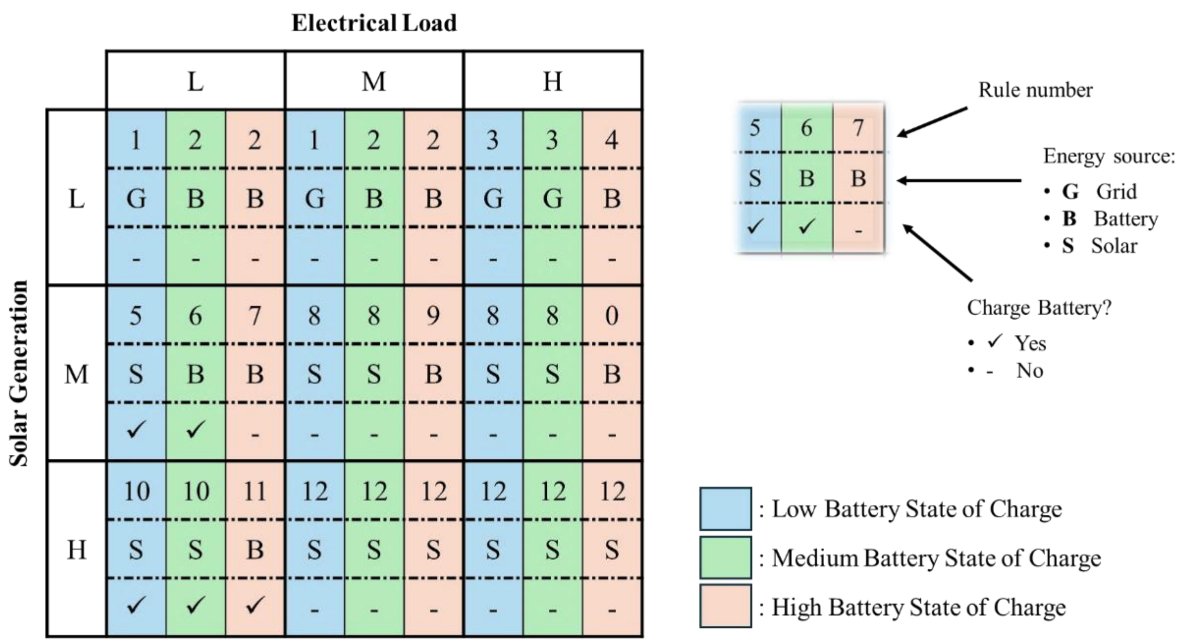


Figure 5.5: The energy source, and battery charging decisions, derived from Table 5.1, focusing on solar generation, overall load, and battery state of charge.

5.4 Evaluating adaptive electricity strategies

Upon completing the development of both the prediction model and the electricity consumption management system, the outcomes are shown in Figure 5.6. Figure 5.6(a) displays the decision-making process over the course of a year, illustrating the various strategies adopted based on solar radiation levels: utilizing the electricity grid, directly employing solar energy from the solar farm, using the solar farm to charge the battery, or consuming the stored energy from the battery directly. This variability highlights the adaptive use of different energy sources in response to solar availability.

Figure 5.6(b) quantifies the extent of reliance on each source, showing that the national grid was independently used to supply electricity 38.95% of the time, the solar farm was used to charge batteries 38.48% of the time and provided direct electricity for 13.48% of the time, with the remaining 9% of energy consumption coming directly from battery storage.

To further evaluate the effectiveness of the system, Figure 5.7 examines a one-week period, focusing on how solar power generation aligns with its use for battery charging and direct supply to industry. The graph illustrates that the management system coordinated solar production with demand in this case study, showing potential for improved integration with optimal use of renewable resources. This synchronization not only reflects the system's capacity to manage fluctuating energy inputs but also its ability to maximize the local consumption of locally produced energy and sustainably meet industrial electricity requirements.

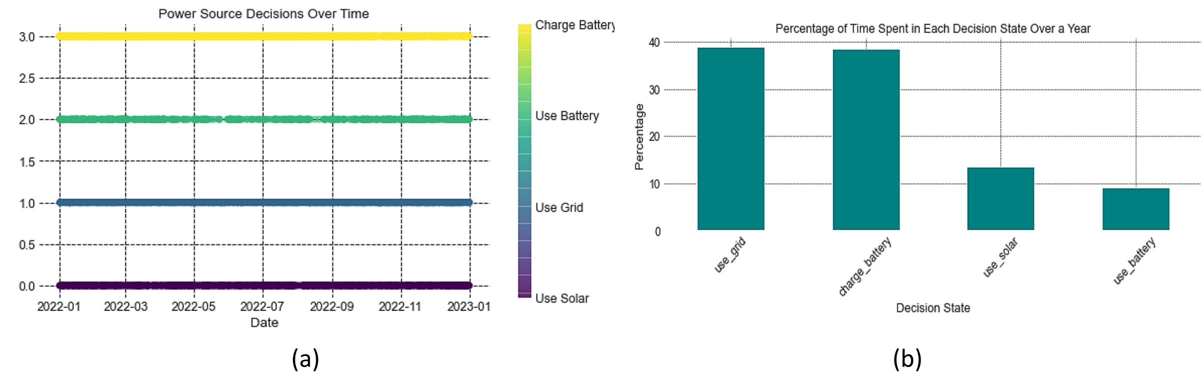


Figure 5.6: Evaluating adaptive electricity strategies: (a) Decision dynamics across a yearly cycle (b) Comparative usage analysis of energy sources.

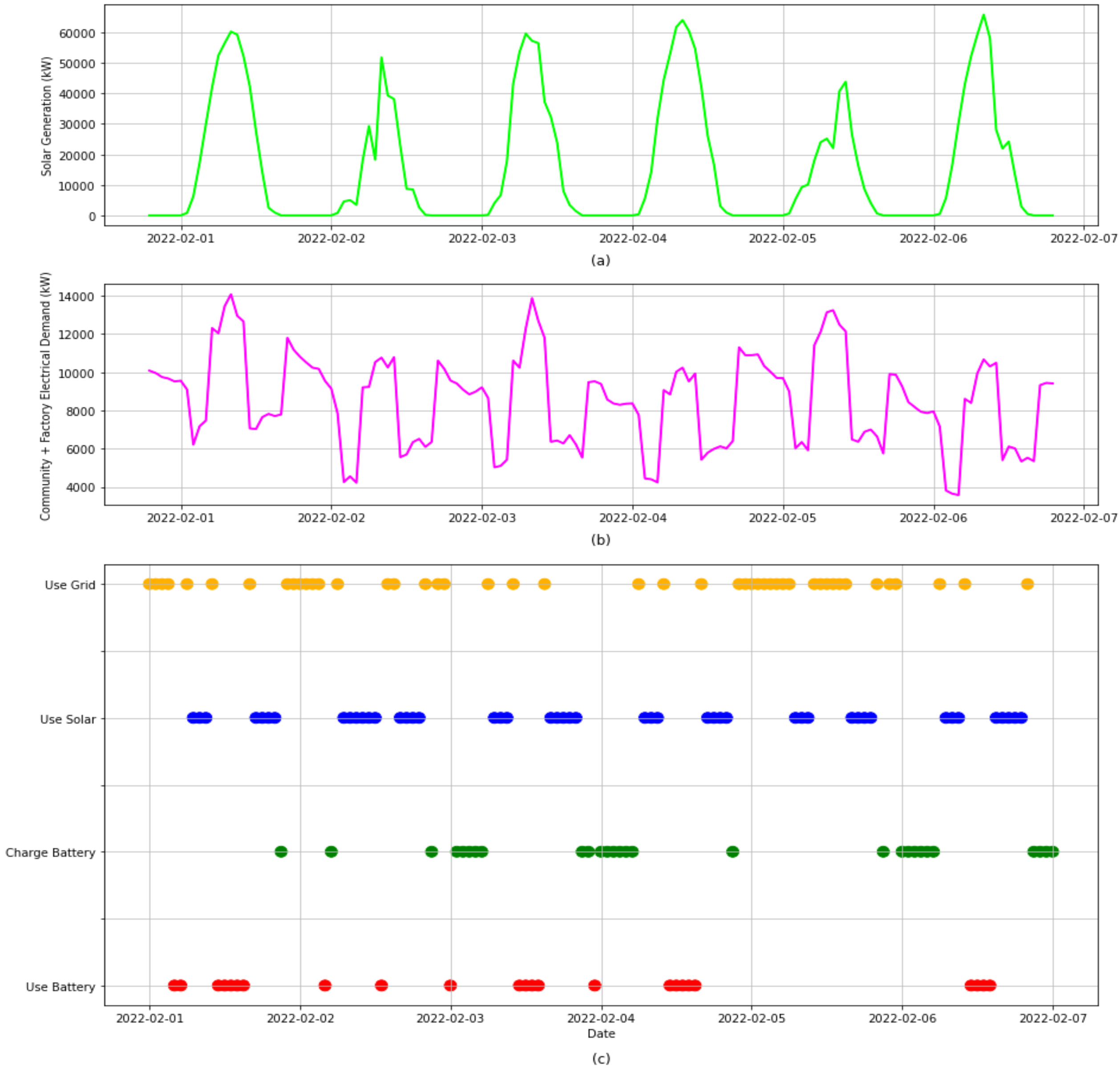


Figure 5.7: Decisions considering solar generation and electrical demand: (a) Weekly Solar Energy Production (kW), (b) Weekly Electrical Demand (kW), (c) Operational decisions on electricity source and storage charging by the fuzzy system.

This study has explored the integration of ANNs and FL to address the inherent variability of renewable electricity sources, specifically in the meat industry. By utilizing ANNs to predict the impact of weather variations on solar irradiance and incorporating FL to refine energy consumption decisions in real-time, the research aimed to reduce reliance on conventional energy grids. The strategies developed focused on optimizing the use of solar-generated electricity, dynamically adjusting between direct consumption, storage, and grid supplementation according to solar availability.

The case study demonstrates that combining ANNs for solar irradiance prediction and FL for adaptive energy management offers a more efficient approach to handling renewable electricity within industrial operations. This method successfully improves energy utilization by aligning electricity consumption with solar generation, reducing the facility's dependence on non-renewable grid energy. Although this approach shows promise in efficiently managing electricity sources, the results indicate potential areas for further improvement. It is important to emphasize that this chapter presents fuzzy logic as a proof-of-concept control strategy in an industrial case study, rather than a claim of definitive superiority over other methods [233]. In this context, Model Predictive Control (MPC) is frequently regarded in the wider literature as a benchmark for microgrid energy management due to its optimization capability and constraint handling. However, MPC typically requires higher computational resources and accurate system models [233],[234], which may limit its practical deployment in some industrial environments. The systems implemented here therefore represent an initial step towards enhancing sustainable practices within industrial settings, while highlighting the need for future comparative studies between fuzzy logic and MPC approaches.

Practical implications of this research suggest that the developed model can be applied to other industrial sectors with similar energy demands, provided that relevant data, such as weather and electricity consumption patterns, are available. The ability to switch dynamically between energy sources based on real-time predictions makes this model adaptable to diverse energy management scenarios, particularly in industries striving for renewable electricity integration. For future research, several areas require exploration to enhance the model's performance. First, comparative studies with alternative energy management techniques, such as Model Predictive Control (MPC), would further substantiate the model's effectiveness across different industrial contexts. In addition, exploring the scalability of the system to larger and more complex industrial facilities could provide insights into its broader applicability. Performing sensitivity analysis on input parameters and incorporating additional variables, such as electricity price, could also lead to improvements in both prediction accuracy and management efficiency. Finally, as AI technologies continue to evolve, future work could integrate advanced ML methods to further optimize the model's adaptability and performance in managing renewable electricity sources, while also conducting direct comparisons with MPC and rule-based controllers under identical conditions to quantify the benefits and limitations of fuzzy logic control.

6. Edge-Driven Electricity Trading in Fractal-Structured Microgrids: A Machine Learning Approach

6.1 Introduction

Efficient energy management in modern power systems is crucial for maintaining grid stability, enhancing renewable energy utilization, and minimizing environmental impact. However, the increasing integration of renewable energy sources, particularly solar power, introduces significant challenges due to their inherent variability and intermittency. These fluctuations result in concept drift, a phenomenon in ML where the statistical properties of energy generation and consumption patterns evolve unpredictably over time [235]. Traditional static ML models struggle to maintain accuracy in such dynamic environments because they are trained with historical data and are not equipped to adapt to ongoing changes. To address these challenges, this research explores the application of SML algorithms to optimize energy management in microgrids. Unlike conventional models, SML continuously updates its predictive models as new data becomes available, ensuring that forecasts remain accurate and relevant in rapidly changing scenarios [236], [237].

Previous research has applied SML to microgrid forecasting and control tasks; however, these studies typically focus on centralized configurations or static learning approaches [238]. The novelty of this research lies in its specific exploration of incremental real-time learning algorithms particularly Hoeffding Trees in decentralized residential microgrids, explicitly addressing dynamic consumption and generation patterns at the household scale.

The incremental learning capability of SML is particularly advantageous in energy systems characterized by fluctuating supply and demand. By updating models incrementally rather than retraining them from scratch, streaming algorithms adapt to changes efficiently, maintaining high predictive accuracy even as other external factors influence consumption and generation patterns. This adaptability is essential for ensuring appropriate real-time decision-making, enabling EMS to respond promptly to sudden shifts in demand and supply [239]. Additionally, streaming models are designed to be computationally lightweight, requiring less memory and processing power compared to traditional batch learning methods. This efficiency makes them ideal for deployment in decentralized environments such as microgrids, where computing resources may be limited but real-time decision-making is critical [240].

By integrating streaming analytics into existing energy management frameworks, this research introduces a novel approach to real-time optimization, enabling intelligent control of power generation, storage, and consumption. This integration allows for dynamic adjustments based on immediate energy availability and demand, enhancing the efficiency of battery storage management and reducing dependency on central grid power. Furthermore, the capability to process continuous data streams supports grid-edge trading, which involves dynamically redistributing or storing surplus solar power for later use. This not only maximizes renewable energy utilization but also minimizes the need for grid imports, thereby promoting energy independence and reduction of grid losses [241].

As suggested earlier, the dynamic nature of SML makes it particularly suitable for environments where energy consumption and generation are highly variable, such as residential microgrids dependent on solar power. These systems must continuously adapt to changes in weather conditions, user behaviour, and other external influences impacting energy availability and demand. By enabling real-time forecasting and adaptive control mechanisms, SML enhances the responsiveness of localized energy systems. In residential settings, where consumption typically comprises both base loads (essential demand) and discretionary loads (flexible or deferable demand such as heating, cooling, or electric vehicle charging), this adaptability provides significant opportunities for load shifting and DSM. Such capabilities significantly improve load balancing and grid reliability, ensuring a stable and efficient power supply even under fluctuating conditions. In this context, Hoeffding Trees [13] and ensemble methods are particularly effective, as they support fast, incremental learning and robust decision-making. These methods have demonstrated high accuracy in predicting short-term fluctuations in solar power generation and electricity usage, thereby minimizing reliance on external power sources for residential microgrids [242].

In addition to enhancing forecasting accuracy, SML also facilitates the automation of energy management processes. By continuously learning from new data, the models can automatically adjust control strategies to optimize energy flows in real time. This includes making intelligent decisions about when to store surplus energy in batteries, when to dis-charge stored energy to meet demand, and when to engage in grid-edge trading to maximize economic benefits. The ability to autonomously manage power generation, consumption, and storage supports the development of decentralized, self-sustaining energy networks. These networks operate independently of traditional grid infrastructures, making them ideal for isolated or remote communities that rely on renewable energy sources. By reducing dependency on centralized power grids, such systems contribute to greater energy security and resilience, especially in regions prone to grid outages or supply disruptions [239], [243].

As energy systems become more complex and interconnected, the need for adaptive, real-time solutions becomes increasingly critical. This research provides a scalable and flexible framework that can be adapted to various energy management scenarios, demonstrating the transformative impact of ML on modern power systems.

In Section 6.2, the key challenges related to temporal energy imbalances within microgrids are identified, traditional static energy management methodologies are critiqued, and SML as a viable real-time adaptive solution is introduced. Section 6.3 then describes the proposed fractal architecture framework for intelligent energy management, detailing its hierarchical structure, the integration of SML for forecasting and adaptive control, and the implementation of three scenario-based battery sizing strategies: centralised, decentralised (distributed), and a hybrid approach. Following this, Section 6.4 provides a comprehensive comparative evaluation of energy management and forecasting strategies, highlighting the predictive accuracy of the SML model, the effectiveness of adaptive control mechanisms, the reduction in grid dependency, and improvements in battery lifecycle management. Section 6.5 summarises the findings, emphasising practical implications, contributions to-ward sustainable and resilient microgrid operations, and future research opportunities.

6.2 Adaptive Energy Management in Smart Microgrids

While the introduction highlighted the strategic value of SML for managing the dynamic nature of renewable-powered microgrids, this section explores the complexities of temporal energy imbalances, critiques the limitations of conventional energy management approaches, and proposes SML as a viable solution for real-time adaptive energy management in smart microgrids.

Challenges of Temporal Energy Imbalances

Microgrids, characterized by their decentralized nature and integration of renewable energy sources, face inherent fluctuations in both energy production and consumption. Renewable sources such as solar and wind are inherently variable, leading to periods of surplus and deficit within short time frames. Traditional Net-Zero Energy models, which focus on achieving an annual balance between energy production and consumption, fail to address these short-term discrepancies [244]. As a result, microgrids may experience frequent transitions between importing energy from the main grid and exporting excess energy back to it, leading to operational inefficiencies and increased dependency on the central grid.

The main challenge lies in the temporal mismatch between energy generation and consumption. For instance, peak solar generation occurs during the day, which may not align with peak consumption periods in residential areas. This misalignment necessitates the need for effective energy storage solutions and real-time management strategies to ensure a continuous and reliable energy supply. Moreover, frequent charging and discharging cycles of battery storage systems, prompted by these imbalances, can lead to accelerated wear and reduced lifespan of the storage infrastructure [245].

Limitations of Conventional Energy Management

Conventional EMS in microgrids predominantly utilize static models that rely heavily on historical data to forecast energy demand and generation. These models are built on predefined rules and statistical patterns observed over extended periods, assuming that future energy behaviors will closely mirror past trends. While this approach provides a foundational understanding of consumption and generation patterns, it is fundamentally limited in its ability to respond to real-time fluctuations and unforeseen changes in energy dynamics. This limitation becomes particularly pronounced in fractal-structured microgrids, where decentralized nodes operate autonomously, and local variations in energy flow can significantly impact overall system stability [246].

In fractal microgrids, energy generation and consumption are highly localized and vary across different hierarchical nodes. This decentralized structure enables localized decision-making and promotes energy exchange between nodes. However, it also introduces significant variability in energy flows due to differences in solar exposure, local load demands, and storage capacities at each node. Static prediction models, which are generally built on aggregated historical data, fail to capture these localized variations, leading to inaccurate forecasts and suboptimal energy management decisions [247].

Moreover, in a fractal architecture, energy flows between nodes are highly dynamic, with nodes exchanging surplus energy or drawing power based on real-time conditions. Static models lack the adaptability needed to account for these inter-node dependencies, resulting in inaccurate load balancing and inefficient resource allocation. For instance, if one node experiences excess solar generation while a neighbouring node faces a deficit, static models may fail to identify this imbalance in real time, leading to unnecessary grid imports or wasted renewable energy [248].

