A HYBRID POWER DEMAND FORECASTING MODEL WITH UNCERTAINTY ESTIMATION UNDER INPUT PERTURBATIONS

KOMEN KIMORYOT FRANCIS

A Thesis Submitted to the Institute of Postgraduate Studies of Kabarak University in Partial Fulfillment of the Requirements for the Award of Doctor of Philosophy in Information Technology

KABARAK UNIVERSITY

NOVEMBER, 2025

DECLARATION

1. I do hereby declare that:

i. This thesis is my original work, and to the best of my knowledge, it has not been presented for the award of a degree in any University or College.
ii. The work has not incorporated material from other works or a paraphrase of such material without due and appropriate acknowledgment.
iii. The work has been subjected to processes of anti-plagiarism and has met Kabarak University's 15% similarity index threshold.

2. I do understand that issues of academic integrity are paramount, and therefore, I may be suspended or expelled from the university or my degree may be recalled for academic dishonesty or any other related academic malpractices.

Signed:________________________________ Date:____________________

Komen Kimoryot Francis

GDI/N/0263/01/21

RECOMMENDATION

To the Institute of Postgraduate Studies:

The thesis entitled "**A Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations"** written by **Komen Kimoryot Francis** is presented to the Institute of Postgraduate Studies of Kabarak University. We have reviewed the thesis and recommend that it be accepted in partial fulfillment of the requirements for the award of Doctor of Philosophy in Information Technology.

Signed:________________________________ Date:____________________

Dr. Moses M. Thiga

Department of Computer Science and Mathematics

Kabarak University

Signed:________________________________ Date:____________________

Dr. Andrew Kipkebut

Department of Computer Science and Mathematics

Kabarak University

COPYRIGHT

ACKNOWLEDGEMENT

This is to humbly acknowledge God for His sufficient grace in my academic journey. I also acknowledge my supervisors, Dr. Moses Thiga and Dr. Andrew Kipkebut, for their invaluable guidance in the development of this thesis. Their steadfastguidance, advice, and correctionsenabled the achievement of the study’s core objectives. I also acknowledge my colleagues for their encouragement throughout this study.

DEDICATION

To my mother, Rosaline Komen, and my late father, John Komen Arap Kiyai, for they raised me with compassionate love in the best way possible. To my wife, Mary, and children, for your love, encouragement, and support. My siblings: Julius Kibet, Kiptum, Jelagat, the late Jepkorir, Kiprotich, Jepchirchir, Kiprop, Kipsang, and Titus Kimutai, each of you for your encouragement and love.

ABSTRACT

Accurate power demand forecasting is crucial for effective electricity planning, resource allocation, and maintaining power grid stability, particularly in the context of dynamic consumption patterns and increasing system complexity. Traditional forecasting models often fall short when exposed to uncertainties caused by input perturbations such as fluctuations in voltage, current, and peak load, resulting in suboptimal forecasts and unreliable planning decisions. To overcome these limitations, this study developed a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations (HPDEF-MUIP). The study objectives were to assess the weaknesses of existing models, to design, implement, and evaluate the HPDEF-MUIP. The model was designed to enhance robustness and forecasting accuracy by fusing three top-performing base models: XGBoost, CatBoost, and Random Forest, each offering complementary strengths in computational scalability, handling of categorical features, and precision in regression tasks. The research employed a quantitative methodology using half-hourly historical power demand data from a Kenya Power and Lighting Company substation, supplemented with meteorological data from the Nakuru Meteorological Station, and critical input features prone to input perturbation. After comprehensive data preprocessing, including cleaning, normalization, and feature engineering, the hybrid model was rigorously trained and validated on real-world data from regional power substations to ensure operational relevance. The findings show that the model achieved an R^2 of 0.9539, an RMSE of 1.7128, and an MAPE of 3.12%, along with a robust F1-score of 0.9112, demonstrating superior performance in both regression and classification metrics. A notable contribution of this study is the incorporation of a structured uncertainty estimation mechanism through input perturbation analysis, which provides critical insights into model reliability under non-ideal operating conditions—a crucial aspect often overlooked in conventional approaches. The HPDEF-MUIP model not only improved accuracy over individual base models but also enhanced interpretability and resilience, making it suitable for deployment in edge computing environments and adaptive smart grid infrastructures. The study concluded that the HPDEF-MUIP represented a significant advancement in forecasting methodologies by effectively bridging the gap between forecasting accuracy, uncertainty estimation, and operational feasibility in a real-world context. Based on these findings, the study recommends the adoption of hybrid uncertainty-aware models within power forecasting systems to enhance the planning and stability of the smart grid. Formulation of supportive policies is also recommended, and further research is suggested to expand the scalability and interpretability of Hybrid models across diverse Power systems.This research contributes a scalable and replicable framework that supports Sustainable Development Goals (SDGs) through more informed decision-making in load balancing, infrastructure planning, and sustainable energy management.

Keywords: *HPDEF-MUIP, Deep Learning, Power Demand Forecasting, Hybrid, Perturbations, Uncertainty Estimation*.

TABLE OF CONTENTS

DECLARATION .... **ii**
RECOMMENDATION .... **iii**
COPYRIGHT .... **iv**
ACKNOWLEDGEMENT .... **v**
DEDICATION .... **vi**
ABSTRACT .... **vii**
TABLE OF CONTENTS .... **viii**
LIST OF TABLES .... **xiv**
LIST OF FIGURES .... **xv**
OPERATIONAL DEFINITION OF TERMS .... **xvi**
CHAPTER ONE **1**
INTRODUCTION **1**
1.1 Introduction 1
1.2 Background to the Study 1
1.3 Statement of the Problem 11
1.4 Objectives of the Study 12
1.4.1 General Objective 12
1.4.2 Specific Objectives 12
1.5 Research Questions 13
1.6 Justification for the Study 13
1.7 Significance of the Study 14
1.8 Scope of the Study 15
1.9 Limitations/Delimitations of the Study 15
CHAPTER TWO **18**
LITERATURE REVIEW **18**
2.1 Introduction 18
2.2 Review ofthe Weaknesses of Existing Hybrid Deep Learning Models 18
2.2.1 Weaknesses of Existing Power Demand Forecasting Models with Uncertainty Estimation Under Input Perturbations for Accurate Forecasting Models 40
2.2.2 Methodology for Identification of the Weaknesses of Existing Deep Learning Forecasting Models 41

2.2.3 Weakness 1: Data Complexity and Sensitivity 42
2.2.4 Weakness 2: Lack of Flexibility 43
2.2.5 Weakness 3: Performance Variability 43
2.2.6 Weakness 4: Risks of Overestimation 43
2.2.7 Weakness 5: Inadequate Temporal Dynamics 43
2.2.8 Weakness 6: Model Complexity 44
2.2.10 Design Recommendations to Address the Identified Weaknesses of Existing Power Demand Forecasting Models 45
2.2.11 Summary of Specific Objective One: Weaknesses of Existing Models 47
2.3 Review of Existing Hybrid Models Design 49
2.4 Review of Previous Study Model's Implementation Methods 51
2.5 Review of Existing Evaluation Methods 64
2.6 Perturbation Theory 72
2.7 Extreme Gradient Boost 75
2.7.1 Advantages of XGBoost 75
2.7.2 Weaknesses of XGBoost 76
2.8 Light Gradient Boosting Machine (LightGBM) 77
2.8.1 Advantages of LightGBM 78
2.8.2 Weaknesses of LightGBM 79
2.9 Categorical Boosting (CATBOOST) 80
2.10 K-Nearest Neighbors Modeling Concept 82
2.11 Artificial Neural Network (ANN) Concept 86
2.11.1 Advantages of ANNs 88
2.11.2 Weaknesses of ANN and Their Role in Hybrid Model 89
2.12 Bayesian Neural Network Concept 90
2.13 Random Forest Concept 97
2.13.1 Advantages of Random Forest 98
2.13.2 Weaknesses of Random Forest 98
2.14 Hybridization Concept in Machine Learning 99
2.15 Bootstrap Aggregating Technique 100
2.15.1 Advantages of Bagging 100
2.15.2 Boosting method 100
2.15.3 Advantages of Boosting 101
2.15.4 Stacking Method 101