A critical challenge that further undermines the effectiveness of static models is concept drift, where the statistical properties of energy consumption and generation change over time [142]. In the context of fractal microgrids, concept drift can occur due to seasonal variations, changes in user behaviour, and the intermittent nature of renewable energy sources. These dynamic patterns introduce non-stationarity into energy flows, making it challenging for static models to maintain accurate predictions. In fractal microgrids, where local nodes experience unique patterns of generation and consumption, concept drift is even more pronounced. Static models, which rely on the assumption that patterns in the data remain constant over time, are ill-equipped to handle these changes, leading to degraded performance and reduced prediction accuracy over time [249].

Figure 6.1 summarizes the sequential limitations inherent to static EMS within fractal microgrids. Starting from suboptimal resource allocation, these deficiencies cascade into inaccurate load balancing and reliance on outdated historical data, ultimately resulting in significant prediction inaccuracies due to unaddressed local variations and concept drift. This cyclical illustration emphasizes the necessity of transitioning from static methodologies toward adaptive ML models for effective energy management.

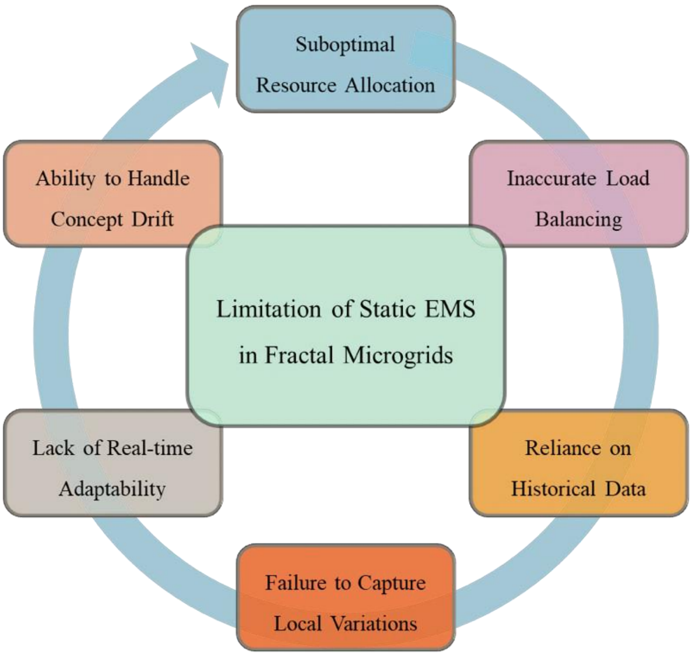


Figure 6.1: Limitations of Static Energy Management Systems in Fractal Microgrids.

Given these challenges, there is a growing consensus on the need for adaptive EMS capable of learning from evolving data streams and adjusting to re-al-time changes. In fractal microgrids, where energy flows are decentralized and highly variable, SML emerges as a promising solution. These algorithms are designed to handle data in motion, continuously updating their parameters to reflect the most recent patterns.

This capability is particularly beneficial for fractal architectures, where energy flows are asynchronous, and local nodes operate autonomously. SML offers several key advantages over static models such as updating models continuously as new data arrives and also maintaining accurate predictions even under dynamic conditions [250].

Streaming Machine Learning for Real-Time Adaptation

Given the challenges of temporal energy imbalances and the limitations of traditional EMS discussed earlier, SML presents a viable adaptive solution, offering the flexibility and scalability essential for managing dynamic conditions in de-centralised renewable energy systems. Unlike batch learning models that rely heavily on historical datasets, SML algorithms incrementally process incoming data, continuously updating their models. This incremental learning capability allows them to maintain high predictive accuracy in rapidly changing environments typical of microgrids powered by renewable energy sources such as solar and wind.

One of the core strengths of SML is its ability to handle concept drift changes in the statistical properties of data over time common in microgrids due to seasonal variations, changes in user behaviour, weather conditions, and integration of new renewable sources. For instance, residential microgrid energy consumption patterns may shift due to household routines, appliance upgrades, or unforeseen events, reducing the effectiveness of static prediction models. Algorithms such as Hoeffding Trees and Adaptive Random Forests, central to SML, are specifically designed to detect and respond to these dynamic shifts by updating predictions incrementally [251]. This ensures sustained accuracy even as energy generation and consumption patterns evolve [235].

By integrating SML into microgrids, operators can perform real-time decision-making to optimize adaptive energy management strategies. Accurate, real-time forecasting of energy production and consumption guides critical operational decisions, including timing for battery storage charging or discharging and engaging in grid-edge trading [252]. For example, microgrids utilizing solar power generation can predict short-term solar availability using SML, enabling proactive storage of surplus energy or intelligent interaction with the main grid. This approach optimizes both energy utilization and economic benefits, ensuring continuous power availability while minimizing reliance on external power sources.

Moreover, real-time adaptability via streaming algorithms directly contributes to optimizing battery storage efficiency. Traditional battery management approaches often suffer inefficiencies due to temporal mismatches in energy generation and consumption cycles. SML-driven forecasts allow microgrids to better synchronize battery operations with actual real-time demands, significantly reducing battery wear and extending storage system lifespans [253]. Additionally, the lightweight and computationally efficient nature of streaming models supports effective deployment even in resource-constrained microgrid environments.

Finally, integrating SML into microgrid systems enhances the capability for efficient grid-edge trading, facilitating dynamic interaction with the central grid. Microgrids can optimize trading decisions based on continuous real-time data streams, ensuring energy exchanges occur at the most beneficial times, thus improving economic returns and over-all energy efficiency [254]. Ultimately, SML represents a transformative approach, significantly enhancing the resilience, sustainability, and autonomy of microgrids, promoting smarter and more adaptable energy infrastructures.

6.3 Fractal Architecture for Intelligent Energy Management

This research introduces a fractal architecture framework that employs SML to achieve dynamic and intelligent control, marking a significant evolution from traditional fractal modelling techniques. By integrating self-similar structures, this framework enables scalable and decentralized management across various operational levels, from individual households to extensive community networks. Each node within this hierarchical system can autonomously perform local energy exchanges and engage in grid-edge trading, enhancing distributed and adaptive control through localized decision-making.

The integration of SML into the fractal architecture significantly enhances its capabilities by enabling real-time predictive analytics and adaptive control mechanisms across the energy management system. Local nodes use these capabilities for immediate responses to fluctuations in energy demand and supply, ensuring efficient operations at the micro-level. Concurrently, higher-level nodes aggregate insights from the grassroots to orchestrate comprehensive energy management strategies across the grid. This establishes a robust multi-level learning hierarchy that enhances decision-making efficacy and system resilience.

The concept of temporal energy balance is central to maintaining stability within the fractal architecture. By utilizing SML, the system dynamically adjusts power generation, storage, and consumption to maintain a continuous balance, optimizing energy flows and minimizing grid dependency. These adaptive control mechanisms are crucial for enhancing the grid's performance, enabling it to swiftly and effectively respond to changes, thus promoting sustainability and efficiency within the energy network.

Multi-level Node Interconnection and Hierarchical Tree Structure

Building on insights from Apperley et al. [255], this research employs a uniform node architecture interconnected through a local grid. Unlike traditional hierarchical architectures where each node often has distinctly different roles, this fractal framework uses identical structural nodes at every level—household, neighbourhood, and community. Each node independently functions as a microgrid, managing its own energy generation, consumption, and storage dynamically. The local grid connectivity ensures efficient and simplified interactions between nodes, facilitating seamless local energy exchanges and grid-edge trading.

In this design, each node within the microgrid network autonomously manages its energy state—balancing local generation with consumption and storage. When surplus energy generation occurs at a household node, the surplus is efficiently distributed to neighbouring nodes experiencing a deficit via a local grid. This localized redistribution of energy substantially enhances the microgrid's overall efficiency and reduces dependence on external central grid infrastructure, significantly improving local energy balance [256].

The local grid interconnection also simplifies communication between nodes, employing concise, standardized pricing signals. These signals facilitate effective local energy transactions and enable nodes to make rapid economic decisions regarding energy buying, selling, or storage actions based on real-time predictive data provided by SML algorithms. This streamlined communication approach aligns well with principles from smart communication protocols, which emphasize simplicity, minimal data overhead, and robust reliability [257].

The hierarchical fractal architecture of the network, featuring uniform nodes inter-connected via a local grid, is shown in Figure 6.2. The diagram emphasizes the uniformity and simplicity of each node, highlighting their capabilities for local energy generation, storage management, and consumption balancing. In the proposed framework, each node operates in one of three clearly defined states, as characterized by Apperley et al. [256]:

- **Deficit state**: Local consumption exceeds local generation and available storage, necessitating energy import from adjacent nodes or higher-level grids.
- **Balanced state**: The node can precisely match local generation and consumption, optimizing local energy use without surplus or deficit.
- **Surplus state**: Local energy generation exceeds consumption and storage capacity, prompting the node to export excess energy to neighboring nodes or higher hierarchical levels.

Through real-time predictive analytics enabled by SML, each node continuously forecasts its future state based on real-time data streams such as weather forecasts, energy usage trends, and storage status. These predictions inform adaptive control decisions, enabling nodes to proactively transition between states. This dynamic state management enhances local energy utilization efficiency, reduces unnecessary reliance on external grids, and significantly improves grid stability and operational resilience.

Thus, the hierarchical tree structure, interconnected by a local grid complemented by intelligent SML-driven forecasting, establishes a powerful, scalable solution for achieving optimal energy management across multiple levels of a fractal-structured microgrid.

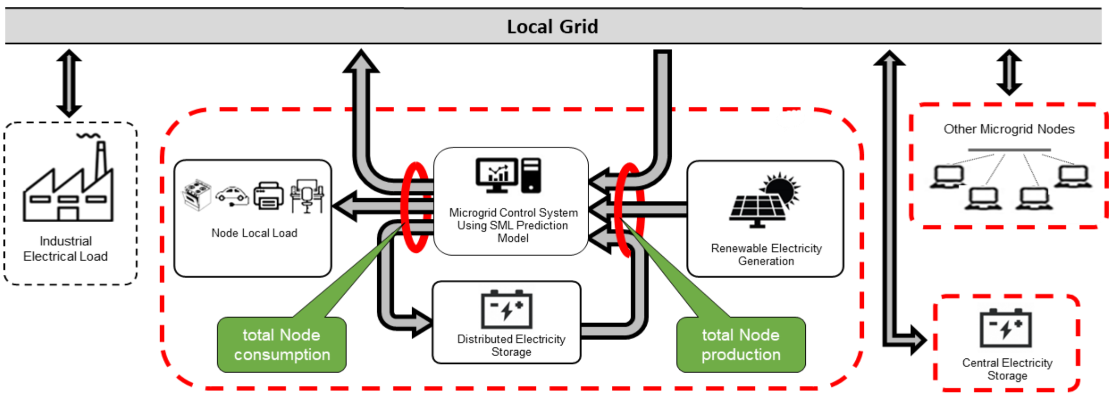


Figure 6.2: A microgrid with a hierarchical fractal architecture with uniform nodes interconnected via a local grid.

Integration of Streaming Machine Learning into Fractal Structures

The integration of SML into fractal microgrid architectures significantly advances autonomous energy management, enabling real-time adaptive control and optimized decision making at individual node levels. In this research, Hoeffding Trees, a prominent incremental learning algorithm, were employed due to their efficiency in processing continuous data streams and rapidly adapting to changes in consumption patterns [258]. Un-like traditional ML methods that require periodic retraining with historical data, Hoeffding Trees incrementally update their predictive models with each incoming data point. This incremental learning capability ensures sustained accuracy and responsiveness, crucial for handling the concept drift inherent in renewable energy generation and consumption dynamics [257].

The dataset used in this study consists of hourly energy consumption data recorded over a full calendar year from multiple residential households situated in an isolated community environment at a similar latitude to Great Barrier Island, New Zealand. Given that the selected community is not connected to a centralized grid, the data profiles were developed from carefully chosen households on the New Zealand mainland with similar characteristics, including occupancy patterns, seasonal variations, appliance utilization. The dataset captured significant variability among households, with additional stochastic elements included to represent localized events such as community gatherings (hui), thereby ensuring the models closely reflect realistic and diverse operational scenarios.

Preprocessing of this high-resolution dataset included several critical steps to enhance predictive accuracy. Lag variables were incorporated to capture immediate temporal dependencies, providing the model with contextual information regarding recent energy consumption. Data normalization was applied to ensure consistency and comparability across diverse household profiles, which was essential for the SML model to accurately respond to the high variability present in the real-time data.

A predictive model utilizing Hoeffding Trees was developed and validated, leveraging ensemble learning through Bagging Regressors [259] to enhance robustness and accuracy. The Hoeffding Tree Regressor used a default MSE loss function, with a grace period of 20 samples and a maximum tree depth of 10, while the ensemble comprised 20 base models. The model continuously updated its predictions as new hourly data became available, implicitly managing concept drift through ongoing incremental learning. Predictive effectiveness was quantitatively assessed using MAE, demonstrating reliable forecasting performance across individual nodes. Sensitivity analysis of the grace period and tree depth parameters confirmed stable model performance across variations, supporting the chosen defaults.

The following pseudocode (Algorithm 1) outlines the real-time implementation logic of the proposed model. After initialising a Hoeffding Tree ensemble with parameters such as maximum depth and grace period, the model receives hourly consumption data in streaming format. Each new data point is pre-processed to generate lag features and normalised for consistency across household profiles. Forecasted energy demand is computed and immediately used to inform energy balancing decisions at the household level. If a surplus is predicted, available excess energy is stored in local batteries within defined operational constraints; if a deficit is anticipated, discharge is triggered provided the battery's state-of-charge exceeds a preset threshold. Grid interaction is dynamically adjusted to either export or import energy depending on the real-time net demand. This pro-cess is repeated continuously, allowing the system to adapt incrementally to evolving consumption and generation patterns. The summarised pseudocode reflects this workflow.

Algorithm 1: Adaptive Streaming Energy Management using Hoeffding Trees

Input: Streaming data of energy consumption $\{x_0, x_1, \ldots, x_n\}$, system parameters θ, battery state B(t) at time t

Output: Forecasted energy demand $\hat{y}(t)$, battery management actions A(t)

1: Initialise Hoeffding Tree Regressor with predefined depth and grace period
2: For each incoming data point x_t in the stream, do
3: Extract feature vector f_t = preprocess(x_t, lag features, normalization)
4: Predict energy demand $\hat{y}_t$ using the current Hoeffding Tree model
5: Update the model incrementally with (f_t, x_t) if feedback is available
6: Compute energy surplus/deficit: $\Delta_t = \text{generation}_t - \hat{y}_t$
7: If $\Delta_t > 0$, then
8: Charge battery B(t) within operational constraints
9: Else if $\Delta_t < 0$, then
10: Discharge battery B(t) if state-of-charge > threshold
11: Else, maintain the current battery state
12: Adjust grid interaction accordingly (export/import power)
13: End for
14: Output updated $\hat{y}(t)$ series and battery control actions A(t)

While more sophisticated ensemble-based streaming models such as Adaptive Random Forests (ARF) may offer improvements in predictive accuracy under certain conditions, they also introduce significantly higher computational overhead, memory consumption, and complexity compared to Hoeffding Trees, making them less suited for lightweight, real-time applications in decentralised microgrid settings. Similarly, adaptive variations such as Adaptive Hoeffding Trees and explicit drift detection methods (e.g., DDM, ADWIN) were not explored in this study, as the primary focus was on evaluating the effectiveness of lightweight incremental learners in practical energy management scenarios. Nevertheless, to support our selection of Hoeffding Trees, a comparative evaluation was conducted with the Adaptive Hoeffding Tree model on the same dataset, demonstrating lower MAE and confirming the effectiveness of our chosen approach. Future studies could further enhance real-time adaptability and robustness by exploring these advanced techniques [260].

Accurate forecasting of energy demand directly enables efficient grid-edge trading by clearly identifying periods of surplus or deficit. Nodes within the fractal microgrid can proactively manage local energy exchanges, enhancing energy autonomy, economic efficiency, and resilience against centralized grid disruptions. This strategic, predictive-based energy management framework significantly reduces dependency on external infrastructure, improving both local and system-wide stability.

Moreover, contemporary studies underline the critical importance of reliable forecasting methods for effective microgrid management. Hosseini et al. [261] introduced the similar pattern algorithm for monthly electricity consumption prediction, emphasizing the value of historical pattern recognition in forecasting accuracy. Similarly, research on fractal smart grids highlights how adaptive ML algorithms effectively respond to real-time fluctuations in renewable energy sources, thereby improving overall grid performance and reliability [261]. These insights further validate the approach adopted in this study, rein-forcing the suitability of Hoeffding Trees for addressing the inherent variability and concept drift encountered in decentralized renewable energy systems.

Recent literature also supports the combined use of predictive analytics with emerging digital technologies, such as digital twins, to optimize energy consumption and storage strategies dynamically. Digital twins, as high-fidelity virtual models reflecting physical energy systems, leverage real-time predictive analytics to enhance operational efficiency and responsiveness, paralleling the SML methods applied here [262]. In this context, integrating SML into digital twin models involves using real-time predictions to continuously update the virtual representation of the physical system. This allows for the timely simulation of future scenarios, supports proactive decision-making, and improves the system's responsiveness to evolving conditions. Such an approach enhances the predictive and adaptive capabilities of digital twins, particularly in managing decentralised and dynamic energy environments.

Adaptive Control Mechanisms and Dynamic Energy Flow Management

Adaptive control mechanisms within the fractal microgrid structure play a pivotal role in optimizing energy utilization and balancing real-time electricity demand and renewable generation. The implemented decision-making logic dynamically controls battery charging and discharging by integrating real-time load predictions from SML and solar generation forecasts. This ensures effective energy management, significantly reducing dependence on external grid resources [32].

The energy management strategy continuously assesses forecasted household consumption relative to expected solar generation. When predicted solar generation exceeds demand, surplus energy is preferentially stored in batteries, considering maximum storage capacity constraints. Conversely, if forecasted consumption exceeds generation, the system evaluates the battery state-of-charge, which should never fall below the critical threshold of 20% of total battery capacity and supplies the necessary power from battery storage accordingly. Only when battery storage drops below this threshold does the system draw electricity from the external grid, ensuring reliability without compromising battery health through excessive deep discharges.

Reducing grid dependency through this adaptive policy also positively impacts the operational lifetime and efficiency of battery storage. Frequent charge-discharge cycles and deep discharging significantly accelerate battery degradation, leading to capacity loss over time. The developed control logic explicitly mitigates these risks by limiting unnecessary cycling and avoiding frequent deep discharges, thus extending battery lifespan and maintaining high operational efficiency. Such practices align well with recent recommendations on maximizing the operational efficiency and lifespan of lithium-ion battery systems, which exhibit high efficiency (85–95%) but require carefully controlled charging and discharging to minimize degradation [33].

Evaluations conducted with real-world consumption profiles and solar data validate the effectiveness of the adaptive management approach. The results demonstrated substantial reductions in grid energy consumption, fewer battery charge-discharge cycles, and enhanced temporal energy balance. This outcome not only improves energy efficiency but also optimizes long-term battery health and reduces maintenance and operational costs [34].

By leveraging predictive analytics and adaptive control mechanisms, the developed framework ensures optimal utilization of renewable energy resources, extends battery system lifespan, and significantly reduces grid reliance, providing a robust and efficient solution for decentralized energy management in modern microgrids.

Scenario-Based Battery Size Optimization

Battery sizing significantly influences the economic viability and operational efficiency of microgrid systems. To thoroughly evaluate battery size optimization, this research explores three distinct scenarios employing detailed analytical methodologies and predictive insights derived from SML. Each scenario applies unique logic, providing a comparative view of the centralized, distributed, and hybrid approaches.

- Scenario 1: Centralized General Battery

In the centralized scenario, a single large-scale battery system is optimized to serve the aggregated energy demands of the entire microgrid community. The Python optimization approach employs a constrained minimization method, explicitly defining the battery capacity required to meet the community's highest observed peak demand. This optimization uses a well-defined objective function that seeks the minimal battery capacity sufficient to handle the most extreme consumption scenario, ensuring uninterrupted power supply throughout varying load conditions.

This centralised battery strategy has inherent advantages, primarily through the economy of scale, allowing for potentially lower unit costs and simplified centralised management. However, centralised systems inherently demand robust distribution infra-structures to manage energy flows across diverse households, which may also lead to in-creased transmission distances, greater cable losses, or the need for higher-specification cabling to reduce those losses. Additionally, reliability hinges on accurate prediction of peak loads, highlighting the importance of real-time analytics provided by SML methods [261]. The optimized general battery size thus represents a carefully balanced decision between capacity, cost, and reliability, acknowledging the critical importance of precise forecasting.

- Scenario 2: Uniform Personal Batteries

Conversely, the second scenario examines a decentralized approach, allocating identical distributed batteries to each household. The sizing logic assesses individual peak demands across various households. It employs three distinct sizing metrics: the maximum peak, the 95th percentile of peak demands, and the average peak demand. Among these, the 95th percentile metric was selected due to its optimal balance between over- and under-capacity pro-visioning. Unlike the absolute maximum, the 95th percentile offers significant resilience without excessive investment, thereby balancing cost-effectiveness and reliability across the community [263].

Adopting a uniform personal battery size simplifies battery management and maintenance but may lead to underutilization in households with lower demands or occasional shortfalls in homes with notably higher peaks. Nonetheless, this uniform strategy significantly reduces operational complexity and provides households autonomy over energy usage, effectively enhancing localized resilience and empowering individual nodes within the fractal architecture [264].

- Scenario 3: Hybrid (General and Personal Batteries)

Recognizing the distinct strengths and limitations of centralized and decentralized systems, the hybrid scenario integrates both a general community-level battery and personal household batteries. Python simulations underpinning this scenario reveal a complementary strategy where personal batteries primarily manage regular household-level fluctuations, while the general battery provides additional backup capacity, strategically sized at a fraction of the combined personal battery capacities.

This dual-layer optimization logic dynamically assesses both solar generation and predicted consumption patterns. In periods of widespread generation surplus, the general battery absorbs excess energy, which personal batteries may not individually accommodate. Conversely, during widespread deficits, particularly under adverse weather conditions or high-demand events, the general battery supplements personal storage, preventing deep discharge cycles and extending the operational lifespan of the household batteries [265].

Visual analysis derived from Python simulations illustrates clear operational benefits. The simulations demonstrate reduced grid usage, optimized battery cycling, and improved energy allocation between the general and personal batteries. Consequently, the general battery substantially enhances system robustness by managing broader community-wide fluctuations, while the distributed personal batteries efficiently address localized, daily energy variations. The combination markedly enhances grid resilience, minimizes total battery cycle frequency [266], and reduces long-term maintenance and operational costs [267].

These scenario-based analyses collectively offer critical insights into optimal battery sizing approaches within fractal-structured microgrids. By leveraging predictive analytics through SML, each scenario is precisely tailored to the microgrid's consumption profiles, ensuring cost-effective, reliable, and efficient energy storage and management solutions.

6.4 Comparative Evaluation of Energy Management and Forecasting Strategies

Accurate forecasting and adaptive energy management strategies are central to enhancing the operational efficiency and reliability of renewable energy-based microgrids. In this context, the developed SML model, leveraging Hoeffding Trees, offers a robust solution to predict real-time electricity consumption for ten residential households and a marae with high precision. A *marae* is a traditional *Māori* gathering place typically consisting of a ceremonial area *marae ātea*, a meeting house *wharenui*, and a kitchen or dining area *wharekai*. It hosts smaller meetings *hui* as well as larger tribal *iwi* events lasting multiple days, with overnight stays usually in the *wharenui* [257]. The effectiveness of this predictive framework, characterized by notably low MSE, forms the foundation for evaluating subsequent energy management policies aimed at minimizing grid dependency and optimizing local energy utilization. The subsequent storage scenarios are presented as a proof-of-concept comparison of storage topologies rather than an engineering design exercise; numerical values are therefore interpreted conceptually, with realism discussed explicitly later in this section.

To validate and illustrate the model's predictive accuracy, Figure 6.3 presents a detailed comparison between actual electricity consumption and predicted values for representative households and the marae. Although there are minor discrepancies in magnitude between predicted and actual values, the model effectively captures the overall trends and fluctuations in electricity demand. Leveraging incremental learning capabilities, the proposed model continuously updates, achieving robust predictive performance with MAE ranging from 54.69 kWh to 320.71 kWh, representing less than 10% of typical hourly consumption values. The prediction model was based on household electricity demand profiles combined with lagged consumption values and normalized to ensure comparability across households. For evaluation, the dataset was partitioned into training and test subsets, and performance was reported on the unseen test data to confirm generalizability. This explicit validation process supports the reliability of the reported MAE values. MAE was selected in this chapter rather than RMSE because it provides a directly interpretable measure of average error in kWh and is less sensitive to extreme outliers, which are common in household demand data. Earlier chapters used RMSE to emphasize penalization of larger errors in turbine prediction, while here MAE aligns better with the practical context of residential microgrids [268].

The close alignment between actual and forecasted values demonstrates the model's capability to consistently deliver reliable predictions, thus enabling proactive and informed decision-making within the energy management framework. These results are consistent with recent studies that highlight the suitability of streaming machine learning methods, such as Hoeffding Trees and their variants, for handling demand forecasting under concept drift in residential microgrids [269]. To quantitatively assess the model's resilience to concept drift during dynamic operating conditions, an Augmented Dickey–Fuller (ADF) test was performed on the pre-diction error series. The results (ADF statistic: −17.06, p-value: 7.99 × 10−30) confirmed the stationarity of the error series, reinforcing the model's robustness to evolving data patterns.

Leveraging this predictive capability, adaptive management algorithms are implemented to dynamically regulate local energy resources, effectively minimising the community's dependence on an external electricity grid. Critical insights into the effectiveness of this strategy are presented through cumulative grid usage metrics, as depicted in Figure 6.4. Here, the performance of the predictive management approach (green line) is compared with traditional management relying solely on historical consumption data without predictive capabilities (red line). As demonstrated, predictive management substantially reduces cumulative grid usage over the evaluated period. This reduction not only underscores improved efficiency in energy allocation but also represents a meaningful step to-ward greater energy independence and sustainability within the microgrid community.

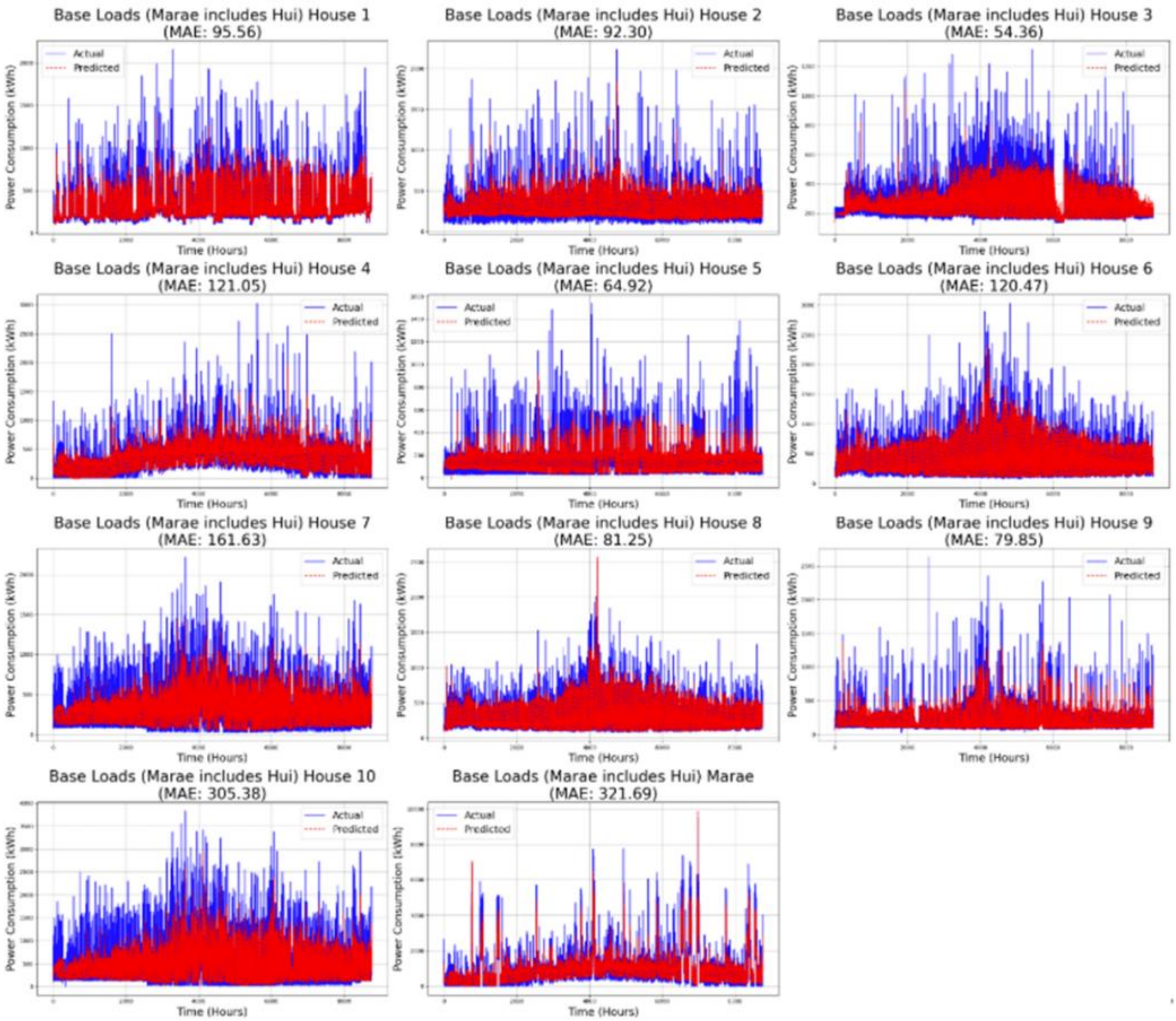


Figure 6.3: Actual versus predicted electricity consumption for each of the ten houses and *Marae*.

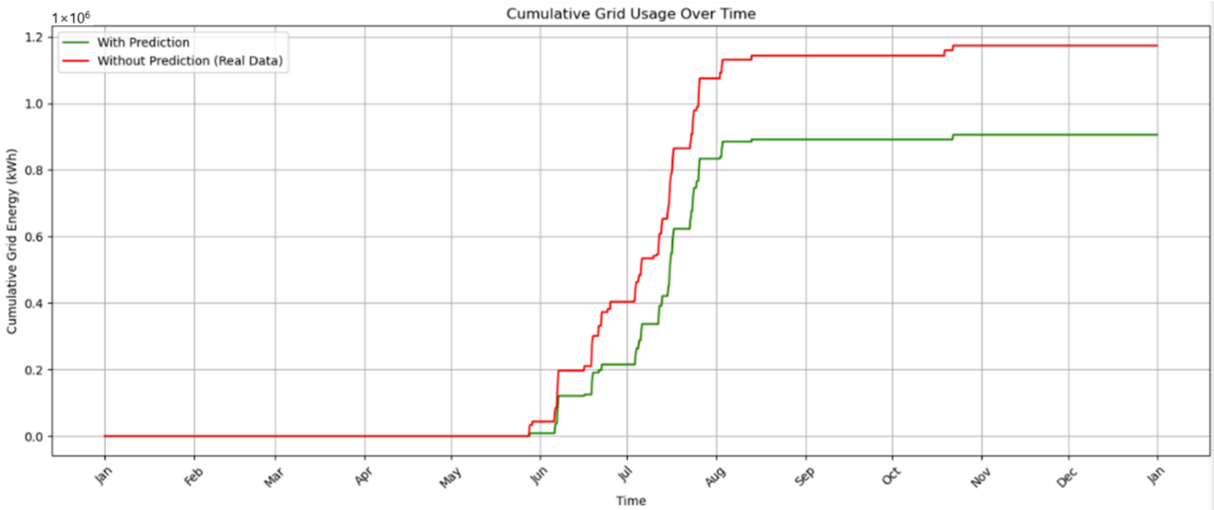


Figure 6.4: Cumulative Grid Usage Over Time (Forecast-based vs. Actual Management).

Further assessment of system resilience is conducted by examining occurrences of energy deficit events, defined as periods when local generation and available battery storage fail to satisfy the community's immediate energy requirements. Figure 6.5 provides a comparative visualization of these deficit events, clearly indicating fewer occurrences under the predictive energy management strategy. The graph illustrates that deficit events predominantly occur during winter months (June–August), aligning with decreased solar generation and increased energy demand. Notably, the predictive strategy, when com-pared to a baseline scenario using the same energy storage infrastructure but operating without real-time forecasting, significantly delays and reduces reliance on grid-supplied energy. By the year's end, the predictive approach results in approximately a 24% reduction in cumulative grid energy use. This reduction in energy shortfalls directly translates into improved reliability and optimized utilization of available energy resources, reinforcing the practical benefits of integrating predictive analytics into real-time adaptive man-agreement processes.

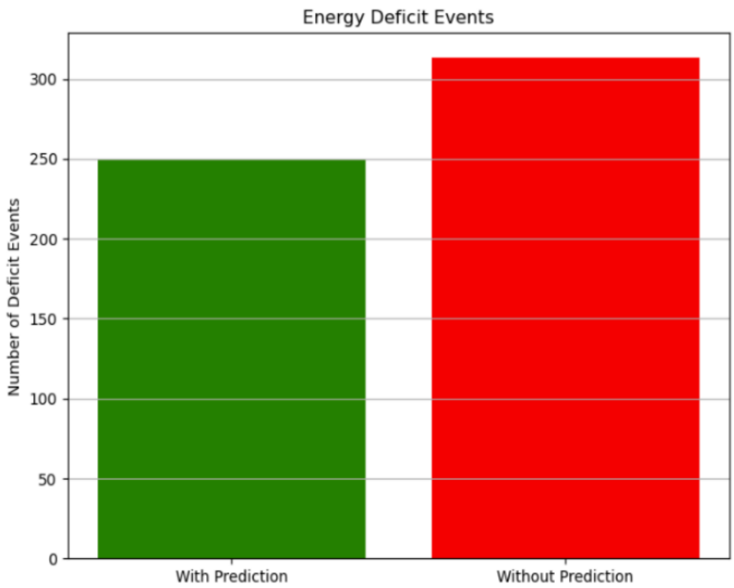


Figure 6.5: Comparison of Energy Deficit Events (Forecast-based vs. Actual Management).

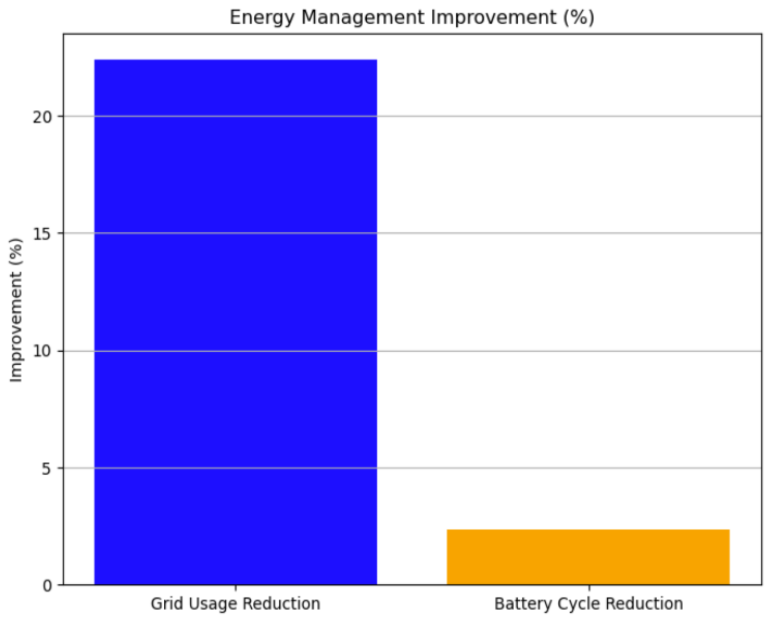
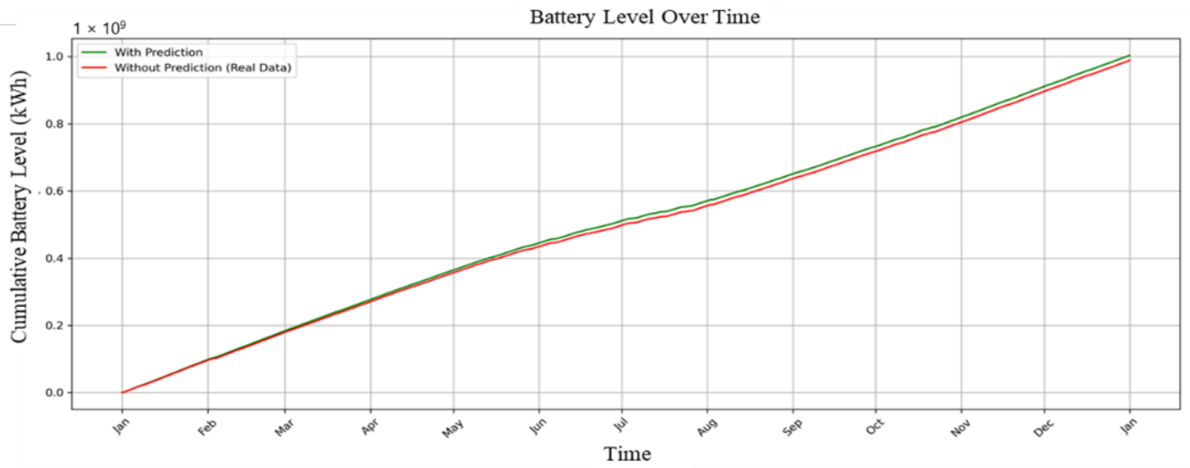


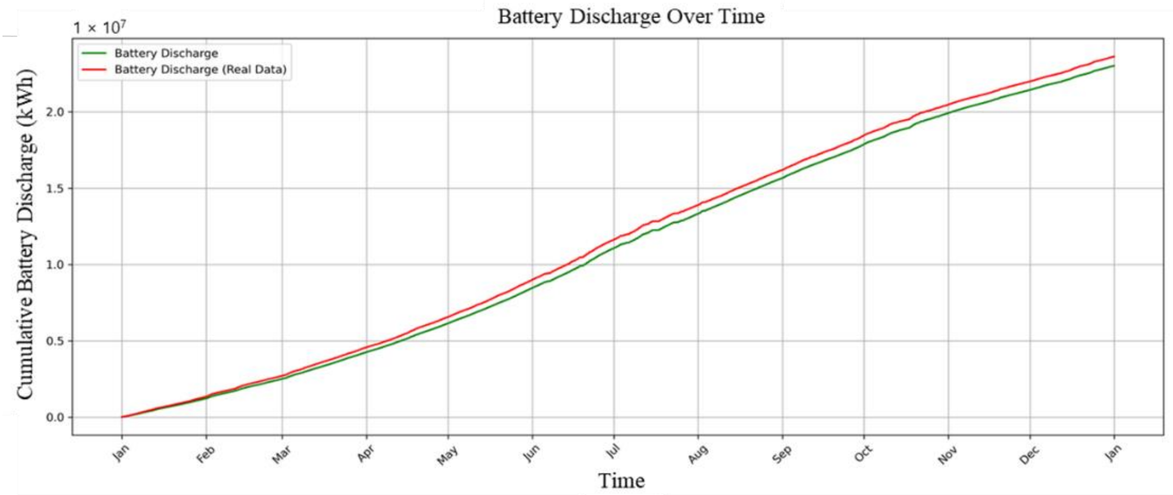
Figure 6.6: Energy Management Improvement, Grid Dependence and Battery Cycle Reduction.