2.15.5 Advantages of Stacking 102
2.15.6 VotingMethod 102
2.15.7 Advantages of Voting 102
2.15.8 Blending Method 103
2.15.9 Advantages of Blending 103
2.16 Research Gap 104
2.17 Research Gap 105
2.18 Proposed Conceptual Framework 107
CHAPTER THREE 108
RESEARCH METHODOLOGY 108
3.1 Introduction 108
3.2 Research Philosophy 108
3.3 Quantitative Research Design 111
3.4 Meta-analysis Literature Review Research Design 111
3.5 Deep Learning Experimental Research Design 113
3.5.1 Data Collection and Pre-processing 115
3.5.2 Feature Selection Engineering 116
3.5.3 Model Training 116
3.5.4 Bayesian Neural Network Training 117
3.5.5 Support Vector Machine (SVM) Optimization 121
3.5.6 Artificial Neural Network (ANN) Learning 123
3.5.7 Extreme Gradient Boosting (XGBOOST) 126
3.6 Design of LightGBM Model 130
3.6.1 Objective Function 131
3.6.2 LightGBM Loss Function 131
3.6.3 Regularization Term 131
3.6.4 Prediction in Gradient Boosting 132
3.6.5 Splitting Criteria for Decision Trees 132
3.6.6 Gradient and Hessian Calculations 132
3.6.7 Feature Importance in LightGBM 133
3.6.8 Adapting LightGBM for Time Series 133
3.7 Estimation of Uncertainty 134
3.7.2 Monte Carlo Simulation 134
3.7.3 Quantile Regression 134

3.8 CATBOOST Model Design 134
3.8.1 Model Representation 135
3.8.2 Objective Function 135
3.8.3 Ordered Boosting 135
3.8.4 Decision Tree Splitting Criterion 136
3.8.5 Handling of Categorical Features 137
3.8.6 Prediction Update Process 137
3.8.7 Time Series Design Methodology 137
3.8.8 Uncertainty Estimation methodology in CATBoost 138
3.9 K-Nearest Neighbors (K-NN)Base Model Design Process 139
3.9.1 Design of a K-NN Base Model 139
3.9.2 K-NN Regression with Time Distance Metric 140
3.9.3 Hybrid Prediction Mechanism 141
3.10 Hybrid Model Methodology 142
3.11 Perturbation Methodology 143
3.12 Model Implementation 144
3.13 Model Evaluation 144
3.14 Model Testing and Interpretation 145
3.15 Ethical Considerations 145
CHAPTER FOUR 148
DATA ANALYSIS, PRESENTATION AND DISCUSSION 148
4.1 Introduction 148
4.2 Design of a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations 148
4.2.1 Data Pre-Processing 149
4.2.2 Feature Selection 150
4.2.3 Interpretation of Feature Importance Results 151
4.2.4 Splitting the Data 153
4.3 Model Design Results 155
4.4 Model Design Performance Results Discussion 156
4.4.1 XGBoost Base Model Design Results Analysis 156
4.4.2 CatBoost Base Model Design Results Analysis 157
4.4.3 Random Forest Base Model Design Results Analysis 158
4.4.4 Design Results Discussion 161

4.5 Model Implementation 164
4.5.1Model Architecture 164
4.5.2 Hybrid Architecture Implementation 164
4.5.3 Hybrid Fusion Methodology 164
4.5.4 Uncertainty Estimation under Input Perturbations 165
4.5.5 Adaptive Learning and Real-Time Feedback 166
4.5.6 Interpretability and Decision Support 166
4.5.8 Operational Significance 167
4.5.9 Model code Implementation 167
4.5.10 Output Forecast 168
4.5.12 Input Perturbation Analysis 169
4.5.13 Comparison Between Actual, Perturbed, and Unperturbed Model Forecast Results Interpretation and Discussion 172
4.5.14 Uncertainty Quantification 173
4.5.15 Uncertainty Estimation Under Input Perturbations 174
4.5.16 Input Perturbation Handling 174
4.5.17 Uncertainty Estimation Implementation 175
4.6 Hybrid Model Evaluation Results 176
4.6.1 Model Performance Results 176
4.6.2 Results for Forecast Demand Versus Actual Demand Interpretation and Discussion 177
4.6.3 Actual Power Demand versus Perturbed Input Demand Forecast Results Interpretation and Discussion 180
4.6.4 Actual Versus Non-Perturbed Model Results Interpretation 183
4.6.6 Comparative Performance of Base Models Versus HPDEF-MUIP 184
4.6.7 HPDEF-MUIP Versus Base Model Results and Discussions 185
4.6.8 Theoretical Impact 188
4.6.9 Practical and Research Implications 188
4.7 Summary of Chapter Four: Data Analysis, Presentation and Discussion 189
CHAPTER FIVE 191
SUMMARY, CONCLUSIONS AND RECOMMENDATIONS 191
5.1 Introduction 191
5.2 Summary 191

5.2.1 Assessing the weaknesses of existing Power Demand Forecasting Models with Uncertainty Estimation under Input Perturbations................ 191
5.2.2 Design of a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations.................................... 192
5.2.3 Implementation of a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations.................................... 194
5.2.4 Evaluation of a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations.................................... 194
5.3 Conclusion 195
5.4 Recommendations.. 198
5.4.1 Recommendations for Objective 1: To assess the weaknesses of existing Power Demand Forecasting Models with Uncertainty Estimation under Input Perturbations 198
5.4.2 Recommendations for Objective 2: To design a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations 198
5.4.3 Recommendations for Objective 3: To implement a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations using Real-World Power System Data 199
5.4.4 Recommendations for Objective 4: To evaluate the accuracy of a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations 199
5.4.1 Policy Recommendations 200
5.4.2 Recommendations for Future Studies... 200
REFERENCES 202
APPENDICES... 220
Appendix 1: Code Listing.. 220
Appendix II: KUREC Clearance Letter 226
Apendix III: NACOSTI Reserch Permit 227
Appendix IV: Evidence of Conference Participation.. 228
Appendix V: List of Publications 229

LIST OF TABLES

Table 1: Summary of Weaknesses of Existing Models 44

Table 2: Research Gap 104

Table 3: Table of Model Design Results 156

Table 4: Top Three Accurately Performing Base Models Design Weights 160

Table 5: Perturbation Performance Assessment Results of Perturbed Input Data Versus Unperturbed Input Data 170

Table 6: Comparative Performance of Base Models Versus HPDEF-MUIP 184

LIST OF FIGURES

Figure I: Conceptual Framework 107

Figure 2: Experimental Research Design 114

Figure 3: Feature Importance 151

Figure 4: Hybrid Model Design Architecture..163

Figure 5: Comparison Between Actual, Perturbed, and Unperturbed Model Forecasts 171

Figure 6: Forecasted Demand Versus Actual Power Demand 177

Figure 7:HPDEF-MUIP Model 179

Figure 8: Results for Actual Power Demand Versus Perturbed Input Demand Forecast 180

Figure 9: Actual Power Demand versus Non-Perturbed Demand Forecast 182

Figure 10:HPDEF-MUIP Vs Base Model Performance Comparison 185

OPERATIONAL DEFINITION OF TERMS

Adversarial training: It involves perturbing the input data during training to make the model more robust to perturbations at test time(Jeon, 2022).

Deep Learning: Deep learning models are a type of Artificial Neural Network that can learn from large amounts of unstructured data, such as images, speech, or text (Wang F. Z., 2022)**.**

GWh: Gigawatt hour (EPRA, 2023)

HPDEF-MUIP: Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations HPDEF-MUIP(Author,2024)

Hybrid :Combination of two or more deep learning methods or techniques (Hong T. Y., 2021).

Input Perturbations: The process of making small, intentional changes to the input data (Wang F. Z., 2021).

MW: Megawatts (EPRA, 2023)

Peak Demand: A measure of the highest load demand in the interconnected network over a given period of time (EPRA, 2023)

SDGs :Sustainability Development Goals (Zakari, 2022)

Uncertainty Estimation:It refers to the process of quantifying the uncertainty associated with a particular measurement, prediction, or model (Ryu, 2019).

CHAPTER ONE

INTRODUCTION

1.1 Introduction

This Chapter provides a brief background of the study and the challenges that inform the development of a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations (HPDEF-MUIP). It then proceeds to state the research problem, outlining the specific research objectives, enumerating the research questions, justifying the study, defining the scope, and specifying the study's significance and limitations.

1.2 Background to the Study

The efficiency, reliability, and stability of the power supply used in industry, local consumption, and in the infrastructure supporting the World Wide Web are critical factors that contribute to the overall performance, sustainability, and resilience of the global Internet ecosystem (Meydani, 2023).The reliability and stability of electricity supply are essential for the smooth and efficient operation of the global Internet ecosystem. Power perturbations, such as voltage fluctuations, frequency variations, and interruptions, can have serious consequences, including equipment damage, data loss, productivity disruptions, and service outages. These issues can significantly affect the availability, performance, and security of web services, online applications, and the underlying infrastructure that supports the Internet. Ensuring a reliable and stable power supply is crucial to maintaining the uninterrupted operation of data centers, network equipment, and other critical components that form the backbone of the World Wide Web. Addressing power uncertainty issues is crucial to supporting a resilient, sustainable, and trusted Internet ecosystem that can reliably serve the needs of billions of users around the world (Wang Z. H., 2023).