Complementary reliability improvements, the evaluation includes a detailed analysis of the broader benefits derived from predictive energy management strategies, particularly regarding reducing grid dependence and optimized battery utilization. As depicted in Figure 6.6, predictive energy management results in a 24% reduction in grid dependency and a 2.5% reduction in battery cycling. Lower battery cycling frequency directly enhances battery longevity, reducing maintenance requirements and operational expenditures over the microgrid's lifespan.

To provide deeper insights into battery usage dynamics, cumulative analyses of battery charging and discharging patterns are presented in Figure 6.7, further demonstrating improved battery management efficiency under predictive conditions. Although visually subtle, the cumulative curves exhibit a gradually increasing divergence over the year, particularly noticeable during high solar generation periods. This divergence indicates fewer unnecessary battery cycles under predictive conditions, resulting in approximately a 2–3% reduction in total charged and discharged energy, or roughly a dozen fewer full battery cycles annually. Such modest improvements substantially contribute to battery longevity by slowing capacity fading, reducing energy losses, and minimizing the frequency of battery replacements, thereby significantly enhancing operational sustainability.



(a)



(b)

Figure 6.7: Battery performance analysis: (a) cumulative battery level vs. time, (b) cumulative battery discharge vs. time.

Additionally, Table 6.1 summarizes the comparative performance of three distinct battery-sizing scenarios: a centralized general battery (Scenario 1), uniform personal batteries (Scenario 2), and a hybrid approach combining both centralized and personal storage (Scenario 3). Scenario 1 employs a centralized battery (198.42 kWh), resulting in moderate total grid usage (~1.41×10^3 kWh) and relatively low battery cycling (421.38 cycles), indicating efficient centralized management but at the expense of significant infrastructure and distribution costs.

Scenario 2 uses identically sized household batteries (14.87 kWh each) controlled for local self-consumption. This capacity falls within the typical range of residential batteries, which are generally sized in the tens of kWh [270]. Compared to centralized storage, the distributed configuration reduces grid reliance but results in higher battery cycling, reflecting the greater frequency of charge–discharge events at the household level.

Scenario 3 combines the distributed rule with a small central buffer (82.14 kWh) for community balancing. The higher grid use and cycling observed here should be interpreted conceptually (trade-offs between autonomy and aggregation) rather than as a capacity-optimal design, given the unequal total storage across scenarios. Although the total grid usage increases notably (2.53×10^3 kWh), partially due to the smaller centralised backup, this hybrid scenario effectively manages community-level energy variability by providing enhanced flexibility and improved robustness against widespread deficits or surpluses. The trade-off observed in Scenario 3, while resulting in higher grid dependence and battery cycling, is particularly justified in operating contexts where ensuring broader system stability, community-level energy balancing, and resilience during abnormal events or critical conditions is prioritised over strict minimisation of grid imports or battery usage frequency. In such scenarios, the hybrid approach offers a pragmatic balance between autonomy and centralised support, which can better accommodate unpredictable generation and demand patterns. Nevertheless, it also records a substantial increase in battery cycling events (3192.74 cycles), highlighting a clear trade-off between system flexibility and battery longevity.

Table 6.1: Limitations of Static Energy Management Systems in Fractal Microgrids.

Scenario #	Distributed Battery Size (kWh)	Central Battery Size (kWh)	Total Battery Capacity (kWh)	Total Grid Usage (kWh)	Total Battery Cycles
1	-	198.42	198.42	1.41×10^3	421.38
2	14.87	-	163.57	6.12×10^2	879.62
3	14.87	82.14	245.71	2.53×10^3	3,192.74

Overall, these scenarios are intended to contrast storage topologies under simple, transparent control rules. Because Scenario 2 employed relatively small household capacities, the three cases do not constitute an equal-capacity benchmark. Interpreted conceptually, the results highlight expected trade-offs, centralized aggregation reduces grid reliance with fewer cycles, while distributed autonomy enhances resilience but at higher cycling cost. In line with empirical sizing studies [270], the adopted household battery capacity of ~15 kWh falls within the practical range for residential applications, reinforcing the realism of the chosen parameters.

6.5 Summary

This study presented an adaptive energy management strategy for fractal-structured residential microgrids, leveraging SML to address the complexities introduced by renewable energy integration and dynamic electricity demand. Employing Hoeffding Trees within a Bagging Regressor ensemble, the proposed method showed promising predictive performance, maintaining real-time accuracy under continuously evolving load and generation conditions. Results suggested operational improvements in this case study, notably reducing cumulative grid dependence by approximately 22.84%, decreasing energy deficit events by 20%, and optimizing battery cycling frequency to extend the lifecycle of storage systems.Three battery sizing scenarios centralised storage, distributed batteries, and a hybrid approach were evaluated, each illustrating distinct benefits and trade-offs.

Centralised storage provided effective load management with fewer battery cycles but introduced challenges related to infrastructure requirements and resilience. Conversely, distributed battery storage improved local autonomy and reduced reliance on centralised systems but led to increased cycling events. The hybrid approach balanced these extremes, offering flexibility in managing community-level variability, though at the cost of increased cycling and grid dependence in this case study.

The findings from this case study suggest the practicality of integrating advanced streaming analytics into existing microgrid management infrastructures, with potential to promote sustainable energy utilisation, resilience, and operational efficiency. The application of real-time forecasting facilitated intelligent decision-making, enabling proactive grid-edge trading and adaptive battery management. Furthermore, the integration of SML aligns closely with digital twin methodologies, providing a scalable and replicable framework that can dynamically respond to evolving conditions within diverse microgrid configurations. Importantly, the introduction of a lightweight yet scalable learning model in this case study points to the potential for a computationally efficient baseline to support real-time microgrid management.

However, the effectiveness of the proposed framework may be influenced by certain limitations inherent in the applied methods and data conditions. Forecast accuracy could be sensitive to the representativeness and variability of the historical data, particularly under unusual or extreme operational scenarios. Moreover, while Hoeffding Trees offer advantages in computational efficiency, their relatively simple structure might restrict their capability to fully capture complex nonlinear relationships compared to more sophisticated ML architectures.

Future studies can build on this foundation by conducting comparative evaluations with other streaming algorithms, such as Adaptive Random Forests, and by implementing more comprehensive hyperparameter tuning and drift adaptation techniques (e.g., ADWIN and DDM) to further improve adaptability and forecasting accuracy under non-stationary conditions. Additional research could also investigate economic optimisation strategies, integrating pricing signals and market dynamics to enhance grid-edge trading. Expanding this framework to accommodate diverse renewable resources and various geographic contexts would further demonstrate the broader applicability and transformative potential of SML in decentralised energy systems.

7. Conclusion

7.1 Thesis Summary

This thesis has presented an integrated and comprehensive approach to addressing the challenges of renewable energy forecasting, management, and optimization within both industrial and residential energy systems. Motivated by the increasing need for adaptive and intelligent energy solutions, especially in the context of New Zealand's transition to a sustainable and low-carbon energy future, this research has applied ML techniques, digital twin frameworks, and adaptive control strategies to enhance the efficiency and resilience of modern power systems.

Chapter 1 of this thesis introduced the central challenge addressed throughout the research: the inherent variability and unpredictability of renewable energy sources such as solar and wind. This core challenge directly relates to Research Question 1 (RQ1), which asked how FL could be used to manage variability and uncertainty in electricity sources. This question was primarily addressed in Chapter5, where FL controllers were integrated with ML-based solar irradiance forecasting to dynamically regulate electricity consumption in an industrial case study. Furthermore, Chapter 6 extended this approach to community microgrids, incorporating adaptive control strategies within fractal energy networks, suggesting that FL can contribute to stabilizing operation under the tested conditions.

RQ2 examined the applicability of ML methods such as LSTM, CNN, and RNN for accurate energy forecasting. This question was explored in depth in Chapters 4 and 5. To ensure narrative continuity, Chapter 4 introduced the use of LSTM and CNN in predicting steam turbine output, outperforming traditional thermodynamic models like the Willans line in capturing time-dependent and nonlinear behaviors. Chapter5 expanded the use of ML to forecast solar radiation, with the results suggesting the potential strengths of deep learning models in improving prediction accuracy. These findings underscore the vital role of ML in transforming energy forecasting by leveraging temporal patterns and complex dependencies.

To answer RQ3, which focused on identifying optimal input parameters for FL-based management systems, a combination of sensitivity and correlation analyses was employed (detailed in Chapter5). These methods helped identify the most impactful variables on system performance—such as solar irradiance, battery state-of-charge, and load profile—enabling more effective rule-based decision-making. This targeted approach to input selection enhanced controller responsiveness and improved system adaptability.

RQ4 concerned the identification and prioritization of renewable electricity sources for community and industrial clusters, with consideration for environmental, operational, and economic factors. This was addressed through a synthesis of literature (Section 2.4.2) and simulation studies in Chapter 6. A comparative evaluation of renewable potentials and integration strategies provided insights into optimal energy mix configurations and resource deployment, balancing performance with sustainability. The use of fractal-structured microgrids in Chapter6 particularly demonstrated how localized and hierarchical planning could maximize renewable use while minimizing storage and infrastructure costs.

RQ5 brought together the full scope of the thesis by asking how ML methods could be used to size, integrate, and manage renewable systems within clusters. This was comprehensively addressed in Chapters 4, 5, and 6. Forecasting models developed in Chapters 4 and 5 fed directly into the control logic and energy management systems presented in Chapter6. There, the integration of SML and digital twins facilitated real-time adaptation, enabling efficient load balancing, local energy trading, and minimized emissions. The results suggest that ML-driven design and control could improve both economic and environmental performance.

A key theme uniting the different aspects of this thesis is the integration of predictive analytics with adaptive control mechanisms, supported by real-time data streams and digital twin frameworks. This synergy enables energy systems to function as adaptive, self-correcting networks, capable of maintaining stability and efficiency in the face of changing conditions. In industrial contexts, the combination of deep learning models and FL controllers allows systems to forecast energy generation or demand and translate these forecasts into actionable control strategies. In residential microgrids, the use of SML models and hierarchical controllers optimizes the operation of distributed renewable resources and storage systems. These adaptive architectures not only enhance operational efficiency but also contribute to reducing dependence on centralized grids and improving the sustainability of energy systems.

By focusing on both industrial and residential energy applications, this thesis demonstrates that the fusion of ML, digital twins, and adaptive control offers a scalable and robust framework for addressing the complexities of renewable energy integration. This research underscores the potential for digital twin methodologies to enable real-time monitoring, predictive modeling, and proactive control in energy systems, supporting the transition to low-carbon and resilient infrastructures.

In conclusion, this thesis presents a comprehensive and integrated approach to optimizing renewable energy utilization through the innovative application of ML and digital twin technologies. By addressing the five core research questions through applied experimentation and theoretical development across multiple chapters, the research provides practical insights and scalable solutions for enhancing energy efficiency, resilience, and sustainability. These contributions align with New Zealand's goals for carbon emission reduction and renewable energy development while offering a model for broader global adoption of intelligent energy management systems.

7.2 Future Work

While this thesis has demonstrated the feasibility and effectiveness of integrating ML and digital twin technologies into renewable energy systems, several avenues for further exploration remain. Future studies can expand the current work by addressing its limitations, exploring emerging technologies, and applying the developed frameworks to broader use cases.

One key direction involves extending the solar monitoring system developed in Chapter 3. Although the system achieved its goal of providing a low-cost, real-time solution for energy monitoring, future work could involve integrating more sophisticated environmental sensors (e.g., wind speed, humidity) and enhancing its communication and computational capabilities. Furthermore, embedding AI for real-time anomaly detection or performance degradation analysis could enhance the digital twin's functionality.

In the predictive modeling domain explored in Chapter 4, future research could explore alternative deep learning architectures such as Transformer-based models or attention-enhanced LSTM variants. These models may offer improved accuracy or interpretability for long-term forecasting under complex temporal dependencies. Additionally, more rigorous methods for data preprocessing and feature engineering—particularly under noisy or incomplete data conditions—could further enhance model robustness.

The integration of FL and ANNs in Chapter 5 has shown promise for adaptive energy management. However, expanding this framework to include reinforcement learning or MPC could enable systems that not only adapt but also learn optimal strategies over time. Comparative evaluations between FL and newer control algorithms would offer valuable insights into operational trade-offs.

In the context of SML for microgrid control, as presented in Chapter 6, future research should focus on refining the model's ability to handle concept drift and evolving consumption patterns more effectively. Incorporating advanced drift detection methods such as ADWIN or DDM and evaluating other SML algorithms such as Adaptive Random Forests or Leveraging Bagging, could improve forecasting reliability under non-stationary conditions. Economic factors—such as electricity pricing and dynamic tariffs—could also be incorporated into the learning process to support financially optimized energy trading.

From a systems perspective, one promising direction is the integration of cyber-physical security into digital twin-based energy frameworks. As smart grids become increasingly connected, addressing vulnerabilities in data integrity and communication will be crucial for operational reliability.

Lastly, to validate the scalability and generalizability of the proposed models, future work should consider deploying them across a broader geographic range and under different climatic, regulatory, or socio-economic contexts. Real-world pilot deployments, potentially in collaboration with local energy utilities or industries, would help bridge the gap between research and implementation.

In summary, the foundation laid by this thesis opens up a rich set of opportunities for technical innovation and practical application in the field of intelligent energy systems. By advancing learning algorithms, integrating multi-objective control, and addressing deployment challenges, future research can continue to build toward smarter, more autonomous, and more sustainable energy infrastructures.

8. References

[1] United Nations, *The Paris Agreement*. [Online]. Available: https://www.un.org/en/climatechange/paris-agreement. Accessed: Jan. 6, 2023.

[2] United Nations, *COP27: Delivering for People and the Planet*. [Online]. Available: https://www.un.org/en/climatechange/cop27. Accessed: Jan. 6, 2023.

[3] Ministry for the Environment, *New Zealand's Greenhouse Gas Inventory 1990–2020 Snapshot*. [Online]. Available: https://environment.govt.nz/publications/new-zealands-greenhouse-gas-inventory-1990-2020-snapshot/. Accessed: Feb. 22, 2023.

[4] Ministry of Business, Innovation & Employment, *Process Heat Current State Fact Sheet*. [Online]. Available: https://www.mbie.govt.nz/dmsdocument/152-process-heat-current-state-fact-sheet-pdf. Accessed: Feb. 27, 2023.

[5] Ministry of Business, Innovation & Employment, *Programmes Funded Through the Strategic Science Investment Fund*. [Online]. Available: https://www.mbie.govt.nz/science-and-technology/science-and-innovation/funding-information-and-opportunities/investment-funds/strategic-science-investment-fund/ssif-funded-programmes/. Accessed: Feb. 9, 2023.

[6] University of Waikato, *Waikato Research to Support Transition to Net-Zero-Carbon Process Heat Sector by 2050*. [Online]. Available: https://www.waikato.ac.nz/news-opinion/media/2020/125m-project-to-reduce-greenhouse-gas-emissions-using-digital-twin-technology. Accessed: Feb. 9, 2023.

[7] M. Pasandideh, M. Taylor, S. R. Tito, M. Atkins, and M. Apperley, "Predicting steam turbine power generation: A comparison of long short-term memory and Willans line model," *Energies*, vol. 17, no. 2, art. no. 352, Jan. 2024, doi: 10.3390/en17020352.

[8] G. Notton, M. Nivet, C. Voyant, A. Fouilloy, F. Motte, and M. Cristofari, "Intermittent and stochastic character of renewable energy sources: Consequences, cost of intermittence and benefit of forecasting," *Renewable and Sustainable Energy Reviews*, vol. 87, pp. 96–105, May 2018, doi: 10.1016/j.rser.2018.02.007.

[9] M. Pasandideh, S. R. Tito, M. Apperley, and M. Atkins, "Design and implementation of a simple low-cost and real-time solar power monitoring system," in *Proc. 2023 IEEE Fifth Int. Conf. DC Microgrids (ICDCM)*, Nov. 2023, pp. 1–6, doi: 10.1109/ICDCM54452.2023.10433596.

[10] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, Jun. 1965, doi: 10.1016/S0019-9958(65)90241-X.

[11] A. G. Olabi and M. A. Abdelkareem, "Renewable energy and climate change," *Renewable and Sustainable Energy Reviews*, vol. 158, p. 112111, Apr. 2022, doi: 10.1016/j.rser.2022.112111.

[12] L. Sun and R. Smith, "Performance modeling of new and existing steam turbines," *Industrial & Engineering Chemistry Research*, vol. 54, no. 6, pp. 1671–1681, Feb. 2015, doi: 10.1021/ie5032309.

[13] P. Domingos and G. Hulten, "Mining high-speed data streams," in *Proc. 6th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, New York, NY, USA, Aug. 2000, pp. 71–80, doi: 10.1145/347090.347107.

[14] K. Szostek, D. Mazur, G. Drałus, and J. Kusznier, "Analysis of the effectiveness of ARIMA, SARIMA, and SVR models in time series forecasting: A case study of wind farm energy production," *Energies*, vol. 17, no. 19, art. no. 4803, Jan. 2024, doi: 10.3390/en17194803.

[15] J. Brownlee, *Deep Learning for Time Series Forecasting: Predict the Future with MLPs, CNNs and LSTMs in Python*, Queensland, Australia: Machine Learning Mastery, 2018.

[16] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 3rd ed. [Online]. Available: https://otexts.com/fpp3/. Accessed: Jan. 6, 2025.

[17] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and Its Applications: With R Examples*, 4th ed. Cham, Switzerland: Springer, 2017, doi: 10.1007/978-3-319-52452-8.

[18] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*, 5th ed. Hoboken, NJ, USA: Wiley, 2015. [Online]. Available: https://www.wiley.com/en-se/Time+Series+Analysis%3A+Forecasting+and+Control%2C+5th+Edition-p-9781118675021. Accessed: Jan. 6, 2025.

[19] Z. Hajirahimi and M. Khashei, "Hybridization of hybrid structures for time series forecasting: A review," *Artificial Intelligence Review*, vol. 56, no. 2, pp. 1201–1261, Feb. 2023, doi: 10.1007/s10462-022-10199-0.

[20] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[21] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, Cambridge, MA, USA: MIT Press, 2016. [Online]. Available: https://mitpress.mit.edu/9780262035613/deep-learning/. Accessed: Jan. 6, 2025.

[22] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should I trust you?': Explaining the predictions of any classifier," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, San Francisco, CA, USA, Aug. 2016, pp. 1135–1144, doi: 10.1145/2939672.2939778.