Perturbation in deep learning forecasting models involves making small, intentional changes to the input data to understand how these changes affect the model's output. Input Perturbations (IP) also refer to intentional or unintentional changes made to the input data of a Deep Learning (DL) model, such as data transformations and adversarial attacks. These perturbations can occur for various reasons, including during data collection, data preprocessing, and data augmentation. Perturbations can significantly affect the performance of deep learning models, and Uncertainty Estimation can be crucial for realizing reliability and robustness. The goal of perturbation is to analyze the model's sensitivity to changes in input data and identify areas where the model may be less robust or reliable (Wang F. Z., 2021).

Electricity consumption exhibits seasonal patterns, characterized by higher power demand for air conditioning in the summer and heating in the winter. Changing seasonal factors in the input data can help the Deep Learning (DL) models adapt to changes in power demand. Weather conditions, such as temperature and humidity, affect electricity consumption, and simulating weather changes can help improve these models. Economic factors such as GDP growth and consumer spending affect electricity demand, and disruptive economic indicators can simulate changes in energy consumption patterns. Government policies, infrastructure changes, and unexpected events can also be simulated in the input data to test the model's response to different scenarios (Zhang Y. R., 2023).

Deep learning models comprise multiple hidden layers, enabling them to learn complex representations of input data and extract relevant features. Deep learning models are used in a wide range of applications, including computer vision, natural language processing, speech recognition, and autonomous driving. Some popular deep learning architectures include Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Transformers (Wan, F, 2022).

Uncertainty Estimation (UE) involves determining the range of possible outcomes that could result from the inherent variability, errors, and assumptions in the data or model being used (Ryu, 2019). Derived from previous research investigations, the Bayesian neural network (BNN) has been utilized for quantifying the uncertainties in the modeling of VRF systems. These uncertainties have been broken down into Aleatoric Uncertainty (AU), which is generated by data noise, and Epistemic Uncertainty (EU), which is generated by the model's incompleteness due to the absence of knowledge about the system. Aleatoric Uncertainty has also been termed Irreducible Uncertainty, Inherent Uncertainty, Variability, and Stochastic Uncertainty, whereas Epistemic Uncertainty has also been termed Reducible Uncertainty, Subjective Uncertainty, and Cognitive Uncertainty.

In the 2023/2024 power cycle, Kenya's total installed power generation capacity reached 3,713 megawatts (MW), with 3,311.1 MW classified as "effective" capacity, representing the amount contracted by Kenya Power and Lighting Company (KPLC) for off-take purposes. Kenya’s monthly peak demand stands at 2,149MW. Electricity consumption increased from 12,652.74 GWh in 2022 to 13,294.63 GWh,representing a7.46% upward growth rate (EPRA, 2023). The total electricity delivered to the national grid increased by 21.84% from 4,953.15 GWh to 6,035.00 GWh, with KENGEN accounting for a market share of 60.74% of the total generated electricity. From 2017 to date, Kenya has experienced a 5% increase, which correlates with the country’s economic growth. The Energy and Petroleum Regulatory Authority (EPRA) published a report showing that a peak electricity demand of 2,149MW in the 2022/2023 cycle represents a 4.5% growth. Power users desire a reliable electric supply. However, the

Customer Average Interruption Index is reported to be 2.25 hours per month,which is attributed to power outages. In March 2023, the longest blackout lasted 2.7 hours per customer, while the shortest blackout, in February 2023, lasted 1.54 hours per customer. These outages contribute to the causes of uncertainty in power supply and reliability.

Uncertainties in power demand forecasts impact power utilities, such as Kenya Power and Lighting Company (KPLC). These uncertainties can affect KPLC's long-term production planning. If power demand is overestimated, power generation utilities may invest in excess production capacity, resulting in underutilization of assets and increased costs. Conversely, if future power demand is underestimated, there may be insufficient generation capacity to meet customer needs, leading to potential outages or load shedding. A study revealed a scenario in which the generation of electricity by Kenyan power utilities was lower than the mandated requirements (Takase M., 2021). Inaccurate demand forecasts can affect the operational efficiency of KPLC's generation, transmission, and distribution systems. Suppose actual demand deviates significantly from forecasted levels. In that case, KPLC may need to adjust production schedules, supply additional power from expensive sources such as diesel-powered generators, or implement contingency measures to balance supply and demand, resulting in increased operating costs and reduced efficiency.

Uncertainties in demand forecasting affect utilities' investment decisions regarding infrastructure upgrades, expansion projects, and network modernization initiatives (Ibrahim, 2023). If demand is expected to grow rapidly, KPLC may need to invest in new generation capacity, transmission lines, and distribution networks to meet future demand. However, if demand growth is slower than expected, KPLC may face stranded assets and financial losses. It is therefore essential that power utilities make accurate forecasts for the future to support informed decision-making. Inaccurate demand forecasts may affect power utilities' revenue management and financial sustainability. If demand is overestimated, KPLC may overcharge customers or impose higher tariffs to cover the cost of excess generation capacity, leading to customer dissatisfaction and potential regulatory scrutiny. Conversely, if demand is underestimated, KPLC may struggle to recoup its fixed costs, leading to revenue shortfalls and financial instability (Waswa, 2024).

Fluctuations in the accuracy of demand forecasting may impact KPLC's ability to deliver reliable and high-quality electricity to customers. If power demand exceeds forecast levels, KPLC may experience grid congestion, voltage fluctuations, and reliability issues, leading to service disruptions and customer complaints. Conversely, if demand is lower than expected, KPLC may underutilize its infrastructure, failing to meet customer expectations for reliability and service quality (Strezoski A., 2021).

Uncertainties in forecasting power demand pose significant challenges for power supply utilities in terms of production planning, operational efficiency, investment decisions, revenue management, and customer service. To mitigate these issues, KPLC may need to enhance its forecasting methodologies, invest in advanced technologies, improve stakeholder collaboration, and adopt flexible and adaptive planning approaches to better anticipateand respond to changes in electricity demand patterns (Eberhard, 2018;Kiptoo, 2023).

Deep learning models are crucial in various fields, including computer vision, natural language processing, and autonomous systems. These models are trained on fixed datasets and evaluated based on their ability to work with new data. However, in real situations, the input data may be perturbed due to factors such as noise, errors, or environmental changes. Therefore, it is important to ensure that deep learning models handle these changes and make accurate power demand forecasts. Evaluating the model's performance under these conditions is crucial for assessing its real-world applicability and reliability (Rice, 2020).

Electricity Demand Forecasting (EDF) is the process of estimating future patterns and levels of electricity consumption. It involves analyzing historical data, considering various factors such as weather, economic conditions, population growth, and technological advancements to predict future electricity demand. Forecasting electricity demand is essential for effective energy planning and management. It helps energy companies and grid operators make informed decisions about power generation capacity, transmission, and distribution. Accurate forecasting enables efficient resource allocation, cost optimization, and planning of power supply infrastructure, ensuring a reliable and stable electricity supply meets the growing demands of consumers (Impram, Nese, & Oral, 2020).

Electricity Demand Forecasting can be accomplished using various methods, including statistical models, Time Series Analysis, Econometric Models, Artificial Intelligence Techniques, and Machine Learning Algorithms. These methods analyze historical data, consider relevant variables, and apply mathematical and statistical techniques to create future forecasts.

Statistical models, such as AutoRegressive (AR), Moving Average (MA), and AutoRegressive Integrated Moving Average (ARIMA), often assume linear relationships between variables, which may not accurately capture the nonlinear and complex relationships present in load data. These models have a limited capacity to represent complex elements and struggle with high-dimensional relationships or interactions. They are sensitive to outliers and anomalies, struggle with long-term forecasts, and are not adaptable to dynamic environments. Additionally, statistical models may not be able to capture the complex dependencies and interactions present in load data (Ashouri, 2022).

Econometric models used in energy demand forecasting have limitations, including the assumption of linear relationships between variables that do not capture non-linear patterns in power demand data. Endogeneity problems can arise when power demand and its determinants interact with each other and with unobserved factors. In addition, econometric models have limited explanatory power, are constrained by data limitations, and struggle with issues such as seasonality, non-stationarity, and long-term forecasting (Onder, 2022).

Time series analysis in electricity demand forecasting presents several limitations. First, the stationary assumption is often violated due to factors such as seasonality, trends, and external influences, leading to non-constant statistical properties over time.

The complex patterns exhibited by electricity demand data, including numerous seasonal effects, irregular fluctuations, and long-term trends, present challenges for traditional time series models to capture these nuances effectively (Kang, 2023). The limited forecast horizon of time series models, which are more suitable for short- to medium-term forecasts, makes it difficult to accurately forecast electricity demand several periods into the future, especially years ahead. In addition, the focus of many time series models is limited to historical demand data, excluding other relevant variables, which can compromise the accuracy and robustness of forecasts. The complexity of time series models also poses the risk of overfitting with overly complex models and biasing predictions with overly simplified ones. Moreover, the sensitivity of time series models to outliers and their limited adaptability to structural changes in the data further affect forecast accuracy (Jiang, 2020).

Artificial intelligence (AI) models provide powerful tools for forecasting electricity demand; however, they have several limitations. For example, neural networks often suffer from overfitting when trained on large datasets, resulting in poor generalization to unseen data (Wang R. L., 2020). Deep Learning models, such as recurrent neural networks (RNNs) and long short-term memory (LSTM) networks, often require large amounts of data for training and may strugglewith long-term forecasting (Bessa et al., 2018;Sarmas, 2023). Artificial Intelligence (AI) models also lack interpretability, making it challenging to understand the underlying factors that influence forecasts. These AI models are sensitive to outliers and noisy data, leading to inaccurate predictions (Ouyang, 2021). The AI models also struggle to capture the complex relationships and non-linearity present in electricity consumption data (Luo, 2020).

Machine learning (ML) algorithms such as Support Vector Machines (SVM), Random Forests, Gradient Boosting Machines (GBM), Artificial Neural Networks (ANN), Long Short-Term Memory (LSTM), and Convolutional Neural Networks (CNN) offer powerful techniques for forecasting electricity demand. However, they also have limitations. SVMs work well for small to medium-sized datasets with high-dimensional feature spaces, but may struggle with large datasets. Random forests are resistant to over-fitting, but may not perform well with highly correlated input variables. GBM often provides high predictive accuracy but is prone to over-fitting (Ge, 2020). ANNs can capture complex patterns and non-linear relationships, but require a large amount of data for training (Khan, 2022). LSTMs are well-suited for time series forecasting tasks; however, they struggle to capture seasonal patterns and sudden changes in the data. CNNs are mainly used for image recognition.

The accuracy of the forecasts depends on the quality of the data, the chosen methodology, and the ability to capture the complex factors influencing electricity demand (Qi, 2020). Forecasting electricity demand presents several challenges, including issues with data quality and availability. The accuracy of forecasts depends on the quality and availability of data. Incomplete or inaccurate data can lead to incorrect or unreliable forecasts. Changes in consumer behavior, such as the adoption of energy-efficient appliances, electric vehicles, and renewable energy sources, can significantly impact electricity demand patterns, making it challenging to accurately predict future demand. Weather conditions play a vital role in electricity demand. Extreme weather events, such as heatwaves, cold snaps, and storms, can cause sudden spikes or drops in electricity demand, making accurate forecasting difficult. Challenges related to economic and regulatory uncertainty, including economic conditions and regulatory policy, also impact electricity demand. Changes in economic growth, energy prices, and government policies can lead to significant shifts in electricity consumption patterns.

Researchers and practitioners have developed various techniques and approaches to address the challenges of electricity demand forecasting, including the inaccuracy of forecasting models. For example, techniques such as data cleaning, data transformation, and imputation are used to improve data quality and completeness. Including external factors such as weather data, economic indicators, and demographic variables can increase the accuracy of forecasts. Ensemble methods, which combine multiple forecast models, can improve accuracy and reduce errors. Deep learning techniques, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have demonstrated promise in improving the accuracy of forecasting electricity demand by capturing complex patterns and relationships within the data (Aurangzeb, 2021).

Developing robust techniques for electric demand forecasting. Uncertainty Estimation in deep learning models under input perturbations has been an active and ongoing area of research in recent years. Various approaches have been proposed, including Bayesian Neural Networks, Deep Ensembles, and Monte Carlo Dropout. These methods provide a measure of uncertainty that can be used to identify out-of-distribution inputs, detect adversarial attacks, and improve the overall robustness of deep learning models (Hosseini, 2020).

In recent years, researchers have shown increasing interest in leveraging Deep Learning (DL) models for Uncertainty Estimation. These models have the potential to provide more accurate and reliable uncertainty estimates by learning directly from the data and capturing the inherent complexities of electricity demand patterns (Qi, 2020). However, the impact of input perturbations on Uncertainty Estimation in DL models for electricity demand forecasting remains largely unexplored.

Machine Learning (ML) approaches such as Regression Analysis (RA), Decision Trees (DT), Support Vector Machines (SVMs), Artificial Neural Networks (ANNs), and DL techniques have been used in electricity demand forecasting. A current challenge with ML approaches is choosing the most appropriate algorithm for a specific problem. Each algorithm has its strengths and weaknesses, and choosing the right one requires a thorough understanding of the problem and the data. The performance of ML models depends on the quality and quantity of data available for training and validation. Insufficient or poor-quality data can lead to over- or under-fitting of models, resulting in inaccurate forecasts (Ghazal, 2022).

Research on diffusion in Neural Networks (NN) has been partially done. Diffusion in Neural Networks (NN) refers to the flow of information through layers during training, which may encounter problems such as vanishing or exploding gradients. Uncertainty Estimation (UE) models can be used to solve diffusion problems by providing a measure of confidence or uncertainty in network predictions. This can help calibrate the model's confidence levels and prevent it from making overly reliable or unreliable predictions. By incorporating UE models, Neural Networks can gain better control over diffusion-related problems, leading to improved performance, robustness, and reliability of their predictions (Huang J. Z., 2020).

Several studies have demonstrated the effectiveness of these techniques in improving the robustness of Deep Learning models. For instance, a recent study on Deep Ensembles (DE) improved the Uncertainty Estimation in image classification tasks (Ren, 2022). The study demonstrated how DE improved Uncertainty Estimation in image classification tasks, especially in identifying out-of-distribution images. However, the study did not compare the proposed method with other state-of-the-art Uncertainty Estimation techniques. The study only focused on image classification tasks, and it is unclear whether deep ensembles would be equally effective in other Deep Learning tasks.

Deep learning approaches have been used to solve problems in power Demand Forecasting. These approaches includeFeature Engineering (FE), the use of Recurrent Neural Networks (RNN) and Long-Short-Term Memory Networks to account for temporal dependencies, Ensemble Methods, Hybrid Models, Hyperparameter Optimization, Transfer Learning, the use of Deep Learning models to capture complex relationships, and Anomaly Detection. Continuous Learning, including exogenous variables and uncertainty quantification, continues to interest researchers. These methods are continually researched and refined to enhance the accuracy of energy demand forecasts; however, their effectiveness varies depending on the specific characteristics of the input data and the context in which they are applied (Ahmad, 2022).

1.3 Statement of the Problem

The Sustainable Development Goals number nine on Industry Innovation and Infrastructure (SDGs 7 and 9) calls for accurate power demand forecasting (Mondejar, 2021). It supports the efficient design and management of power infrastructure for digital sustainability. This enables the development of smart grids, optimized energy storage solutions, and the integration of renewable energy sources, fostering innovation in the energy sector (Zakari, 2022). Forecasting Electricity Demand using Deep Learning models has increasinglyattracted scholarly attention from many researchersdue to their superior capacity for learning nonlinear consumption patterns and adapting to complex energy system dynamics. However, the application of Deep Learning models has failed to address uncertainties arising from input perturbations in model training, resulting in inaccurate and unreliable power demand forecasts (Smith A. B., 2021). This shortcoming underscores the need for more resilient and uncertainty-aware forecasting frameworks that can reliably perform under real-world perturbations.This provides precise knowledge for quantifying uncertainty in Deep Learning (DL) models, offering a straightforward way to factor in the challenges of diffusion, generalization, and transferability.

1.4 Objectives of the Study

1.4.1 General Objective

To develop, implement, and evaluate a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbationsin order to improve the accuracy, robustness, and reliability of short-term electricity demand forecasting.

1.4.2 Specific Objectives

i. To assess the weaknesses of existing Power Demand Forecasting Models with Uncertainty Estimation under Input Perturbations

ii. To design a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations

iii. To implement a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbationsusing real-world power system data.

iv. To evaluate the accuracy of a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations.

1.5 Research Questions

i. How can the weaknesses of existing Power Demand Forecasting Models withUncertainty Estimation under Input Perturbations be assessed?
ii. How can the Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations be designed?
iii. How can the Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations be implemented?
iv. How canthe accuracy of a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations be evaluated?

1.6 Justification for the Study

Accurate power demand forecasting is critically important not only for optimizing resource allocation and improving operational efficiency within electricity utilities, but also for advancing Kenya's broader national energy goals. In alignment with Kenya Vision 2030 and the SDG 9 goal of Industry Innovation and Infrastructure, reliable forecasting supports the transition towards smarter, more sustainable, and resilient power systems. Enhanced predictive capability enables utilities to balance demand and supply more effectively, reduce system losses, manage peak-load pressures, and minimize reliance on costly standby generation. Improved forecasting accuracy strengthens evidence-based decision-making for grid expansion, renewable energy integration, and policy formulation,ultimately contributing to universal electrification, economic growth, and long-term energy security for the country.