[23] P. Goodwin, "The Holt-Winters approach to exponential smoothing: 50 years old and going strong," *Foresight: The International Journal of Applied Forecasting*, no. 19, pp. 30–33, 2010.

[24] P. R. Winters, "Forecasting sales by exponentially weighted moving averages," *Management Science*, vol. 6, no. 3, pp. 324–342, 1960.

[25] B. Park, H. Kim, Y. Choi, and J. Yoon, "Time series ARIMA model for prediction of daily and monthly average global solar radiation: The case study of Seoul, South Korea," Symmetry, vol. 11, no. 2, p. 240, Feb. 2019. [Online]. Available: https://www.mdpi.com/2073-8994/11/2/240. Accessed: Jun. 15, 2025.

[26] E. Chodakowska, J. Nazarko, Ł. Nazarko, H. S. Rabayah, R. M. Abendeh, and R. Alawneh, "ARIMA models in solar radiation forecasting in different geographic locations," *Energies*, vol. 16, no. 13, art. no. 5029, Jan. 2023, doi: 10.3390/en16135029.

[27] E. Zhao, J. Zhao, L. Liu, Z. Su, and N. An, "Hybrid wind speed prediction based on a self-adaptive ARIMAX model with an exogenous WRF simulation," *Energies*, vol. 9, no. 1, art. no. 7, Jan. 2016, doi: 10.3390/en9010007.

[28] A. Karakan, "Predicting energy production in renewable energy power plants using deep learning," *Energies*, vol. 17, no. 16, art. no. 4031, Jan. 2024, doi: 10.3390/en17164031.

[29] Y. Zhang, Y. Peng, X. Qu, J. Shi, and E. Erdem, "A finite mixture GARCH approach with EM algorithm for energy forecasting applications," *Energies*, vol. 14, no. 9, art. no. 2352, Jan. 2021, doi: 10.3390/en14092352.

[30] J.-P. Lai, Y.-M. Chang, C.-H. Chen, and P.-F. Pai, "A survey of machine learning models in renewable energy predictions," Appl. Sci., vol. 10, no. 17, p. 5975, Sep. 2020, doi: 10.3390/app10175975.

[31] Z. Yao, S. Huang, Z. Li, J. Wang, X. Chen, and J. Zhang, "Machine learning for a sustainable energy future," Nat. Rev. Mater., vol. 8, pp. 996–1018, Jul. 2023, doi: 10.1038/s41578-022-004905.

[32] IBM, "What are Neural Networks?", [Online]. Available: https://www.ibm.com/topics/neural-networks. Accessed: Dec. 30, 2022.

[33] N. J. Sairamya, L. Susmitha, S. T. George, and M. S. P. Subathra, "Hybrid approach for classification of electroencephalographic signals using time–frequency images with wavelets and texture features," in *Intelligent Data Analysis for Biomedical Applications*, D. J. Hemanth, D. Gupta, and V. E. Balas, Eds., in *Intelligent Data-Centric Systems*, Academic Press, 2019, pp. 253–273, doi: 10.1016/B978-0-12-815553-0.00013-6.

[34] R. Kruse, S. Mostaghim, C. Borgelt, C. Braune, and M. Steinbrecher, "Multi-layer perceptrons," in *Computational Intelligence: A Methodological Introduction*, in *Texts in Computer Science*, Cham: Springer, 2022, pp. 53–124, doi: 10.1007/978-3-030-42227-1_5.

[35] S. Ibrić, J. Djuriš, J. Parojčić, and Z. Djurić, "Artificial neural networks in evaluation and optimization of modified release solid dosage forms," *Pharmaceutics*, vol. 4, no. 4, pp. 531–550, Oct. 2012, doi: 10.3390/pharmaceutics4040531.

[36] D. J. Matich, "Redes Neuronales: Conceptos Básicos y Aplicaciones," *Universidad Tecnológica Nacional*, 2001. [Online]. Available: https://www.frro.utn.edu.ar/5_anio/monografias. Accessed: Jun. 18, 2025.

[37] H. Hu, S. Gong, and B. Taheri, "Energy demand forecasting using convolutional neural network and modified war strategy optimization algorithm," *Heliyon*, vol. 10, no. 6, p. e27353, Mar. 2024, doi: 10.1016/j.heliyon.2024.e27353.

[38] M. Alhussein, K. Aurangzeb, and S. I. Haider, "Hybrid CNN-LSTM model for short-term individual household load forecasting," *IEEE Access*, vol. 8, pp. 180544–180557, 2020, doi: 10.1109/ACCESS.2020.3028281.

[39] J. Chung and B. Jang, "Accurate prediction of electricity consumption using a hybrid CNN-LSTM model based on multivariable data," *PLOS ONE*, vol. 17, no. 11, p. e0278071, Nov. 2022, doi: 10.1371/journal.pone.0278071.

[40] E. M. Al-Ali, Y. Hajji, Y. Said, M. Hleili, A. M. Alanzi, A. H. Laatar, and M. Atri, "Solar Energy Production Forecasting Based on a Hybrid CNN-LSTM-Transformer Model," *Mathematics*, vol. 11, no. 3, art. no. 676, Jan. 2023, doi: 10.3390/math11030676. [Online]. Available: https://doi.org/10.3390/math11030676. Accessed: Jun. 18, 2025.

[41] D. Salman, C. Direkoglu, M. Kusaf, and M. Fahrioglu, "Hybrid deep learning models for time series forecasting of solar power," *Neural Comput. Appl.*, vol. 36, no. 16, pp. 9095–9112, Jun. 2024, doi: 10.1007/s00521-024-09558-5.

[42] "Papers with Code – LSTM explained." [Online]. Available: https://paperswithcode.com/method/lstm. Accessed: Dec. 27, 2022.

[43] J. Park, D. Yi, and S. Ji, "Analysis of recurrent neural network and predictions," *Symmetry*, vol. 12, no. 4, Apr. 2020, doi: 10.3390/sym12040615.

[44] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," *Physica D*, vol. 404, p. 132306, Mar. 2020, doi: 10.1016/j.physd.2019.132306.

[45] O. Almqvist, *A comparative study between algorithms for time series forecasting on customer prediction: An investigation into the performance of ARIMA, RNN, LSTM, TCN and HMM*, B.Sc. thesis, University of Skövde, 2019. [Online]. Available: http://urn.kb.se/resolve?urn=urn:nbn:se:his:diva-16974

[46] C. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Clim. Res.*, vol. 30, pp. 79–82, 2005, doi: 10.3354/cr030079.

[47] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature," *Geosci. Model Dev.*, vol. 7, no. 3, pp. 1247–1250, Jun. 2014, doi: 10.5194/gmd-7-1247-2014.

[48] Y. Zhao and L. K. Foong, "Predicting electrical power output of combined cycle power plants using a novel artificial neural network optimized by electrostatic discharge algorithm," *Measurement*, vol. 198, p. 111405, Jul. 2022, doi: 10.1016/j.measurement.2022.111405.

[49] D. N. Moriasi, J. G. Arnold, M. W. Van Liew, R. L. Binger, R. D. Harmel, and T. L. Veith, "Model evaluation guidelines for systematic quantification of accuracy in watershed simulations," *Trans. ASABE*, vol. 50, no. 3, pp. 885–900, 2007, doi: 10.13031/2013.23153.

[50] F. Saeed and S. Aldera, "Adaptive renewable energy forecasting utilizing a data-driven PCA-transformer architecture," *IEEE Access*, vol. 12, pp. 109269–109280, 2024, doi: 10.1109/ACCESS.2024.3440226.

[51] O. Tsvetkova and T. B. M. J. Ouarda, "A review of sensitivity analysis practices in wind resource assessment," *Energy Convers. Manag.*, vol. 238, p. 114112, 2021, doi: 10.1016/j.enconman.2021.114112.

[52] E. Trillas, "Lotfi A. Zadeh: On the man and his work," *Sci. Iran.*, vol. 18, no. 3, pp. 574–579, Jun. 2011, doi: 10.1016/j.scient.2011.05.001.

[53] M. R. H. Mohd Adnan, A. Sarkheyli, A. Mohd Zain, and H. Haron, "Fuzzy logic for modeling machining process: A review," *Artif. Intell. Rev.*, vol. 43, no. 3, pp. 345–379, Mar. 2015, doi: 10.1007/s10462-012-9381-8.

[54] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Trans. Syst. Man Cybern.*, vol. SMC-15, no. 1, pp. 116–132, Jan. 1985, doi: 10.1109/TSMC.1985.6313399.

[55] V. S. Sharma, S. K. Sharma, and A. K. Sharma, "Cutting tool wear estimation for turning," *J. Intell. Manuf.*, vol. 19, no. 1, pp. 99–108, Feb. 2008, doi: 10.1007/s10845-007-0048-2.

[56] M. Kor, E. Abkhoshk, D. Tao, G. L. Chen, and H. Modarres, "Modeling and optimization of high chromium alloy wear in phosphate laboratory grinding mill with fuzzy logic and particle swarm optimization technique," *Miner. Eng.*, vol. 23, no. 9, pp. 713–719, Aug. 2010, doi: 10.1016/j.mineng.2010.04.009.

[57] A. Mellit, S. A. Kalogirou, L. Hontoria, and S. Shaari, “Artificial intelligence techniques for sizing photovoltaic systems: A review,” *Renew. Sustain. Energy Rev.*, vol. 13, no. 2, pp. 406–419, Feb. 2009, doi: 10.1016/j.rser.2008.01.006.

[58] M. A. Alghassab, “Fuzzy-based smart energy management system for residential buildings in Saudi Arabia: A comparative study,” *Energy Rep.*, vol. 11, pp. 1212–1224, Jun. 2024, doi: 10.1016/j.egyr.2023.12.039.

[59] Q. Zhang, N. Tang, J. Lu, W. Wang, L. Wu, and W. Kuang, “A hybrid RBF neural network based model for day-ahead prediction of photovoltaic plant power output,” *Frontiers in Energy Research*, vol. 11, Jan. 2024, doi: 10.3389/fenrg.2023.1338195.

[60] P. K. Shukla and S. P. Tripathi, “A review on the interpretability-accuracy trade-off in evolutionary multi-objective fuzzy systems (EMOFS),” *Information*, vol. 3, no. 3, pp. 256–277, Sep. 2012, doi: 10.3390/info3030256.

[61] M. S. Manshahia, V. Kharchenko, G.-W. Weber, and P. Vasant, Eds., *Advances in Artificial Intelligence for Renewable Energy Systems and Energy Autonomy*, in *EAI/Springer Innovations in Communication and Computing*, Cham: Springer, 2023, doi: 10.1007/978-3-031-26496-2.

[62] P. R. Shukla, J. Skea, R. Slade, A. Al Khourdajie, R. van Diemen, D. McCollum, M. Pathak, S. Some, P. Vyas, R. Fradera, M. Belkacemi, A. Hasija, G. Lisboa, S. Luz, and J. Malley, Eds., *Climate Change 2022: Mitigation of Climate Change. Working Group III Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge: Cambridge University Press, Jul. 2023, doi: 10.1017/9781009157926.

[63] H. Ritchie, “Global inequalities in CO_2 emissions,” *Our World in Data*, Aug. 2023. [Online]. Available: https://ourworldindata.org/inequality-co2

[64] Intergovernmental Panel on Climate Change, “Summary for Policymakers,” *AR6 Synthesis Report*, 2023. [Online]. Available: https://www.ipcc.ch/report/ar6/syr/summary-for-policymakers/

[65] S. M. Thomas, A.-G. Ausseil, J. Guo, A. Herzig, E. Khaembah, A. Renwick, E. Teixeira, T. van der Weerden, S. J. Wakelin, and I. Vetharaniam, "Exploring the role of high-value crops to reduce agricultural greenhouse gas emissions in New Zealand," *Regional Environmental Change*, vol. 24, no. 3, p. 105, Jul. 2024, doi: 10.1007/s10113-024-02267-8.

[66] J. Laubach, T. K. Flesch, C. Ammann, M. Bai, Z. Gao, L. Merbold, D. I. Campbell, J. P. Goodrich, S. L. Graham, J. E. Hunt, A. M. Wall, and L. A. Schipper, "Methane emissions from animal agriculture: Micrometeorological solutions for challenging measurement situations," *Agricultural and Forest Meteorology*, vol. 350, p. 109971, May 2024, doi: 10.1016/j.agrformet.2024.109971.

[67] E. Kliejunas, A. Cavadino, B. Kidd, C. Cleghorn, J. Drew, C. Ni Mhurchu, and K. E. Bradbury, "Quantifying the greenhouse gas emissions of New Zealand households’ food purchases: An analysis by demographic variables," *Journal of Cleaner Production*, vol. 430, p. 139699, Dec. 2023, doi: 10.1016/j.jclepro.2023.139699.

[68] S. M. Hickey, W. E. Bain, T. P. Bilton, G. J. Greer, S. Elmes, B. Bryson, C. S. Pinares-Patiño, J. Wing, A. Jonker, E. A. Young, K. Knowler, N. K. Pickering, K. G. Dodds, P. H. Janssen, J. C. McEwan, and S. J. Rowe, "Impact of breeding for reduced methane emissions in New Zealand sheep on maternal and health traits," *Frontiers in Genetics*, vol. 13, Sep. 2022, doi: 10.3389/fgene.2022.910413.

[69] International Monetary Fund, “New Zealand: Selected Issues,” May 2022. [Online]. Available: https://www.elibrary.imf.org/view/journals/002/2022/139/article-A003-en.xml

[70] Ministry for the Environment, “New Zealand’s Greenhouse Gas Inventory 1990–2023: Snapshot.” [Online]. Available: https://environment.govt.nz/publications/new-zealands-greenhouse-gas-inventory-19902023-snapshot/

[71] Ministry for the Environment, “Climate Change Response (Zero Carbon) Amendment Act 2019.” [Online]. Available: https://environment.govt.nz/acts-and-regulations/acts/climate-change-response-amendment-act-2019/

[72] Intergovernmental Panel on Climate Change, “Chapter 2: Emissions trends and drivers,” *AR6 WGIII Report*, 2023. [Online]. Available: https://www.ipcc.ch/report/ar6/wg3/chapter/chapter-2/

[73] Climate Change & Nature NZ, “Methane,” 2025. [Online]. Available: https://climateandnature.org.nz/climate-wiki/causes/greenhouse-gases/methane/

[74] NZ embracing renewable electricity future," *Beehive – The Official Website of the New Zealand Government*, Apr. 16, 2019. [Online]. Available: https://www.beehive.govt.nz/release/nz-embracing-renewable-electricity-future

[75] E. F. Viglizzo, F. E. Bert, M. A. Taboada, and B. J. R. Alves, “Editorial: Finding paths to net-zero carbon in climate-smart food systems,” *Front. Sustain. Food Syst.*, vol. 7, Nov. 2023, doi: 10.3389/fsufs.2023.1322803.

[76] E. M. Ungerfeld, “Opportunities and hurdles to the adoption and enhanced efficacy of feed additives towards pronounced mitigation of enteric methane emissions from ruminant livestock,” *Methane*, vol. 1, no. 4, Art. no. 4, Dec. 2022, doi: 10.3390/methane1040021.

[77] Ministry of Business, Innovation & Employment, “New Zealand Energy Quarterly – March 2023.” [Online]. Available: https://www.mbie.govt.nz/building-and-energy/energy-and-natural-resources/energy-statistics-and-modelling/energy-publications-and-technical-papers/new-zealand-energy-quarterly/previous-energy-quarterly-editions/new-zealand-energy-quarterly-march-2023

[78] International Energy Agency, “Evolution of wind power in New Zealand (January 2020 - October 2023).” [Online]. Available: https://www.iea.org/data-and-statistics/charts/evolution-of-wind-power-in-new-zealand-january-2020-october-2023

[79] Ministry of Business, Innovation & Employment, “Electricity statistics.” [Online]. Available: https://www.mbie.govt.nz/building-and-energy/energy-and-natural-resources/energy-statistics-and-modelling/energy-statistics/electricity-statistics

[80] Electricity Authority NZ, “Solar generation now and in the future.” [Online]. Available: http://www.ea.govt.nz/news/eye-on-electricity/solar-generation-now-and-in-the-future/

[81] Z. Zhang, X. Liu, D. Zhao, S. Post, and J. Chen, “Overview of the development and application of wind energy in New Zealand,” *Energy Built Environ.*, vol. 4, no. 6, pp. 725–742, Dec. 2023, doi: 10.1016/j.enbenv.2022.06.009.

[82] I. P. Pincelli, J. Hinkley, and A. Brent, "Developing onshore wind farms in Aotearoa New Zealand: carbon and energy footprints," *Journal of the Royal Society of New Zealand*, vol. 55, no. 4, pp. 1005–1027, 2024, doi: 10.1080/03036758.2024.2344785.

[83] R. Hale, D. Thompson, T. Brough, L. Kregting, M. Hayden, D. Parsons, S. D. Nodder, J. Beaumont, O. Anderson, and C. Stevens, "Environmental implications of future offshore renewable energy development in Aotearoa New Zealand," *Journal of the Royal Society of New Zealand*, vol. 55, no. 4, pp. 912–945, Oct. 2024, doi: 10.1080/03036758.2024.2406829.

[84] F. Sun and T. Jin, “A hybrid approach to multi-step, short-term wind speed forecasting using correlated features,” *Renew. Energy*, vol. 186, pp. 742–754, Mar. 2022, doi: 10.1016/j.renene.2022.01.041.

[85] A. C. Brent, J. Hinkley, D. Burmester, and R. Rayudu, “Solar Atlas of New Zealand from satellite imagery,” *J. R. Soc. N. Z.*, vol. 50, no. 4, pp. 572–583, Oct. 2020, doi: 10.1080/03036758.2020.1763409.

[86] My Solar Quotes, “Solar Power Potential in New Zealand.” [Online]. Available: https://www.mysolarquotes.co.nz/blog/how-solar-power-works/solar-power-potential-in-new-zealand/

[87] PV Magazine Australia, “Harmony secures consent for almost 150 MW of solar in NZ,” Sep. 2024. [Online]. Available: https://www.pv-magazine-australia.com/2024/09/10/harmony-secures-consent-for-almost-150-mw-of-solar-in-nz/

[88] I. P. Pincelli, A. C. Brent, J. Hinkley, and R. Sutherland, "Scaling up solar and wind electricity: empirical modelling and a disruptive scenario for their deployments in Aotearoa New Zealand," *Journal of the Royal Society of New Zealand*, vol. 55, no. 4, pp. 775–797, 2024, doi: 10.1080/03036758.2024.2381750.

[89] A. Tandon, A. Awasthi, K. C. Pattnayak, A. Tandon, T. Choudhury, and K. Kotecha, "Machine learning-driven solar irradiance prediction: advancing renewable energy in Rajasthan," *Discover Applied Sciences*, vol. 7, no. 2, p. 107, Jan. 2025, doi: 10.1007/s42452-025-06490-8.

[90] J. Lawless, B. Van Campen, and J. Randle, *Future Geothermal Generation Stack (New Zealand)*, Technical Report, Mar. 2020. [Online]. Available: https://doi.org/10.13140/RG.2.2.34460.92800

[91] Contact Energy, “New Zealand’s energy security boosted as Contact opens new geothermal power station,” Apr. 2025. [Online]. Available: https://contact.co.nz/investor-centre/news/2024/new-zealands-energy-security-boosted-as-contact-opens-new-geothermal-power-station

[92] A. Al-Fakih, A. Abdulraheem, and S. Kaka, “Application of machine learning and deep learning in geothermal resource development: Trends and perspectives,” *Deep Undergr. Sci. Eng.*, vol. 3, no. 3, pp. 286–301, 2024, doi: 10.1002/dug2.12098.

[93] L. C. A. Gutiérrez-Negrín, “Evolution of worldwide geothermal power 2020–2023,” *Geotherm. Energy*, vol. 12, no. 1, p. 14, May 2024, doi: 10.1186/s40517-024-00290-w.

[94] K. Phetpasak, G. Cant, and E. Challies, “Doing dams better? A case study of a proposed pumped hydro scheme at Lake Onslow, Central Otago,” *N. Z. Geogr.*, vol. 80, no. 3, pp. 220–232, 2024, doi: 10.1111/nzg.12404.

[95] A. Zisos, G.-K. Sakki, and A. Efstratiadis, “Mixing renewable energy with pumped hydropower storage: Design optimization under uncertainty and other challenges,” *Sustainability*, vol. 15, no. 18, Art. no. 13313, Jan. 2023, doi: 10.3390/su151813313.

[96] M. K. Majeed, “Evaluating the potential for a multi-use seasonal pumped storage scheme in New Zealand’s South Island,” Ph.D. dissertation, Univ. of Waikato, Hamilton, New Zealand, 2019. [Online]. Available: https://hdl.handle.net/10289/12423

[97] W. E. Bardsley, “Note on the pumped storage potential of the Onslow-Manorburn Depression, New Zealand,” *J. Hydrol. (N. Z.)*, vol. 44, no. 2, pp. 131–135, 2005. [Online]. Available: http://www.jstor.org/stable/43944921. Accessed: Jun. 20, 2025.

[98] Ch. V. Raghavendran, G. N. Satish, V. Krishna, and R. V. S. Lalitha, “Forecasting hourly electrical energy output of a power plant using parametric models,” in *Soft Computing and Signal Processing*, V. S. Reddy, V. K. Prasad, J. Wang, and K. T. V. Reddy, Eds., in *Advances in Intelligent Systems and Computing*, Singapore: Springer Nature, 2022, pp. 479–490, doi: 10.1007/978-981-16-7088-6_44.

[99] M. Bai, J. Liu, Y. Ma, X. Zhao, Z. Long, and D. Yu, “Long short-term memory network-based normal pattern group for fault detection of three-shaft marine gas turbine,” *Energies*, vol. 14, no. 1, Art. no. 13, Jan. 2021, doi: 10.3390/en14010013.

[100] J. Arriagada, M. Genrup, A. Loberg, and M. Assadi, “Fault diagnosis system for an industrial gas turbine by means of neural networks,” *Combustion*, Jan. 2003, pp. 1–6.

[101] M. S. Alsabban, N. Salem, and H. M. Malik, “Long short-term memory recurrent neural network (LSTM-RNN) power forecasting,” in *2021 13th IEEE PES Asia Pacific Power and Energy Engineering Conference (APPEEC)*, Nov. 2021, pp. 1–8, doi: 10.1109/APPEEC50844.2021.9687681.

[102] M. Yuvaraju and M. Divya, “Residential load forecasting by recurrent neural network on LSTM model,” in *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, May 2020, pp. 395–400, doi: 10.1109/ICICCS48265.2020.9121116.

[103] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, “Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network,” *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 841–851, Jan. 2019, doi: 10.1109/TSG.2017.2753802.

[104] K. Ijaz, Z. Hussain, J. Ahmad, S. F. Ali, M. Adnan, and I. Khosa, “A novel temporal feature selection based LSTM model for electrical short-term load forecasting,” *IEEE Access*, vol. 10, pp. 82596–82613, 2022, doi: 10.1109/ACCESS.2022.3196476.

[105] K. Pal and V. Tiwari, “Short term load forecasting for CPP using ANN,” in *Intelligent Renewable Energy Systems*, John Wiley & Sons, Ltd, 2022, pp. 391–408, doi: 10.1002/9781119786306.ch14.

[106] K. Gajowniczek and T. Ząbkowski, “Short term electricity forecasting using individual smart meter data,” *Procedia Comput. Sci.*, vol. 35, pp. 589–597, Jan. 2014, doi: 10.1016/j.procs.2014.08.140.

[107] J. Lei, C. Ren, W. Li, L. Fu, Z. Li, Z. Ni, Y. Li, C. Liu, H. Zhang, Z. Chen, and T. Yu, “Prediction of crucial nuclear power plant parameters using long short-term memory neural networks,” *Energy Reports*, early access, Apr. 2022, doi: 10.1002/er.7873.

[108] J. Zhang, P. Wang, R. Yan, and R. X. Gao, “Long short-term memory for machine remaining life prediction,” *J. Manuf. Syst.*, vol. 48, pp. 78–86, Jul. 2018, doi: 10.1016/j.jmsy.2018.05.011.

[109] Q. Cao, S. Chen, Y. Zheng, Y. Ding, Y. Tang, Q. Huang, K. Wang, and W. Xiang, "Classification and prediction of gas turbine gas path degradation based on deep neural networks," *International Journal of Energy Research*, vol. 45, no. 7, pp. 10513–10526, 2021, doi: 10.1002/er.6539.

[110] M. Talaat, M. H. Gobran, and M. Wasfi, “A hybrid model of an artificial neural network with thermodynamic model for system diagnosis of electrical power plant gas turbine,” *Eng. Appl. Artif. Intell.*, vol. 68, pp. 222–235, Feb. 2018, doi: 10.1016/j.engappai.2017.10.014.

[111] H. Nikpey, M. Assadi, and P. Breuhaus, “Development of an optimized artificial neural network model for combined heat and power micro gas turbines,” *Appl. Energy*, vol. 108, pp. 137–148, Aug. 2013, doi: 10.1016/j.apenergy.2013.03.016.

[112] K. T. Kabengele, L. K. Tartibu, and I. O. Olayode, “Modelling of a combined cycle power plant performance using artificial neural network model,” in *2022 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD)*, Durban, South Africa, 2022, pp. 1–7, doi: 10.1109/icABCD54961.2022.9856095.

[113] Y. D. Arferiandi, W. Caesarendra, and H. Nugraha, “Heat rate prediction of combined cycle power plant using an artificial neural network (ANN) method,” *Sensors*, vol. 21, no. 4, Art. no. 1022, Jan. 2021, doi: 10.3390/s21041022.

[114] M. H. K. Roni and M. A. G. Khan, “An artificial neural network based predictive approach for analyzing environmental impact on combined cycle power plant generation,” in *2017 2nd International Conference on Electrical & Electronic Engineering (ICEEE)*, Rajshahi, Bangladesh, 2017, pp. 1–4, doi: 10.1109/CEEE.2017.8412913.

[115] D. Adams, D.-H. Oh, D.-W. Kim, C.-H. Lee, and M. Oh, “Prediction of SOx–NOx emission from a coal-fired CFB power plant with machine learning: Plant data learned by deep neural network and least square support vector machine,” *J. Clean. Prod.*, vol. 270, p. 122310, Oct. 2020, doi: 10.1016/j.jclepro.2020.122310.

[116] S. Anvari, H. Taghavifar, R. K. Saray, S. Khalilarya, and S. Jafarmadar, “Implementation of ANN on CCHP system to predict trigeneration performance with consideration of various operative factors,” *Energy Convers. Manag.*, vol. 101, pp. 503–514, Sep. 2015, doi: 10.1016/j.enconman.2015.05.045.

[117] H. Asgari, X. Chen, M. B. Menhaj, and R. Sainudiin, “Artificial neural network–based system identification for a single-shaft gas turbine,” *J. Eng. Gas Turbines Power*, vol. 135, no. 9, Jul. 2013, doi: 10.1115/1.4024735.

[118] N. Pachauri and C. W. Ahn, “Electrical energy prediction of combined cycle power plant using gradient boosted generalized additive model,” *IEEE Access*, vol. 10, pp. 24566–24577, 2022, doi: 10.1109/ACCESS.2022.3153720.

[119] D. A. Wood, “Combined cycle gas turbine power output prediction and data mining with optimized data matching algorithm,” *SN Appl. Sci.*, vol. 2, no. 3, p. 441, Feb. 2020, doi: 10.1007/s42452-020-2249-7.

[120] S. Abbas, M. A. Khan, L. E. Falcon-Morales, A. Rehman, Y. Saeed, and M. Zareei, "Modeling, simulation and optimization of power plant energy sustainability for IoT enabled smart cities empowered with deep extreme learning machine," *IEEE Access*, vol. 8, pp. 39982–39997, 2020, doi: 10.1109/ACCESS.2020.2976452.

[121] E. A. Elfaki and A. H. Ahmed, “Prediction of electrical output power of combined cycle power plant using regression ANN model,” *J. Power Energy Eng.*, vol. 6, no. 12, Dec. 2018, doi: 10.4236/jpee.2018.612002.

[122] P. Tüfekci, “Prediction of full load electrical power output of a base load operated combined cycle power plant using machine learning methods,” *Int. J. Electr. Power Energy Syst.*, vol. 60, pp. 126–140, Sep. 2014, doi: 10.1016/j.ijepes.2014.02.027.

[123] The Weka Workbench, [Online]. Available: https://ml.cms.waikato.ac.nz/weka/. Accessed: Jun. 18, 2025.

[124] H. Kaya, P. Tüfekci, and S. F. Gürgen, “Local and global learning methods for predicting power of a combined gas & steam turbine,” in *Proc. Int. Conf. Emerging Trends in Computer and Electronics Engineering*, Dubai, UAE, Mar. 24–25, 2012, pp. 13–18.

[125] H. Moayedi and A. Mosavi, “Electrical power prediction through a combination of multilayer perceptron with water cycle ant lion and satin bowerbird searching optimizers,” *Sustainability*, vol. 13, no. 4, Art. no. 2336, Jan. 2021, doi: 10.3390/su13042336.

[126] R. Siddiqui, H. Anwar, F. Ullah, R. Ullah, M. A. Rehman, N. Jan, and F. Zaman, “Power prediction of combined cycle power plant (CCPP) using machine learning algorithm-based paradigm,” *International Journal of Distributed Sensor Networks*, vol. 17, no. 1, pp. 1–12, Dec. 2021, doi: 10.1155/2021/9966395.

[127] H. M. Ali, “Prediction of energy generated from composite cycle power plant in smart cities,” *Periodic Eng. Nat. Sci.*, vol. 9, no. 4, Art. no. 4, Sep. 2021, doi: 10.21533/pen.v9i4.2330.

[128] Z. Chen, F. Zou, L. Liao, S. Gao, M. Zhang, and J. Chun, "Prediction of electrical energy output for combined cycle power plant with different regression models," in *Proc. 5th Euro-China Conf. Intell. Data Anal. Appl.*, P. Krömer, Ed., Cham, Switzerland: Springer, 2019, pp. 558–566, doi: 10.1007/978-3-030-03766-6_63.

[129] A. A. Islikaye and A. Cetin, “Performance of ML methods in estimating net energy produced in a combined cycle power plant,” in *2018 6th Int. Istanbul Smart Grids Cities Congr. Fair (ICSG)*, Apr. 2018, pp. 217–220, doi: 10.1109/SGCF.2018.8408976.

[130] M. G. R. Shuvo, N. Sultana, L. Motin, and M. R. Islam, “Prediction of hourly total energy in combined cycle power plant using machine learning techniques,” in *2021 1st Int. Conf. Artif. Intell. Data Anal. (CAIDA)*, Apr. 2021, pp. 170–175, doi: 10.1109/CAIDA51941.2021.9425308.

[131] L. X. Niu and X. J. Liu, “Multivariable generalized predictive scheme for gas turbine control in combined cycle power plant,” in *2008 IEEE Conf. Cybern. Intell. Syst.*, Sep. 2008, pp. 791–796, doi: 10.1109/ICCIS.2008.4670947.

[132] A. Wankhede and V. Ghate, “Prediction of power of a combine gas and steam turbine using artificial neural network,” in *2018 3rd IEEE Int. Conf. Recent Trends Electron. Inf. Commun. Technol. (RTEICT)*, May 2018, pp. 1103–1108, doi: 10.1109/RTEICT42901.2018.9012143.

[133] G. Ahn and S. Hur, “Continuous conditional random field model for predicting the electrical load of a combined cycle power plant,” *Ind. Eng. Manag. Syst.*, vol. 15, no. 2, pp. 148–155, 2016, doi: 10.7232/iems.2016.15.2.148.

[134] C. M. Bartolini, F. Caresana, G. Comodi, L. Pelagalli, M. Renzi, and S. Vagni, "Application of artificial neural networks to micro gas turbines," *Energy Conversion and Management*, vol. 52, no. 1, pp. 781–788, Jan. 2011, doi: 10.1016/j.enconman.2010.08.003.

[135] Z. Liu and I. A. Karimi, “Gas turbine performance prediction via machine learning,” *Energy*, vol. 192, p. 116627, Feb. 2020, doi: 10.1016/j.energy.2019.116627.

[136] N. E. Benti, M. D. Chaka, and A. G. Semie, “Forecasting renewable energy generation with machine learning and deep learning: Current advances and future prospects,” *Sustainability*, vol. 15, no. 9, Art. no. 7087, Jan. 2023, doi: 10.3390/su15097087.

[137] Q. Paletta, G. Terrén-Serrano, Y. Nie, B. Li, J. Bieker, W. Zhang, L. Dubus, S. Dev, and C. Feng, "Advances in solar forecasting: Computer vision with deep learning," *Advanced Applied Energy*, vol. 11, Aug. 2023, Art. no. 100150, doi: 10.1016/j.adapen.2023.100150.

[138] C. N. Obiora, A. Ali, and A. N. Hasan, “Implementing extreme gradient boosting (XGBoost) algorithm in predicting solar irradiance,” in *2021 IEEE PES/IAS PowerAfrica*, Aug. 2021, pp. 1–5, doi: 10.1109/PowerAfrica52236.2021.9543159.

[139] P. Gupta and R. Singh, “Combining a deep learning model with multivariate empirical mode decomposition for hourly global horizontal irradiance forecasting,” *Renew. Energy*, vol. 206, Feb. 2023, doi: 10.1016/j.renene.2023.02.052.

[140] J. Pang and S. Dong, “A novel ensemble system for short-term wind speed forecasting based on hybrid decomposition approach and artificial intelligence models optimized by self-attention mechanism,” *Energy Convers. Manag.*, vol. 307, p. 118343, May 2024, doi: 10.1016/j.enconman.2024.118343.

[141] Y. Nie, Q. Paletta, A. Scott, L. M. Pomares, G. Arbod, S. Sgouridis, J. Lasenby, and A. Brandt, "Sky image-based solar forecasting using deep learning with heterogeneous multi-location data: Dataset fusion versus transfer learning," *Applied Energy*, vol. 369, p. 123467, Sep. 2024, doi: 10.1016/j.apenergy.2024.123467.

[142] Q. Xiang, L. Zi, X. Cong, and Y. Wang, “Concept drift adaptation methods under the deep learning framework: A literature review,” *Appl. Sci.*, vol. 13, no. 11, Art. no. 6515, Jan. 2023, doi: 10.3390/app13116515.

[143] C. Gu and H. Li, “Review on deep learning research and applications in wind and wave energy,” *Energies*, vol. 15, no. 4, Art. no. 415, Jan. 2022, doi: 10.3390/en15041510.

[144] J. H. Kim and W. B. Powell, “Optimal energy commitments with storage and intermittent supply,” *Oper. Res.*, vol. 59, no. 6, pp. 1347–1360, Dec. 2011, doi: 10.1287/opre.1110.0971.

[145] G. Xu and L. Xia, “Short-term prediction of wind power based on adaptive LSTM,” in *2018 2nd IEEE Conf. Energy Internet Energy Syst. Integr. (EI2)*, Oct. 2018, pp. 1–5, doi: 10.1109/EI2.2018.8582536.

[146] M. A. Prósper, C. Otero-Casal, F. Canoura Fernández, and G. Miguez-Macho, "Wind power forecasting for a real onshore wind farm on complex terrain using WRF high resolution simulations," *Renewable Energy*, vol. 135, pp. 674–686, May 2019, doi: 10.1016/j.renene.2018.12.047.

[147] J. F. Manwell, J. G. McGowan, and A. L. Rogers, *Wind Energy Explained: Theory, Design and Application*. John Wiley & Sons, 2010.

[148] J. Zhang, X. Jiang, X. Chen, X. Li, D. Guo, and L. Cui, "Wind power generation prediction based on LSTM," in *Proc. 2019 4th Int. Conf. Math. Artif. Intell. (ICMAI)*, New York, NY, USA, Apr. 2019, pp. 85–89, doi: 10.1145/3325730.3325735.

[149] W. Wu, K. Chen, Y. Qiao, and Z. Lu, “Probabilistic short-term wind power forecasting based on deep neural networks,” in *2016 Int. Conf. Probabilistic Methods Appl. Power Syst. (PMAPS)*, Oct. 2016, pp. 1–8, doi: 10.1109/PMAPS.2016.7764155.

[150] J. Zhang, D. Liu, Z. Li, X. Han, H. Liu, C. Dong, J. Wang, C. Liu, and Y. Xia, "Power prediction of a wind farm cluster based on spatiotemporal correlations," *Applied Energy*, vol. 302, p. 117568, Nov. 2021, doi: 10.1016/j.apenergy.2021.117568.

[151] J. Zhang, J. Yan, D. Infield, Y. Liu, and F. Lien, “Short-term forecasting and uncertainty analysis of wind turbine power based on long short-term memory network and Gaussian mixture model,” *Appl. Energy*, vol. 241, pp. 229–244, May 2019, doi: 10.1016/j.apenergy.2019.03.044.

[152] A. Gensler, J. Henze, B. Sick, and N. Raabe, “Deep learning for solar power forecasting — An approach using AutoEncoder and LSTM neural networks,” in *2016 IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2016, pp. 2858–2865, doi: 10.1109/SMC.2016.7844673.

[153] S. Lawrence and C. L. Giles, “Overfitting and neural networks: Conjugate gradient and backpropagation,” in *Proc. IEEE-INNS-ENNS Int. Joint Conf. Neural Networks (IJCNN)*, Jul. 2000, vol. 1, pp. 114–119, doi: 10.1109/IJCNN.2000.857823.

[154] T. Ahmad and D. Zhang, “A data-driven deep sequence-to-sequence long-short memory method along with a gated recurrent neural network for wind power forecasting,” *Energy*, vol. 239, p. 122109, Jan. 2022, doi: 10.1016/j.energy.2021.122109.

[155] Y. Chen, Y. Wang, Z. Dong, J. Su, Z. Han, D. Zhou, Y. Zhao, and Y. Bao, “2-D regional short-term wind speed forecast based on CNN-LSTM deep learning model,” *Energy Conversion and Management*, vol. 244, p. 114451, Sep. 2021, doi: 10.1016/j.enconman.2021.114451.

[156] L. Lin, M. Li, L. Ma, A. Baziar, and Z. M. Ali, “Hybrid RNN-LSTM deep learning model applied to a fuzzy based wind turbine data uncertainty quantization method,” *Ad Hoc Netw.*, vol. 123, p. 102658, Dec. 2021, doi: 10.1016/j.adhoc.2021.102658.

[157] W. Dong, H. Sun, J. Tan, Z. Li, J. Zhang, and H. Yang, “Multi-degree-of-freedom high-efficiency wind power generation system and its optimal regulation based on short-term wind forecasting,” *Energy Convers. Manag.*, vol. 249, p. 114829, Dec. 2021, doi: 10.1016/j.enconman.2021.114829.