1.7 Significance of the Study

This study is significant as it advances the scientific and practical understanding of hybrid power demand forecasting under uncertainty, an area that remains insufficiently addressed in contemporary energy analytics research. The study contributes to theory by integrating hybrid model learning with perturbation-based uncertainty estimation, thereby extending existing scholarly work on robust and resilient forecasting models beyond accuracy optimization to include reliability under noisy and volatile real-world conditions. This integration enhances theoretical insight into model accuracy, input variability, and uncertainty quantification, providing a replicable methodological foundation for future studies across forecasting domains. Practically, the model enhances operational decision-making for power utilities by improving the accuracy and dependability of short-term forecasts used in generation planning, grid stability management, demand-response scheduling, and cost optimization.

Additionally, it supports deployment within smart grid and edge computing environments for real-time power system intelligence. From a policy and societal perspective, the study provides an empirical basis for data-driven energy planning, resource allocation, tariff structuring, and infrastructure investment, aligning with national energy strategies and global sustainability priorities, particularly SDG 7 (Affordable and Clean Energy) and SDG 9 (Industry, Innovation, and Infrastructure). Methodologically, the research presents a novel uncertainty-aware hybrid forecasting framework that can be adapted to other sectors, such as finance, transport, climate, and water resource forecasting, which are faced with volatile and uncertain data environments. Collectively, this study strengthens both academic inquiry and practical energy sector transformation by bridging the gap between theoretical modeling advances and real-world implementation, thereby facilitating the development of sustainable, intelligent, and resilient power systems.

1.8 Scope of the Study

The scope of the study was limited to design and implementation, focusing on forecasting power demand with uncertainty estimation under input perturbations using power load data.This study focused on improving the accuracy of electricity demand forecasting models by addressing the uncertainties associated with input perturbations. The study analyzed the uncertainties associated with deep learning models and developed strategies to mitigate their impact on prediction accuracy. The study also investigated the effects of variations in input data on model predictions and used the results to increase the reliability and accuracy of Power Demand forecasts in the Kenyan Central Rift Region.

The ultimate goal of the study was to enhance the accuracy of electricity demand forecasting models and apply the resulting knowledge to benefit the electric power industry. The methodological approach included theoretical analysis, empirical investigation, and computational modeling to design and test new algorithms and techniques for uncertainty estimation, thereby improving the model. The study acknowledges that it could not cover all potential sources of uncertainty.Data were collected using Power Spreadsheets from KPLC Central Rift Region smart meters for the period between June 2018 and September 2024. Meteorological data were collected for the same period at the Nakuru Meteorological station. Both secondary and primary data were collected and used in the study.

1.9 Limitations/Delimitations of the Study

This study was subject to several methodological and contextual limitations. First, the implementation of a hybrid ensemble combined with uncertainty-estimation techniques introduced significant computational complexity, making the experimentation process resource-intensive and time-consuming. Second, the dataset contained missing values arising from prolonged power outages and meter downtime, which necessitated imputation and may have affected data fidelity. Third, the performance of the proposed model cannot be generalized to all deep learning architectures or forecasting scenarios, as model effectiveness may vary depending on the dataset characteristics, model architecture, and forecast horizon. Furthermore, the reliability of the model was highly dependent on the availability and quality of historical input data; limited or low-quality datasets constrained optimal training and validation.

The study also faced constraints related to computing power, which restricted the scale of hyperparameter optimization and the number of perturbation simulations. Finally, the findings possess limited external validity beyond short-term electricity demand forecasting, and caution should be exercised when extrapolating the outcomes to other sectors, countries, or energy forecasting contexts without further empirical validation. In order to maintain a clear scope and methodological feasibility, this study was intentionally delimited in several ways. It focused specifically on assessing weaknesses in existing power demand forecasting models, and on designing, implementing, and evaluating a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations (HPDEF-MUIP).

The study employed electricity demand and meteorological data only, excluding other potential predictors such as demographic, economic, or industrial activity indicators. Additionally, the research concentrated on short-term forecasting and did not extend to medium- or long-term forecasting horizons. Only selected machine learning models were considered. Although multiple models were initially explored, the hybrid architecture was limited to the three best-performing algorithms: XGBoost, CatBoost, and Random

Forest. The evaluation focused on enhancing predictive accuracy and quantifying uncertainty. Deployment considerations, such as integration with operational grid management systems or real-time forecasting platforms, were beyond the scope of the present study and are proposed for future research.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

This chapter reviews literature on the weaknesses of existing forecasting models, designing of HPDEF-MUIP, implementation of HPDEF-MUIP and Evaluation of HPDEF-MUIP. Research gaps and the study's Proposed Conceptual Framework is also discussed.

2.2 Review ofthe Weaknesses of Existing Hybrid Deep Learning Models

Literature on Uncertainty Estimation Techniques (UET) has been reviewed to understand the existing different techniques for Uncertainty Estimation in deep learning models. These techniques include Bayesian Methods, Ensemble Methods, and Deep Learning-Based Methods such as dropout and Stochastic Regularization.Existing hybrid energy demand forecasting models exhibit several shortcomings that hinder their performance. These models often deal with complex nonlinear relationships, multivariate dependencies, and long-term dependencies that are critical for accurate forecasting in dynamic environments. Many of the models fail to adequately capture the stochastic nature of renewables, resulting in less than desirable performance in short-term power forecasting (Huang A. B., 2024).

Many hybrid models do not effectively manage the complex non-linear interactions between electricity demand and influencing factors, resulting in reduced accuracy. Some models lack the ability to capture the long-term dependencies and temporal dynamics necessary for reliable forecasting. The performance of existing models can vary significantly depending on the geographic context and specific characteristics of the datasets used, suggesting a lack of generalizability (Liu M. L., 2023).

A research study was done to investigate the need for advanced probabilistic load forecasting techniques, it highlighted the poor effectiveness of hybrid models in dealing with uncertainty, and improves forecast accuracy in different scenarios. The weakness of this model was that Single-model approaches run into inherent flaws during forecasting. The researchers recommended that further research would explore the incorporation of machine learning algorithms to improve predictive capabilities. In addition, it was recommended that the impact of consumer behavior patterns on load forecasting should be investigated to further refine the accuracy of the model (Yang Y. X., 2024).

Previous studies on hybrid energy demand forecasting models often face challenges related to inefficient model training and increased sensitivity to outliers, which lead to increased operating costs and potential disruptions in energy distribution. Recent improvements, particularly the improved whale optimization algorithm, have shown a 7% increase in performance. Their hybrid model showed promising improvements, indicating a positive direction for future research. For the further development of these models, researchers observed that it is necessary to focus on improving their model robustness and adaptability in dynamic environments (Yang Y. T., 2022). However, limitations such as low model training efficiency and sensitivity to outliers remain significant obstacles to be addressed.

The hybrid robust system developed by Yang et al. (2022) addressed outliers in electrical load forecasting, but had some limitations. A key weakness lies in the deterministic nature of the model, as it aims to increase accuracy by reducing the impact of outliers without explicitly accounting for uncertainty in predictions. This means that while the model performs well under typical conditions, it may struggle in highly volatile scenarios or when encountering unseen anomalies where uncertainty quantification could play a key role.

Incorporating an uncertainty estimate into the model could significantly improve its performance. Uncertainty-aware models can provide not only point forecasts, but also probability intervals that offer a clearer understanding of the reliability of the forecasts. This is essential in electricity load forecasting, where unforeseen events or changes (e.g. extreme weather conditions or demand peaks) can significantly affect network operation. Methods such as Bayesian Neural Networks (BNN) or Monte Carlo Dropout could be integrated into a hybrid framework providing predictive intervals that help system operators prepare for a range of possible load outcomes rather than a single deterministic forecast (Yang Y. T., 2022).Quantifying uncertainty can enable better risk management. In network operation, this enables more robust decision-making and ensures that even in cases of low confidence in forecasts, operators can adjust strategies (e.g., improving the hybrid model using uncertainty estimation techniques would increase its applicability and reliability in real-world, high-stakes scenarios, thus improving the overall resilience of energy systems.

Bo, Nie, and Wang (2020) emphasize that accurate energy load forecasting is essential to increase grid efficiency and reduce energy costs. They blame the existing models for the inability to effectively integrate individual prognostic methods and the neglect of the importance of data pre-processing, which ultimately leads to a decrease in accuracy. In response, their study on a hybrid model combines singular spectral analysis (SSA) preprocessing with multiple algorithms, including Backpropagation (BP), Support Vector Machines (SVM), General Regression Neural Networks (GRNN), and ARIMA, along with a weighting approach. Increase the accuracy of predictions. Validation using Australian performance data demonstrated that this innovative approach significantly outperformed individual models (Bo, 2020).Despite its robustness, the hybrid model has significant weaknesses, notably the inability to account for uncertainties arising from input disturbances. This oversight is critical in situations where energy demand experiences significant fluctuations due to factors such as weather changes, economic changes, or unforeseen events. Without addressing these uncertainties, even a highly accurate model risks underperforming when confronted with volatile input data.