[158] M. Zhou, B. Wang, S. Guo, and J. Watada, “Multi-objective prediction intervals for wind power forecast based on deep neural networks,” *Inf. Sci.*, vol. 550, pp. 207–220, Mar. 2021, doi: 10.1016/j.ins.2020.10.034.

[159] X. Yan, Y. Liu, Y. Xu, and M. Jia, “Multistep forecasting for diurnal wind speed based on hybrid deep learning model with improved singular spectrum decomposition,” *Energy Convers. Manag.*, vol. 225, p. 113456, Dec. 2020, doi: 10.1016/j.enconman.2020.113456.

[160] Y.-L. Hu and L. Chen, “A nonlinear hybrid wind speed forecasting model using LSTM network, hysteretic ELM and Differential Evolution algorithm,” *Energy Convers. Manag.*, vol. 173, pp. 123–142, Oct. 2018, doi: 10.1016/j.enconman.2018.07.070.

[161] W. Chen, H. Liu, Y. Lin, W. Li, Y. Sun, and D. Zhang, “LSTM-NN yaw control of wind turbines based on upstream wind information,” *Energies*, vol. 13, no. 6, Art. no. 1482, Jan. 2020, doi: 10.3390/en13061482.

[162] S. A. Pourmousavi Kani and M. M. Ardehali, “Very short-term wind speed prediction: A new artificial neural network–Markov chain model,” *Energy Convers. Manag.*, vol. 52, no. 1, pp. 738–745, Jan. 2011, doi: 10.1016/j.enconman.2010.07.053.

[163] Z. Dzulfikri, N. Nuryanti, and Y. Erdani, “Design and implementation of artificial neural networks to predict wind directions on controlling yaw of wind turbine prototype,” *J. Robot. Control (JRC)*, vol. 1, no. 1, Art. no. 1, Jan. 2020, doi: 10.18196/jrc.1105.

[164] I. Delgado and M. Fahim, “Wind turbine data analysis and LSTM-based prediction in SCADA system,” *Energies*, vol. 14, no. 1, Art. no. 125, Jan. 2021, doi: 10.3390/en14010125.

[165] J. Zhang, X. Jiang, X. Chen, X. Li, D. Guo, and L. Cui, "Wind power generation prediction based on LSTM," in *Proc. 4th Int. Conf. Math. Artif. Intell. (ICMAI)*, Apr. 2019, pp. 85–89. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/3325730.3325735. Accessed: Jan. 4, 2023.

[166] J. Wang, S. Qin, Q. Zhou, and H. Jiang, “Medium-term wind speeds forecasting utilizing hybrid models for three different sites in Xinjiang, China,” *Renew. Energy*, vol. 76, pp. 91–101, Apr. 2015, doi: 10.1016/j.renene.2014.11.011.

[167] I. A. Araya, C. Valle, and H. Allende, “LSTM-based multi-scale model for wind speed forecasting,” in *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications (CIARP 2018)*, R. Vera-Rodriguez, A. F. Frangi, J. Ortega-Garcia, and V. Ruiz-Rodarte, Eds., Cham, Switzerland: Springer, 2019, pp. 38–45, doi: 10.1007/978-3-030-13469-3_5.

[168] Y. Cao and L. Gui, “Multi-step wind power forecasting model using LSTM networks, similar time series and LightGBM,” in *2018 5th Int. Conf. Syst. Informatics (ICSAI)*, Nov. 2018, pp. 192–197, doi: 10.1109/ICSAI.2018.8599498.

[169] K. Lu, W. X. Sun, X. Wang, X. R. Meng, Y. Zhai, H. H. Li, and R. G. Zhang, “Short-term wind power prediction model based on encoder–decoder LSTM,” *IOP Conf. Ser. Earth Environ. Sci.*, vol. 186, no. 5, p. 012020, Sep. 2018, doi: 10.1088/1755-1315/186/5/012020.

[170] V. D. Casas, F. L. Pena, R. J. Duro, and A. Lamas, “Automatic aerodynamic design of a wind turbine through evolutionary techniques,” in *2005 IEEE Intell. Data Acquis. Adv. Comput. Syst.: Technol. Appl.*, Sep. 2005, pp. 454–459, doi: 10.1109/IDAACS.2005.283023.

[171] M. Konstantinou, S. Peratikou, and A. G. Charalambides, “Solar photovoltaic forecasting of power output using LSTM networks,” *Atmosphere*, vol. 12, no. 1, Art. no. 124, Jan. 2021, doi: 10.3390/atmos12010124.

[172] Y. Jung, J. Jung, B. Kim, and S. Han, “Long short-term memory recurrent neural network for modeling temporal patterns in long-term power forecasting for solar PV facilities: Case study of South Korea,” *J. Clean. Prod.*, vol. 250, p. 119476, Mar. 2020, doi: 10.1016/j.jclepro.2019.119476.

[173] H. M. A. Rayess and A. Ü. Keskin, “Forecasting the hydroelectric power generation of GCMs using machine learning techniques and deep learning (Almus Dam, Turkey),” *GEOFIZIKA*, vol. 38, no. 1, Art. no. 1, Jun. 2021, doi: 10.15233/gfz.2021.38.4.

[174] E. Rossi, I. Pecorini, and R. Iannelli, “Multilinear regression model for biogas production prediction from dry anaerobic digestion of OFMSW,” *Sustainability*, vol. 14, no. 8, Art. no. 4393, Jan. 2022, doi: 10.3390/su14084393.

[175] D. Yongsheng, J. Fengshun, Z. Jie, and L. Zhikeng, “A short-term power output forecasting model based on correlation analysis and ELM-LSTM for distributed PV system,” *J. Electr. Comput. Eng.*, vol. 2020, p. e2051232, Jun. 2020, doi: 10.1155/2020/2051232.

[176] F. Wang, Z. Xuan, Z. Zhen, K. Li, T. Wang, and M. Shi, “A day-ahead PV power forecasting method based on LSTM-RNN model and time correlation modification under partial daily pattern prediction framework,” *Energy Convers. Manag.*, vol. 212, p. 112766, May 2020, doi: 10.1016/j.enconman.2020.112766.

[177] B. Chen, P. Lin, Y. Lai, S. Cheng, Z. Chen, and L. Wu, “Very-short-term power prediction for PV power plants using a simple and effective RCC-LSTM model based on short term multivariate historical datasets,” Electronics, vol. 9, no. 2, Art. no. 289, Feb. 2020, doi: 10.3390/electronics9020289.

[178] M. Gao, J. Li, F. Hong, and D. Long, “Short-term forecasting of power production in a large-scale photovoltaic plant based on LSTM,” *Appl. Sci.*, vol. 9, no. 15, Art. no. 3192, Jan. 2019, doi: 10.3390/app9153192.

[179] F. Mei, J. Gu, J. Lu, J. Lu, J. Zhang, and Y. Jiang, “Day-ahead nonparametric probabilistic forecasting of photovoltaic power generation based on the LSTM-QRA ensemble model,” *IEEE Access*, vol. 8, pp. 166138–166149, 2020, doi: 10.1109/ACCESS.2020.3021581.

[180] R. Abhinav and N. M. Pindoriya, "Grid integration of wind turbine and battery energy storage system: Review and key challenges," in *2016 IEEE 6th Int. Conf. Power Syst. (ICPS)*, Mar. 2016, pp. 1–6, doi: 10.1109/ICPES.2016.7583998.

[181] H. Chen, T. N. Cong, W. Yang, C. Tan, Y. Li, and Y. Ding, "Progress in electrical energy storage system: A critical review," *Prog. Nat. Sci.*, vol. 19, no. 3, pp. 291–312, Mar. 2009, doi: 10.1016/j.pnsc.2008.07.014.

[182] S. Belaid, D. Rekioua, A. Oubelaid, D. Ziane, and T. Rekioua, "A power management control and optimization of a wind turbine with battery storage system," *J. Energy Storage*, vol. 45, p. 103613, Jan. 2022, doi: 10.1016/j.est.2021.103613.

[183] M. A. Hannan, S. B. Wali, P. J. Ker, M. S. Abd Rahman, M. Mansor, V. K. Ramachandaramurthy, K. M. Muttaqi, T. M. I. Mahlia, and Z. Y. Dong, "Battery energy-storage system: A review of technologies, optimization objectives, constraints, approaches, and outstanding issues," *J. Energy Storage*, vol. 42, p. 103023, Oct. 2021, doi: 10.1016/j.est.2021.103023.

[184] Y. Zheng, J. Zhao, Y. Song, F. Luo, K. Meng, and J. Qiu, "Optimal operation of battery energy storage system considering distribution system uncertainty," *IEEE Trans. Sustain. Energy*, vol. 9, no. 3, pp. 1051–1060, Jul. 2018, doi: 10.1109/TSTE.2017.2762364.

[185] Z. Yang, K. Li, and A. Foley, "Computational scheduling methods for integrating plug-in electric vehicles with power systems: A review," *Renew. Sustain. Energy Rev.*, vol. 51, pp. 396–416, Nov. 2015, doi: 10.1016/j.rser.2015.06.007.

[186] J. García-Villalobos, I. Zamora, J. I. San Martín, F. J. Asensio, and V. Aperribay, "Plug-in electric vehicles in electric distribution networks: A review of smart charging approaches," *Renew. Sustain. Energy Rev.*, vol. 38, pp. 717–731, Oct. 2014, doi: 10.1016/j.rser.2014.07.040.

[187] W. Tang and H.-T. Yang, "Optimal Operation and Bidding Strategy of a Virtual Power Plant Integrated With Energy Storage Systems and Elasticity Demand Response," *IEEE Access*, vol. 7, pp. 79798–79809, 2019, doi: 10.1109/ACCESS.2019.2922700.

[188] M. R. Jannesar, A. Sedighi, M. Savaghebi, and J. M. Guerrero, "Optimal placement, sizing, and daily charge/discharge of battery energy storage in low voltage distribution network with high photovoltaic penetration," *Appl. Energy*, vol. 226, pp. 957–966, Sep. 2018, doi: 10.1016/j.apenergy.2018.06.036.

[189] C. Wadia, P. Albertus, and V. Srinivasan, "Resource constraints on the battery energy storage potential for grid and transportation applications," *J. Power Sources*, vol. 196, no. 3, pp. 1593–1598, Feb. 2011, doi: 10.1016/j.jpowsour.2010.08.056.

[190] T. Mesbahi, F. Khenfri, N. Rizoug, P. Bartholomeüs, and P. L. Moigne, "Combined Optimal Sizing and Control of Li-Ion Battery/Supercapacitor Embedded Power Supply Using Hybrid Particle Swarm–Nelder–Mead Algorithm," *IEEE Trans. Sustain. Energy*, vol. 8, no. 1, pp. 59–73, Jan. 2017, doi: 10.1109/TSTE.2016.2582927.

[191] Y. Yang, S. Bremner, C. Menictas, and M. Kay, "Battery energy storage system size determination in renewable energy systems: A review," *Renew. Sustain. Energy Rev.*, vol. 91, pp. 109–125, Aug. 2018, doi: 10.1016/j.rser.2018.03.047.

[192] F. Díaz-González, A. Sumper, O. Gomis-Bellmunt, and R. Villafáfila-Robles, "A review of energy storage technologies for wind power applications," *Renew. Sustain. Energy Rev.*, vol. 16, no. 4, pp. 2154–2171, May 2012, doi: 10.1016/j.rser.2012.01.029.

[193] H. Zhao, Q. Wu, S. Hu, H. Xu, and C. N. Rasmussen, “Review of energy storage system for wind power integration support,” *Appl. Energy*, vol. 137, pp. 545–553, Jan. 2015, doi: 10.1016/j.apenergy.2014.04.103.

[194] B. R. Barricelli, E. Casiraghi, and D. Fogli, “A survey on digital twin: Definitions, characteristics, applications, and design implications,” *IEEE Access*, vol. 7, pp. 167653–167671, 2019, doi: 10.1109/ACCESS.2019.2953499.

[195] M. Gopal, T. C. Prakash, N. V. Ramakrishna, and B. P. Yadav, “IoT Based Solar Power Monitoring System,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 981, no. 3, p. 032037, Dec. 2020, doi: 10.1088/1757-899X/981/3/032037.

[196] A. A. Felix, O. Hope, O. Joseph, and N. Chukwuebuka, “Design and Implementation of a Solar Energy Measurement and Monitoring System,” in *2018 IEEE PES/IAS PowerAfrica*, pp. 833–840, Jun. 2018, doi: 10.1109/PowerAfrica.2018.8521048.

[197] J. K., R. S. M., S. S. V., and K. K., “Solar power remote monitoring and controlling using IoT,” in *2023 2nd International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)*, Coimbatore, India, 2023, pp. 1–5, doi: 10.1109/ICAECA56562.2023.10200775.

[198] A. Kekre and S. K. Gawre, “Solar photovoltaic remote monitoring system using IOT,” in *2017 International Conference on Recent Innovations in Signal processing and Embedded Systems (RISE)*, pp. 619–623, Oct. 2017, doi: 10.1109/RISE.2017.8378227.

[199] B. Shrihariprasath and V. Rathinasabapathy, “A smart IoT system for monitoring solar PV power conditioning unit,” in *2016 World Conference on Futuristic Trends in Research and Innovation for Social Welfare (Startup Conclave)*, Coimbatore, India, 2016, pp. 1–5, doi: 10.1109/STARTUP.2016.7583930.

[200] D. Saravanan and T. Lingeshwaran, “Monitoring Of Solar Panel Based On IOT,” in *2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN)*, pp. 1–5, Mar. 2019, doi: 10.1109/ICSCAN.2019.8878814.

[201] K. N. Kumar, P. Badrinath, S. Vickraman, and G. Satheesan, “Charging Station for E-Vehicle Using Solar with IoT,” in *2022 Fourth International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT)*, pp. 1–5, Dec. 2022, doi: 10.1109/ICERECT56837.2022.10060451.

[202] H. Eteruddin, D. Setiawan, and Atmam, “Web Based Raspberry Monitoring System Solar Energy Power Plant,” *IOP Conf. Ser. Earth Environ. Sci.*, vol. 469, no. 1, p. 012051, Apr. 2020, doi: 10.1088/1755-1315/469/1/012051.

[203] “IoT Analytics - ThingSpeak Internet of Things.” *Accessed: Nov. 19, 2024*. [Online]. Available: https://thingspeak.mathworks.com/

[204] S. Ayesh, P. Ramesh, and S. Ramakrishnan, “Design of wireless sensor network for monitoring the performance of photovoltaic panel,” in *2017 Trends in Industrial Measurement and Automation (TIMA)*, Chennai, India: IEEE, pp. 1–6, Jan. 2017, doi: 10.1109/TIMA.2017.8064826.

[205] A. Haider-e-Karar, A. A. Khuwaja, and A. Sattar, "Solar power remote monitoring and controlling using Arduino, LabVIEW and web browser," in *Proc. Power Gener. Syst. Renewable Energy Technol. (PGSRET)*, Islamabad, Pakistan, 2015, pp. 1–4, doi: 10.1109/PGSRET.2015.7312240.

[206] J. Greulich, M. Glatthaar, and S. Rein, “Fill factor analysis of solar cells’ current–voltage curves,” *Prog. Photovolt. Res. Appl.*, vol. 18, no. 7, pp. 511–515, 2010, doi: 10.1002/pip.979.

[207] K. Kushiya, Y. Tanaka, H. Hakuma, Y. Goushi, S. Kijima, T. Aramoto, and Y. Fujiwara, “Interface control to enhance the fill factor over 0.70 in a large-area CIS-based thin-film PV technology,” *Thin Solid Films*, vol. 517, no. 7, pp. 2108–2110, Feb. 2009, doi: 10.1016/j.tsf.2008.10.125.

[208] C.-H. Chiang and C.-G. Wu, “Bulk heterojunction perovskite–PCBM solar cells with high fill factor,” *Nat. Photonics*, vol. 10, no. 3, pp. 196–200, Mar. 2016, doi: 10.1038/nphoton.2016.3.

[209] A. R. Mollick, “Why it’s crucial to understand the fill factor of solar cell? How it’s affecting the efficiency of PV cells? – EcoWowLife.” *Accessed: Nov. 25, 2024*. [Online]. Available: https://ecowowlife.com/fill-factor-of-solar-cell/

[210] S. Lanfredi, G. Palacio, F. S. Bellucci, C. V. Colin, and M. A. L. Nobre, “Thermistor behaviour and electric conduction analysis of Ni-doped niobate ferroelectric: the role of multiple β parameters,” *J. Phys. Appl. Phys.*, vol. 45, no. 43, p. 435302, Oct. 2012, doi: 10.1088/0022-3727/45/43/435302.

[211] K. C. Narasimhamurthy, B. Malaviya, K. R. Pranesh, A. Jayashree, and H. Kamatham, “Analysis of Operational-Amplifier Inverting and Non-inverting Amplifiers in Remote Lab,” in *Cyber-physical Systems and Digital Twins*, M. E. Auer and K. Ram B., Eds., Cham: Springer International Publishing, 2020, pp. 595–604. doi: 10.1007/978-3-030-23162-0_53.

[212] S. Pasha, “Thingspeak Based Sensing and Monitoring System for IoT with Matlab Analysis,” *Int. J. New Technol. Res.*, vol. 2, no. 6, p. 263492, Jun. 2016.

[213] A. F. Agarap, “Deep learning using rectified linear units (ReLU),” *arXiv*, Mar. 2018, doi: 10.48550/arXiv.1803.08375.

[214] “Adam – Cornell University Computational Optimization Open Textbook – Optimization Wiki,” *Accessed: Jun. 19, 2023*. [Online]. Available: https://optimization.cbe.cornell.edu/index.php?title=Adam

[215] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," J. Mach. Learn. Res., vol. 3, pp. 1157–1182, Mar. 2003.

[216] R. Kohavi and G. H. John, "Wrappers for feature subset selection," Artif. Intell., vol. 97, no. 1–2, pp. 273–324, Jan. 1997.

[217] “Welcome to CoolProp — CoolProp 6.4.3 Documentation,” *Accessed: Jun. 7, 2023*. [Online]. Available: http://www.coolprop.org/

[218] “GEKKO Optimization Suite — GEKKO 1.0.5 Documentation,” *Accessed: Jun. 19, 2023*. [Online]. Available: https://gekko.readthedocs.io/en/latest/

[219] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," Int. J. Forecast., vol. 22, no. 4, pp. 679–688, Oct.–Dec. 2006.

[220] S. R. Tito, M. Walmsley, and M. Apperley, “Impact of thermal energy storage on optimal sizing of a grid-connected wind-photovoltaic-battery system for an industry-edged community,” in *Proc. 2023 IEEE 5th Int. Conf. DC Microgrids (ICDCM)*, Nov. 2023, pp. 1–6, doi: 10.1109/ICDCM54452.2023.10433591.

[221] “Greenhouse gas emissions targets and reporting,” *Ministry for the Environment*, *Accessed: Sep. 8, 2024*. [Online]. Available: https://environment.govt.nz/…/greenhouse-gas-emissions-targets-and-reporting/

[222] “Emissions Reduction Plan,” *Ministry of Business, Innovation & Employment*, *Accessed: Sep. 8, 2024*. [Online]. Available: https://www.mbie.govt.nz/.../emissions-reduction-plan

[223] D. Maradin, “Advantages and disadvantages of renewable energy sources utilization,” *International Journal of Energy Economics and Policy*, vol. 11, no. 3, pp. 176–183, 2021.

[224] L. Suganthi, S. Iniyan, and A. A. Samuel, “Applications of fuzzy logic in renewable energy systems – A review,” *Renew. Sustain. Energy Rev.*, vol. 48, pp. 585–607, Aug. 2015, doi: 10.1016/j.rser.2015.04.037.

[225] S.-C. Necula, “Assessing the potential of artificial intelligence in advancing clean energy technologies in Europe: A systematic review,” *Energies*, vol. 16, no. 22, Art. no. 27633, Jan. 2023, doi: 10.3390/en16227633.

[226] “Solarview,” *Accessed: Apr. 23, 2024*. [Online]. Available: https://solarview.niwa.co.nz/

[227] Z. Qadir, S. I. Khan, E. Khalaji, H. S. Munawar, F. Al-Turjman, M. A. P. Mahmud, A. Z. Kouzani, and K. Le, “Predicting the energy output of hybrid PV–wind renewable energy system using feature selection technique for smart grids,” *Energy Rep.*, vol. 7, pp. 8465–8475, Nov. 2021, doi: 10.1016/j.egyr.2021.01.018.

[228] G. Notton, C. Voyant, A. Fouilloy, J. L. Duchaud, and M. L. Nivet, “Some applications of ANN to solar radiation estimation and forecasting for energy applications,” *Appl. Sci.*, vol. 9, no. 1, Art. no. 1, Jan. 2019, doi: 10.3390/app9010209.

[229] *Solar Collectors and Panels: Theory and Applications*, Croatia: IntechOpen, 2010. [Online]. Available: https://www.google.co.nz/books/edition/Solar_Collectors_and_Panels/9sCODwAAQBAJ?hl=en&gbpv=0

[230] A. Mellit and A. M. Pavan, “A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy,” *Sol. Energy*, vol. 84, no. 5, Art. no. 5, May 2010, doi: 10.1016/j.solener.2010.02.006.

[231] N. Yildiz, “Layered feedforward neural network is relevant to empirical physical formula construction: A theoretical analysis and some simulation results,” *Phys. Lett. A*, vol. 345, no. 1, Art. no. 1, Sep. 2005, doi: 10.1016/j.physleta.2005.06.116.

[232] J. Bergstra and Y. Bengio, “Random search for hyper-parameter optimization,” *J. Mach. Learn. Res.*, vol. 13, no. 10, Art. no. 10, 2012.

[233] A. R. Abbasi et al., "Recent developments of energy management strategies in microgrids: An updated and comprehensive review and classification," Energy Convers. Manag., vol. 297, p. 117723, Mar. 2023, doi: 10.1016/j.enconman.2023.117723.

[234] K. S. Joshal, N. Pal, and S. Srivastava, "Microgrids with model predictive control: A critical review," Energies, vol. 16, no. 13, p. 4851, Jul. 2023, doi: 10.3390/en16134851.

[235] J. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy, and A. Bouchachia, “A survey on concept drift adaptation,” *ACM Comput. Surv.*, vol. 46, no. 4, Art. no. 44:1-44:37, Mar. 2014, doi: 10.1145/2523813.

[236] T. Lan, K. Jermsittiparsert, S. T. Alrashood, M. Rezaei, L. Al-Ghussain, and M. A. Mohamed, “An advanced machine learning based energy management of renewable microgrids considering hybrid electric vehicles’ charging demand,” *Energies*, vol. 14, no. 3, Art. no. 30569, Jan. 2021, doi: 10.3390/en14030569.

[237] L. Lei, Y. Tan, G. Dahlenburg, W. Xiang, and K. Zheng, “Dynamic energy dispatch based on deep reinforcement learning in IoT-driven smart isolated microgrids,” *IEEE Internet Things J.*, vol. 8, no. 10, pp. 7938–7953, May 2021, doi: 10.1109/JIOT.2020.3042007.

[238] A. Joshi, S. Capezza, A. Alhaji, and M.-Y. Chow, "Survey on AI and machine learning techniques for microgrid energy management systems," IEEE/CAA J. Autom. Sinica, vol. 10, no. 7, pp. 1513–1529, Jul. 2023, doi: 10.1109/JAS.2023.123657.

[239] Y. Gupta and M. Amin, “A neural network-based energy management system for PV-battery based microgrids,” *arXiv*, Jun. 14, 2022, doi: 10.48550/arXiv.2206.06716.

[240] S. Mohseni, R. Khalid, and A. C. Brent, “Data-driven, metaheuristic-based off-grid microgrid capacity planning optimisation and scenario analysis: Insights from a case study of Aotea-Great Barrier Island,” *arXiv*, Sep. 21, 2022, doi: 10.48550/arXiv.2209.10668.

[241] P. Mohammadi, R. Darshi, S. Shamaghdari, and P. Siano, “Comparative analysis of control strategies for microgrid energy management with a focus on reinforcement learning,” *IEEE Access*, vol. 12, pp. 171368–171395, 2024, doi: 10.1109/ACCESS.2024.3495032.

[242] W. Guo, Y. Li, X. Zhou, C. Xu, and Y. Zhang, “A deep learning-based microgrid energy management method under the Internet of Things architecture,” *Int. J. Grid Cloud Comput. Manag. Syst. (IJGCMS)*, vol. 16, no. 1, pp. 1–19, 2024, doi: 10.4018/IJGCMS.336288.

[243] M. Apperley, H. Viggers, M. Walmsley, R. Chapman, P. Howden-Chapman, G. Penny, I. Shearer, and P. Taptiklis, "The role of smart community microgrids in Aotearoa’s energy future," *J. R. Soc. N. Z.*, vol. 55, pp. 1147–1166, 2025.

[244] A. Parkin, M. Herrera, and D. A. Coley, “Net-zero buildings: When carbon and energy metrics diverge,” *Build. Cities*, vol. 1, no. 1, May 2020, doi: 10.5334/bc.27.

[245] N. Shamarova, K. Suslov, P. Ilyushin, and I. Shushpanov, “Review of battery energy storage systems modeling in microgrids with renewables considering battery degradation,” *Energies*, vol. 15, no. 19, Art. no. 6967, Jan. 2022, doi: 10.3390/en15196967.

[246] Y. Wan, J. Qin, X. Yu, T. Yang, and Y. Kang, “Price-based residential demand response management in smart grids: A reinforcement learning-based approach,” *IEEECAA J. Autom. Sin.*, vol. 9, no. 1, pp. 123–134, Jan. 2022, doi: 10.1109/JAS.2021.1004287.

[247] N. Rahbari-Asr, U. Ojha, Z. Zhang, and M.-Y. Chow, “Incremental welfare consensus algorithm for cooperative distributed generation/demand response in smart grid,” *IEEE Trans. Smart Grid*, vol. 5, no. 6, pp. 2836–2845, Aug. 2014, doi: 10.1109/TSG.2014.2346511.

[248] M. A. Velasquez, J. Barreiro-Gomez, N. Quijano, A. I. Cadena, and M. Shahidehpour, “Intra-hour microgrid economic dispatch based on model predictive control,” *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 1968–1979, May 2020, doi: 10.1109/TSG.2019.2945692.

[249] M. H. Alabdullah and M. A. Abido, “Microgrid energy management using deep Q-network reinforcement learning,” *Alex. Eng. J.*, vol. 61, no. 11, pp. 9069–9078, Nov. 2022, doi: 10.1016/j.aej.2022.02.042.

[250] N. Rahbari-Asr, Y. Zhang, and M.-Y. Chow, "Consensus-based distributed scheduling for cooperative operation of distributed energy resources and storage devices in smart grids," *IET Gener. Transm. Distrib.*, vol. 10, no. 5, pp. 1268–1277, Apr. 2016, doi: 10.1049/iet-gtd.2015.0159.

[251] H. M. Gomes, A. Bifet, J. Read, J. P. Barddal, F. Enembreck, B. Pfharinger, G. Holmes, and T. Abdessalem, "Adaptive random forests for evolving data stream classification," *Mach. Learn.*, vol. 106, no. 9, pp. 1469–1495, Oct. 2017, doi: 10.1007/s10994-017-5642-8.

[252] W. Dong, Q. Yang, W. Li, and A. Y. Zomaya, "Machine-learning-based real-time economic dispatch in islanding microgrids in a cloud-edge computing environment," *IEEE Internet Things J.*, vol. 8, no. 17, pp. 13703–13711, Sep. 2021, doi: 10.1109/JIOT.2021.3067951.

[253] S. Matrone, A. Pozzi, E. Ogliari, and S. Leva, "Deep learning-based predictive control for optimal battery management in microgrids," *IEEE Access*, vol. 12, pp. 141580–141593, Sep. 2024, doi: 10.1109/ACCESS.2024.3458435.

[254] G. Gao, Y. Wen, and D. Tao, "Distributed energy trading and scheduling among microgrids via multiagent reinforcement learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 12, pp. 10638–10652, Dec. 2023, doi: 10.1109/TNNLS.2022.3170070.

[255] M. Apperley, "Modelling fractal-structured smart microgrids: Exploring signals and protocols," presented at *Energy 2019: 9th Int. Conf. Smart Grids, Green Commun. & IT Energy-aware Technol.*, Athens, Greece, Jun. 2019, pp. 13–17. *Accessed: Sep. 12, 2022*. [Online]. Available: https://researchcommons.waikato.ac.nz/handle/10289/13037

[256] T. Mindra, "Fractal smart grids – A new challenge for power systems operation," in *2023 24th Int. Conf. Control Syst. & Comput. Sci. (CSCS)*, May 2023, pp. 301–306, doi: 10.1109/CSCS59211.2023.00054.

[257] M. Apperley and T. Toki, "An islanded community solar microgrid with capability of future fractal growth," presented at *Energy 2023: 13th Int. Conf. Smart Grids, Green Commun. & IT Energy-aware Technol.*, Barcelona, Spain, Mar. 2023, pp. 29–35. *Accessed: Mar. 11, 2025*. [Online]. Available: https://www.publichousingresearch.org.nz/resources/islanded-community-solar-microgrid-capability-future-fractal-growth

[258] G. Hulten, L. Spencer, and P. Domingos, "Mining time-changing data streams," in *Proc. 7th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. (KDD '01)*, New York, NY, USA, Aug. 2001, pp. 97–106, doi: 10.1145/502512.502529.

[259] N. Altman and M. Krzywinski, "Ensemble methods: Bagging and random forests," *Nat. Methods*, vol. 14, no. 10, pp. 933–934, Oct. 2017, doi: 10.1038/nmeth.4438.

[260] U. Adhikari, T. H. Morris, and S. Pan, "Applying Hoeffding adaptive trees for real-time cyber-power event and intrusion classification," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4049–4060, Sep. 2018, doi: 10.1109/TSG.2017.2647778.

[261] B. Hosseini, "Forecasting the monthly electricity consumption of household subscribers using the Similar Pattern Algorithm," *Social Science Research Network*, Rochester, NY, Apr. 22, 2024, Work No. 4803434, doi: 10.2139/ssrn.4803434.

[262] T. G. Walmsley, P. Patros, W. Yu, B. R. Young, S. Burroughs, M. Apperley, J. K. Carson, I. A. Udugama, H. Aeowjaroenlap, M. J. Atkins, and M. R. W. Walmsley, "Adaptive digital twins for energy-intensive industries and their local communities," *Digit. Chem. Eng.*, vol. 10, p. 100139, Mar. 2024, doi: 10.1016/j.dche.2024.100139.

[263] A. Monforti Ferrario, A. Bartolini, G. Comodi, F. J. Vivas Fernández, S. Ciapetti, M. Morini, and M. Renzi, “Optimal sizing of battery and hydrogen energy storage systems configurations in a hybrid renewable microgrid,” *E3S Web Conf.*, vol. 238, p. 09002, Jan. 2021, doi: 10.1051/e3sconf/202123809002.

[264] A. C. Duman, H. S. Erden, Ö. Gönül, and Ö. Güler, “Optimal sizing of PV-BESS units for home energy management system-equipped households considering day-ahead load scheduling for demand response and self-consumption,” *Energy Build.*, vol. 267, p. 112164, Jul. 2022, doi: 10.1016/j.enbuild.2022.112164.

[265] F. Klausmann and A.-L. Klingler, “Adaptive control strategy for stationary electric battery storage systems with reliable peak load limitation at maximum self-consumption of locally generated energy,” *Energies*, vol. 16, no. 9, Art. no. 9, Jan. 2023, doi: 10.3390/en16093964.

[266] A. Geslin, L. Xu, D. Ganapathi, K. Moy, W. C. Chueh, and S. Onori, “Dynamic cycling enhances battery lifetime,” *Nature Energy*, vol. 10, pp. 172–180, Feb. 2025, doi: 10.1038/s41560-024-01675-8.

[267] X. Jiao, L. Rao, J. Yap, C.-Y. Yu, and J.-H. Kim, “Stabilizing cathode-electrolyte interphase of $LiNi_{0.5}Mn_{1.5}O_4$ high-voltage spinel by blending garnet solid electrolyte in lithium-ion batteries,” *J. Power Sources*, vol. 561, p. 232748, Mar. 2023, doi: 10.1016/j.jpowsour.2023.232748.

[268] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE," Geosci. Model Dev., vol. 7, no. 3, pp. 1247–1250, May 2014, doi: 10.5194/gmd-7-1247-2014.

[269] J. Haug, K. Broelemann, and G. Kasneci, "Dynamic model tree for interpretable data stream learning," arXiv preprint arXiv:2203.16181, Mar. 2022. [Online]. Available: https://arxiv.org/abs/2203.16181.

[270] J. Mair, K. Suomalainen, D. M. Eyers, and M. W. Jack, "Sizing domestic batteries for load smoothing and peak shaving based on real-world demand data," Energy Build., vol. 247, p. 111109, Sep. 2021, doi: 10.1016/j.enbuild.2021.111109.

9. Appendix: Co-authorship Forms

THE UNIVERSITY OF WAIKATO
Te Whare Wānanga o Waikato

Co-Authorship Form

School of Graduate Research
The University of Waikato
Private Bag 3105
Hamilton 3240, New Zealand
Phone +64 7 838 5096
Email: SGR@waikato.ac.nz
Website: http://www.waikato.ac.nz/students/research-degre

This form is to accompany the submission of any PhD that contains research reported in published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in your appendices for all the copies of your thesis submitted for examination and library deposit (including digital deposit).

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.
Chapter 3: Design and Implementation of a Simple Low-Cost and Real-Time Solar Power Monitoring System
Name of the article:"Design and Implementation of a Simple Low-Cost and Real-Time Solar Power Monitoring System"
Where published: 2023 IEEE Fifth International Conference on DC Microgrids (ICDCM)
When published: November 15, 2023

Nature of contribution by PhD candidate	Conceived the idea, performed experiments, wrote the paper
Extent of contribution by PhD candidate (%)	70

CO-AUTHORS

Name	Nature of Contribution
Mark Apperley	Supervision, discussion, paper revision
Shafiqur Rahman Tito	Supervision, discussion, paper revision

Certification by Co-Authors

The undersigned hereby certify that:

- the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- that the candidate wrote all or the majority of the text.

Name	Signature	Date
Mark Apperley		06 June 2025
Shafiqur Rahman Tito		10/06/2025

July 2015

School of Graduate Research
The University of Waikato
Private Bag 3105
Hamilton 3240, New Zealand
Phone +64 7 838 5096
Email: SGR@waikato.ac.nz
Website: http://www.waikato.ac.nz/students/research-degre

THE UNIVERSITY OF WAIKATO
Te Whare Wānanga o Waikato

Co-Authorship Form

This form is to accompany the submission of any PhD that contains research reported in published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in your appendices for all the copies of your thesis submitted for examination and library deposit (including digital deposit).

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

Chapter 4:Data-Driven Predictive Modelling of Steam Turbines: Insights from Monitoring Systems ,
Name of the article:"Predicting Steam Turbine Power Generation: A Comparison of Long Short-Term Memory and Willans Line Model"
Where published: Energies — a peer-reviewed, open access journal published by MDPI (Multidisciplinary Digital Publishing Institute)
When published: January 10, 2024

Nature of contribution by PhD candidate	Conceived the idea, performed experiments, wrote the paper
Extent of contribution by PhD candidate (%)	70

CO-AUTHORS

Name	Nature of Contribution
Mark Apperley	Supervision, discussion, paper revision
Martin Atkins	Supervision, discussion, paper revision
Shafiqur Rahman Tito	Supervision, discussion, paper revision
Matthew Taylor	Performed experiments, contributed to paper

Certification by Co-Authors

The undersigned hereby certify that:

- the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- that the candidate wrote all or the majority of the text.

Name	Signature	Date
Mark Apperley		06 June 2025
Martin Atkins		10 June 2025
Shafiqur Rahman Tito		10/06/2025
Matthew Taylor		12 June 2025

July 2015

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

Co-Authorship Form

School of Graduate Research
The University of Waikato
Private Bag 3105
Hamilton 3240, New Zealand
Phone +64 7 838 5096
Email: SGR@waikato.ac.nz
Website: http://www.waikato.ac.nz/students/research-degree

This form is to accompany the submission of any PhD that contains research reported in published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in your appendices for all the copies of your thesis submitted for examination and library deposit (including digital deposit).

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.
Chapter 5: Enhancing Industrial Energy Efficiency with Predictive Analytics and Fuzzy Logic
Name of the article: "Enhancing Industrial Energy Efficiency with Predictive Analytics and Fuzzy Logic: A Case Study of Renewable Energy Management in the Meat Processing Industry" International Conference on Neural Information Processing (ICONIP 2024)
Published in 18th March 2025

Nature of contribution by PhD candidate	Conceived the idea, performed experiments, wrote the paper
Extent of contribution by PhD candidate (%)	70

CO-AUTHORS

Name	Nature of Contribution
Mark Apperley	Supervision, discussion, paper revision
Martin Atkins	Supervision, discussion, paper revision
Jason Kurz	Supervision, discussion, paper revision

Certification by Co-Authors

The undersigned hereby certify that:

- the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- that the candidate wrote all or the majority of the text.

Name	Signature	Date
Mark Apperley		06 June 2025
Martin Atkins		10 June 2025
Jason Kurz		09 June 2025

July 2015

School of Graduate Research
The University of Waikato
Private Bag 3105
Hamilton 3240, New Zealand
Phone +64 7 838 5096
Email: SGR@waikato.ac.nz
Website: http://www.waikato.ac.nz/students/research-degre

Co-Authorship Form

This form is to accompany the submission of any PhD that contains research reported in published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in your appendices for all the copies of your thesis submitted for examination and library deposit (including digital deposit).

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

Chapter 6: Edge-Driven Electricity Trading in Fractal-Structured Microgrids: A Machine Learning Approach
Name of the article: "Energy Management and Edge-Driven Trading in Fractal-Structured Microgrids: A Machine Learning Approach"
Where published:Energies — a peer-reviewed, open access journal published by MDPI (Multidisciplinary Digital Publishing Institute)
When published: May 30, 2025

Nature of contribution by PhD candidate	Conceived the idea, performed experiments, wrote the paper
Extent of contribution by PhD candidate (%)	70

CO-AUTHORS

Name	Nature of Contribution
Mark Apperley	Supervision, discussion, paper revision
Jason Kurz	Supervision, discussion, paper revision

Certification by Co-Authors

The undersigned hereby certify that:

- the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- that the candidate wrote all or the majority of the text.

Name	Signature	Date
Mark Apperley		06 June 2025
Jason Kurz		10 June 2025

July 2015