Research conducted by Eugene (2023) suggests that to mitigate data volatility shortcomings, a viable solution is to integrate uncertainty estimation into the hybrid model. Techniques such as probabilistic forecasting or Bayesian Neural Networks (BNNs) could be used to predict a range of outcomes accompanied by confidence intervals. This improvement would allow network operators to understand not only the most likely demand scenarios, but also the spectrum of potential variations. By addressing the uncertainty in the face of input disturbances, the prediction model would provide a more robust and accurate solution that would better suit the dynamic conditions of the real world (Eugene, 2023).

Fan (2023) conducted a study that integrated multiple methods, including Variational Mode Decomposition (VMD), Convolutional Neural Networks (CNN), and Bi-Directional Long-Short-Term Memory (BLSTM), which made the model overly complex. This complexity hindered model interpretability and required significant computational resources for both training and implementation. To mitigate these issues, the inclusion of uncertainty estimation techniques could streamline the model by identifying the most influential components and reducing unnecessary complexity. This approach would not only improve interpretability but also optimize resource allocation, making the model more practical for real-world applications (Fan, 2023).

The work of T. Zhang et al. researches the complexities of long-term energy demand and energy peak forecasting, addressing several critical research gaps. Traditionally, these two aspects have been predicted separately, leading to increased complexity. To streamline this process, the paper proposes a unified Sequential eXtreme Gradient Boosting (XGBoost) model. The model acknowledges the intricate nature of power consumption profiles and employs a sequential multi-input/multi-output approach that considers multiple influencing factors to enhance accuracy. To further enhance the accuracy of peak energy consumption prediction, the model employs a hybrid direct-recursive sequential configuration that utilizes predicted energy consumption as input. The effectiveness of the model is validated using data from the New South Wales power grid and compared to state-of-the-art models that demonstrated superior performance (Tingze, 2023). Their model appears promising; however, potential research gaps remain, such as scalability to larger datasets, real-time forecasting capabilities, integration with renewable energy sources, comprehensive economic and environmental impact assessments, and handling of uncertainties.

Extreme Gradient Boosting (XGBoost) is a machine learning algorithm widely used for regression, classification, and ranking tasks. It is based on the gradient boosting framework and includes advanced features to increase performance and efficiency (Kumari P. &., 2021). It has (i) A Framework for transition amplification. Transition boosting involves sequentially building a series of decision trees, where each tree corrects the mistakes of the previous one. Trees are added to minimize a specified loss function using the gradient descent method. (ii) Extreme transition gain, XGBoost optimizes gradient boosting by incorporating regularization, better handling of missing values, and efficient computation. It includes techniques such as sparsity-aware algorithms, parallel tree construction, and weighted quantile rendering, which enable high accuracy and speed. The goal of XGBoost is to minimize an objective function that combines a loss function and a regularization term (Tingze, 2023):

$$L(\emptyset) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{k} \Omega(f_k) \quad (1)$$

Where;

$L(\emptyset)$: Total model loss.

$l(y_i, \hat{y}_i)$: Loss function (e.g, root mean squared error for regression).

$\Omega(f_k)$: A control term for the complexity of each tree.

Regulatory Term

The regulation term controls overfitting and is defined as

$$\Omega(f) = r^T + \frac{\lambda}{2} \sum_{j=1}^{T} w_j^2 \quad \text{...} \quad (2)$$

Where;

$$\vartheta_i = \frac{\partial l\left(y_i, \hat{y}_i^{(t-1)}\right)}{\partial \hat{y}^{(t-1)}} \text{and}$$

$$h_i = \frac{\partial_i^2 l\left(y_i, \hat{y}_i^{(t-1)}\right)}{\partial\left(\hat{y}_i^{(t-1)}\right)^2}$$

ϑ_i -First order derivative (gradient): the direction of improvement

h_i -Second order derivative (Hessian) Curvature of the loss

Tree splitting criterion

To grow the tree, the algorithm selects the best split based on the gain.

$$Gain = \frac{1}{2}\left(\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} + \frac{(G_L + G_R)^2}{H_L + H_R + \lambda}\right) - \gamma \quad \text{...} \quad (5)$$

Where;

G_L, H_L: Sum of gradients and Hessians for the left child

G_R, H_R: Sum of gradients and Hessians for the right child

λ : L2 Regularization

γ : Complexity penalty

A split is accepted if Gain > 0. Meaning it improves performance. Some of the advantages of XGBoost include the regulation of L1 and L2 norms that exceed limits. It handles missing data efficiently. It is optimized for parallel and distributed systems. It allows the use of custom-defined loss functions.

XGBoost is highly efficient and has been applied to tasks such as Regression problems (e.g., housing price prediction). Classification tasks (e.g., spam detection). Ranking issues (e.g., search engine ranking).

A study by Fan et al. (2022), titled "Application of Multivariate Surface Random Forests for Short-Term Load Forecasting", published in the International Journal of Electrical Power & Energy Systems, investigated the effectiveness of Random Forest (RF) models for short-term load forecasting. Findings reveal that Random Forest (RF) models significantly improve prediction accuracy compared to traditional methods, utilizing a multivariate response surface approach to capture complex relationships between multiple input variables and loads. Research underscores the importance of feature selection, as focusing on relevant features leads to improved model performance.

In addition, the RF model exhibits robustness to outliers and noise, indicating strong generalizability across different datasets and suggesting its applicability in various scenarios. The study's implications suggest that RF models can help power utilities and companies optimize the management of their energy resources. The researcher recommended that future research avenues include simplifying model complexity, exploring automated feature engineering, developing hybrid models that integrate RF with other techniques, enhancing real-time prediction capabilities, and evaluating model generalization across various domains and conditions (Guo F. F., 2022).

The Random Forest algorithm is a powerful machine learning technique that uses ensemble learning to increase prediction accuracy. It constructs several decision trees, each trained on a random subset of the data, and combines their predictions. This approach reduces overfitting and improves generalizability. Key features include robustness to noise, the ability to process high-dimensional data, and the ability to judge feature importance (Guo F. F., 2022). The algorithm involves creating bootstrap samples, creating decision trees with random selection of features, and aggregating the predictions of these trees. The final output is either a majority class for classification or an average prediction for regression.

A study titled "Optimism-Based Adaptive Control of Linear-Quadratic Systems" focused on the development of an adaptive control framework for linear-quadratic systems (Faradonbeh, 2020). The goal of the study was to design an adaptive controller that could regulate the system's dynamics and simultaneously optimize the quadratic performance criterion. The proposed controller employed an optimistic-based approach that incorporated a model-free learning algorithm to estimate the system dynamics and adaptively update the controller parameters. The authors demonstrated the effectiveness of their approach through theoretical analysis and numerical simulations. They compared the performance of the optimism-based adaptive controller with that of other existing control methods. They demonstrated that it achieves better tracking performance and robustness in the presence of uncertainties. Further research is needed to analyze the robustness of systems or models to input perturbations (Faradonbeh, 2020).

A study by Fang et al. (2022) investigated a Learning to Learn Transferable Attack (LLTA) method for generating adversarial perturbations that can be applied to a victim model by learning from both data and model augmentation. The authors noted that existing adversarial attack transfer methods had limited portability because adversarial perturbations tend to overlap with a single surrogate model and specific data patterns. To address this limitation, the LLTA method incorporated random resizing and padding for data augmentation, as well as random backpropagation for model augmentation, which helped mitigate the influence on model prediction. Their study demonstrates the effectiveness of the LLTA method on the used datasets. However, it will be valuable to further evaluate the method on a more diverse set of datasets and models to confirm its generalization capabilities (Fang, 2022).

A study was done on “A controller synthesis method to achieve independent reference tracking performance and disturbance rejection performance.” This study aimed to design a controller that achieves independent tracking of reference values and suppresses interference: mathematical Modeling, Control Synthesis, Optimization, and Verification. Coupling Decoupling Strategies, Optimization Techniques, and robustness analysis were used to refine the controller design. The study did not do a comparison of their method with existing methods in the literature to help analyze model performance (Shi, 2022).

Electricity demand forecasting is a critical task in the power sector, as accurateorecasts are essential for the efficient and reliable operation of the electricity grid (Wang B. Z., 2019). Various approaches have been proposed to forecast electricity demand, ranging from Statistical Methods to Machine Learning Techniques. One of the most widely used approaches to electricity demand forecasting is Time Series Analysis, which involves modeling historical demand data and using it to make future forecasts (Zhao, 2017). Autoregressive Integrated Moving Average (ARIMA) models are commonly used for this purpose, as they can capture trends, seasonality, and cyclical patterns in data. However, ARIMA models have limitations in handling non-linear relationships and cannot include external factors such as weather data (Pokou, 2024).

To overcome the limitations of ARIMA models, Machine Learning (ML) techniques such as Artificial Neural Networks (ANN), Support Vector Regression (SVR), and Random Forest (RF) have been applied to electricity demand forecasting (Wang B. Z., 2019). These ML techniques captured complex relationships between input variables and output demand;however,theydid not include external factors such as weather data and economic indicators. It was seen that this model is prone to overfitting, where it performs well on training data but poorly on unseen data; hence, there is a need for enhanced model validation.

A study by Hong T. P. (2016) stated that forecasting electricity demand presents several challenges. A significant challenge is the uncertainty and variability of the input data. Factors such as weather conditions and economic indicators, which are key to forecasting demand, often fluctuate unpredictably, making it difficult to make accurate forecasts. Additionally, there is the challenge of insufficient training data, particularly pronounced in developing countries where data collection efforts may be limited (Ahmad T. Z., 2020). In addition, changes in consumer behavior, such as changes in energy use patterns and the increasing integration of renewables into the grid, also add to the complexity and potential inaccuracies of electricity demand forecasts (Hong T. P., 2016), (Mir, 2020).

Forecasting electricity demand is a crucial aspect of energy management and planning (Liu, J.. Current approaches primarily rely on Machine Learning Techniques, such as Neural Networks and Regression models, to produce accurate predictions (Hong T. P., 2020). However, these methods often overlook uncertainties associated with input perturbations, leading to inaccuracies and unreliability in Demand Forecasts (Smith A. B., 2021).One major challenge lies in the absence of a comprehensive framework for quantifying uncertainty in machine learning-based electricity demand prediction (Wang & Zhang, 2018). Load forecasting models often suffer from uncertainty, resulting in inaccurate and unreliable predictions. Therefore, there is an urgent need to develop more robust forecasting models that account for uncertainties in input variables (Chen, 2020). This study proposes to design an innovative hybrid model that incorporates uncertainty estimation during model training.

A researcher conducted a study using a Bayesian approach for uncertainty estimation, which improved the robustness of neural networks against adversarial attacks. (Jeon, 2022). The study demonstrated the potential of the Bayesian approach for Uncertainty Estimation to improve the robustness of neural networks against adversarial attacks. A weakness of this study is that it did not evaluate the method in real-world applications, making it unclear how effective it would be in practice. Additionally, the study did not examine the computational cost of using the Bayesian approach, which can be a limiting factor in certain applications. Our study will evaluate the accuracy of the proposed model in forecasting future demand and compare it with the actual selected future demand.

Rashid's (2022) study on Adversarial Training for Intrusion Detection Systemsoncluded that Adversarial training could be used to estimate model uncertainty by measuring the model's sensitivity to input perturbations (Rashid, 2022). The study showed that using adversarial training enhanced the robustness of Deep Learning models by improving Uncertainty Estimation. The weakness of his study is that it did not evaluate the proposed method against other state-of-the-art techniques for improving the robustness of deep learning models. The study only focused on the effect of adversarial attacks on intrusion detection systems in the Internet of Things (IoT), and it is unclear how effective the proposed method would be against other types, such as input perturbations, in our proposed study.

The Stein Variational Gradient Descent algorithm was used in a study by Liu and Wang in their paper, "Stein Variational Gradient Descent." Stein Variational Gradient Descent

(SVGD) is a technique for sampling from probability distributions using the gradient descent method. SVGD was used to estimate uncertainty by generating samples from a posterior distribution over model parameters, which was used to compute the variance or entropy of the predictions. (Liu X. T., 2021) . The research outcome showed that SVGD outperformed other probabilistic optimization methods in terms of optimization efficiency and accuracy. It was found to be particularly effective for high-dimensional and multimodal distributions, where other methods may struggle to find the global optimum. The study also showed that SVGD is relatively robust to different initialization strategies and hyperparameters, making it a versatile optimization tool.

However, the study also identified the main limitation as the method's sensitivity to the choice of kernel function, which significantly impacted the algorithm's performance. (Liu X. T., 2021)The method was sensitive to the choice of optimization parameters and required careful tuning to achieve optimal performance.

To address these gaps and limitations, the study recommended several potential solutions, including the use of adaptive kernel functions, the development of more efficient algorithms for large datasets, and the improvement of optimization parameter tuning strategies. The study also recommended further research to be explored on the limitations and potential extensions of SVGD, particularly in the context of more complex optimization problems and real-world applications(Liu X. T., 2021).

A research study was done on Deep Kernel Learning (DKL). DKL involves combining Neural Networks with kernel methods to learn complex nonlinear relationships in the data. Deep kernel learning was used to estimate uncertainty by treating the kernel function as a measure of similarity between data points. It was used to compute the variance or entropy of the predictions (Yu, 2020). The weakness of the study is that it did not evaluate the proposed method against other state-of-the-art techniques for improving the robustness of deep learning models. The study focused solely on a specific type of kernel function, and it remains unclear how effective the study method would be in addressing the model diffusion challenges.

A research study was done on Conformal Prediction. Conformal prediction is a framework for constructing prediction regions that provably contain the true prediction with a specified probability. Conformal prediction estimated uncertainty by constructing prediction regions for individual predictions. It was used in conjunction with Ensemble Methods to estimate uncertainty across multiple predictions. The study evaluated the performance of multi-split conformal prediction, a variant of conformal prediction that uses multiple training sets to construct prediction regions (Solari, 2022). The weakness is that the performance of multi-split conformal prediction was heavily dependent on the number and distribution of training sets used to construct the prediction regions. Inadequate or biased sampling leads to incorrect uncertainty estimates and reduced prediction accuracy.

The study by Zhu et al. (2022) examined the uncertainty and sensitivity analysis of cooling and heating loads in building energy planning. The authors found that their model could provide estimates of uncertainty, which would help building managers in making informed decisions about energy management and efficiency. The study provided valuable insights into the robustness of the proposed techniques; however, it only focused on input perturbations and did not consider other sources of uncertainty, such as model parameters and modeling assumptions (Zhu, 2022). It was recommended that future research investigate the impact of these sources of uncertainty on the proposed techniques.

The study by Cibin et al. (2023) investigated the Machine Learning-Based (MLA) approach for estimating the thermal dynamics of residential buildings and explored the potential for enhancing energy flexibility and efficiency. In addition, the authors investigated the impact of input perturbations on the accuracy of their models, a crucial consideration for deep learning models in real-world settings.

The authors used a dataset of residential buildings to train Deep Learning models to estimate the thermal resistance and heat capacity of the buildings. They then perturbed the input variables, such as outdoor temperature and energy consumption, to assess the models' sensitivity to these changes. However, the authors did not provide detailed information on the dataset used to train and test the models, which limits the generalizability of the findings (Cibin, 2023). The authors did not provide information on the practical implementation of their models or the potential barriers to adoption.

A study was conducted to investigate the application of Artificial Neural Networks (ANNs) for Electricity Load Forecasting in Nigeria. The study employed a Multi-Layer Perceptron (MLP) Neural Network to predict electricity demand and evaluated the model's performance under various scenarios, including changes in temperature and humidity. The study found that the ANN model was able to provide accurate predictions under different input perturbations; however, it idid not explicitly address Uncertainty Estimation challenges (Ozigis, 2021).

The study by Mosavi (2019) on energy consumption prediction found that using deep learning made a valuable contribution to the field of energy forecasting. However, when considering the context of Uncertainty Estimation in Deep Learning models under input perturbations, the study did not analyze the effect of input perturbations on the performance of Deep Learning models. Future studies could explore the impact of input perturbations on the generalizability and accuracy of energy consumption prediction models (Mosavi, 2019).

The study by Bhamidipati et al. (2021) demonstrated the importance of local value capture in the context of the energy transition, particularly in developing countries where the benefits of renewable energy development may not always accrue locally. The study did not directly address Uncertainty Estimation or input perturbations in deep learning models, where principles of local value capture could help to build more resilient and sustainable energy systems that are better able to address uncertainty and adapt to changing conditions (Bhamidipati, 2021).

A Research study by Wang (2020) focused on energy management in off-grid rural electrification projects. The study proposed a deep learning approach for predicting energy demand and optimizing energy supply in these projects. While the study did not explicitly address Uncertainty Estimation, it suggested that input perturbations could have a significant impact on the performance of deep learning models for energy management (Wang, 2020)

The study by Hong (2022) shows that Electricity Demand Forecasting is a critical component of efficient energy planning, operations, and decision-making. Deep learning models, such as Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs), have demonstrated remarkable performance in capturing complex patterns and relationships in electricity demand data (Hong, 2021). However, these models often overlook the uncertainties associated with input perturbations, which can lead to potential inaccuracies and inadequate risk assessment in demand forecasts.

Uncertainty Estimation plays a vital role in Electricity Demand Forecasting as it provides a measure of confidence or variability in the forecasted demand values (Hong, 2021). Traditional approaches to Uncertainty Estimation, such as statistical methods or regression models, struggle to capture the complex dynamics and nonlinear relationships present in electricity demand data.

Understanding the effect of input perturbations on Uncertainty Estimation is crucial for enhancing the robustness and reliability of deep learning models in the face of changing conditions and unforeseen events (Qi, 2020). By systematically perturbing input variables such as temperature, time of day, day of the week, holidays, and other relevant factors, researchers can assess the resulting changes in forecasted demand and corresponding Uncertainty Estimates.

Investigating input perturbations on Uncertainty Estimation can lead to the development of a hybrid model that combines deep learning techniques with improved methods for uncertainty quantification. This hybrid model would enable more accurate risk assessment and decision-making processes for energy grid operators and stakeholders (Yu Z. W., 2022). By capturing the uncertainties associated with different input perturbations, the hybrid model would provide valuable insights into the reliability and robustness of electricity demand forecasts under various conditions.

The integration of deep learning models with Uncertainty Estimation techniques and input perturbations has the potential to advance the field of electricity demand forecasting significantly. This research aims to bridge the existing gap in the literature by developing a hybrid model that investigates the impact of input perturbations on Uncertainty Estimation in deep learning models. The findings from this study will contribute to the development of more accurate and reliable forecasting models, enhancing the efficiency and stability of energy grids and supporting effective decision-making in the energy sector (Zhang Y. K., 2024).

Developing robust techniques for Uncertainty Estimation in deep learning models under input perturbations will be essential for enhancing reliability and robustness in real-world scenarios. The existing approaches, including Bayesian Neural Networks, Deep Ensembles, and Monte Carlo Dropout, have shown promising results and are likely to be further developed in the future (Lin, et al., 2021). Deep learning models have shown promise in electricity demand forecasting, but there is a need to understand the effect of input perturbations on Uncertainty Estimation within these models. The existing literature overlooks this aspect, leading to potential inaccuracies and inadequate risk assessment. Hence, there is a need to comprehensively investigate the impact of different types and magnitudes of input perturbations. This will enhance the robustness of Uncertainty Estimation in deep learning models for electricity demand forecasting (Gu, 2023).

By combining deep learning techniques with uncertainty quantification methods, the proposed hybrid model aims to provide more enhanced, robust, and reliable forecasts in the presence of input perturbations. The model will systematically perturb input variables such as temperature, time of day, day of the week, holidays, and other relevant factors to assess the resulting changes in forecasted demand and corresponding uncertainty estimates. This investigation will yield valuable insights into the robustness and reliability of deep learning models under varying conditions and perturbations (Zhao Gu, 2023).

Several studies have demonstrated the effectiveness of these techniques in improving the robustness of deep learning models. For instance, a recent study on deep ensembles has improved uncertainty estimation in image classification tasks. (Ren, 2022). The study demonstrated how deep ensembles improved Uncertainty Estimation in image classification tasks, especially in identifying out-of-distribution images. However, the study did not compare the proposed method with other state-of-the-art Uncertainty Estimation techniques. Additionally, the study focused solely on image classification tasks, and it remains unclear whether deep ensembles would be equally effective in other deep learning tasks.

Furthermore, the development of this hybrid model will enable the exploration of novel techniques and approaches for capturing the uncertainties associated with different input perturbations. By incorporating Uncertainty Estimation as an integral part of the forecasting process, the hybrid model will facilitate improved risk management strategies and decision-making processes for energy grid operators and stakeholders.

Despite the remarkable performance of deep learning models, such as recurrent neural networks (RNNs) and convolutional neural networks (CNNs), in electricity demand forecasting, the uncertainties associated with input perturbations have been largely overlooked, leading to potential inaccuracies and inadequate risk assessment in demand forecasts. While Uncertainty Estimation is crucial for providing reliable forecasts, traditional approaches often struggle to capture the complexities present in electricity demand data (Zhao Gu, 2023). As a result, there is a significant need to research and gain an understanding of the impact of input perturbations on Uncertainty Estimation in deep learning models for electricity demand forecasting.

A growing body of literature investigates the effect of input perturbations on the accuracy and reliability of Uncertainty Estimates produced by existing machine learning methods. In particular, this study will extensively investigate the research question of how different types and levels of input perturbations affect uncertainty estimates. The inherent characteristics of video data encompass not only spatial alterations but also temporal changes, as perturbations applied to videos. This leads to an expansion of the search space, escalating the demands on processing time and computational resources required to pinpoint the temporal segment of a sequence that has the most significant influence on the output of the Deep Neural Network (DNN) classifier (Li et al.(2020) Li, 2020).

The majority of previously published research has utilized backpropagation-based techniques. However, more recently, perturbation-based methodologies have been implemented for video data, building upon the concept of meaningful perturbation introduced by researchers (Fong, 2017). These methods extend the external perturbation approach from two dimensions (images) to three dimensions (videos). In contrast to the substantial advancements in image attribution research, the number of studies focusing on attribution techniques for videos remains limited. The predominant direction of research is centered on adapting existing image attribution methods to suit video inputs.

French (2019) did a study on the importance of strong, High-Dimensional Perturbations in Semi-Supervised Semantic Segmentation. Researchers have been investigating how applying input perturbations can improve the performance of segmentation models when labeled data is limited. The study included experiments using different levels of perturbations on a Supervised Segmentation task. Researchers employed techniques such as data augmentation to generate additional training samples with varying characteristics. By introducing strong perturbations, they focused on evaluating the model's ability to generalize and accurately segment objects in images.

The study's findings suggest that strong, high-dimensional perturbations benefited supervised semantic segmentation in several ways. These perturbations enhanced model feature extraction capabilities, facilitated the exploration of challenging regions in the input space, and reduced switching by providing a more diverse training dataset. The study focuses on the impact of strong perturbations on semi-supervised semantic segmentation. However, there is a research gap in evaluating the generalization of models trained with strong perturbations to different datasets with different characteristics (French, 2019).

A similar study approach was utilized in RISE, where input images were occluded with random occlusion patterns generated by sampling small (7x7 pixels) binary masks and subsequently magnifying them to a higher resolution using bilinear interpolation. The main limitation observed in both the Occlusion and RISE techniques was their failure to consider the morphology of objects within the image, leading to the production of only approximate outcomes. LIME employed occlusions of superpixels and approximated networks using linear models. Due to the coarse nature of super pixels, the generated saliency maps exhibited relatively low precision. In order to enhance spatial accuracy, the meaningful perturbations method introduced an optimization process for shaping perturbation masks. The objective was to minimize the blurring effect on the input image while still maximizing the reduction in class score (Petsiuk, 2018).

A study was conducted to understand uncertainty estimation in Semantic Segmentation (Garg, 2020). The researchers focused on addressing the challenges associated with Uncertainty Estimation in Semantic Segmentation tasks. Semantic segmentation involves assigning a class label to each pixel in an image, and Uncertainty Estimation plays a crucial role in quantifying the confidence or reliability of the predicted segmentation. One of the gaps in Semantic Segmentation is the lack of accurate and reliable Uncertainty Estimation methods.

A researcher studied"Understanding and Improving Uncertainty Estimation in Semantic Segmentation." The focus was on the crucial task of Uncertainty Estimation in semantic segmentation (Hoyer, 2021). The researchers noted that training semantic segmentation models with deep networks necessitates a substantial quantity of labeled training data, posing a significant practical challenge due to the labor-intensive nature of manually labeling segmentation masks.

A study was conducted on a novel framework for Convolutional Spectral Kernel Learning, which incorporated spectral graph theory into the design of CNNs. By exploiting the spectral properties of data, their approach aimed to enhance the network's ability to capture intrinsic patterns and improve generalization performance (Wang J. L., 2022). The experimental evaluations conducted by Wang J. L. (2022) on various benchmark datasets supported the efficacy of their approach. They compared the performance of their Convolutional Spectral Kernel Learning method with standard CNNs and other state-of-the-art methods. The results demonstrated improved generalization capabilities, showing the potential of spectral methods in enhancing the performance of CNNs.

A research study done on Uncertainty Estimation for Deep Regression via Perturbation-Based Ensembles (Chandak, 2023). In this work, the authors presented PEG (Perturbed Ensemble via Gradient Updates), a simple and efficient approach that constructs deep ensembles by utilizing gradients computed over validation data. Experiments demonstrated that PEG was capable of generating an ensemble using a pre-trained model. The research study acknowledged the potential for improvement in terms of scalability and computational efficiency. As Deep Learning models continue to grow in complexity and size, it becomes crucial to optimize the PEG approach to handle larger datasets and reduce the computational burden without compromising its performance (Chandak, 2023).

The study focuses on Robust and explainable fault diagnosis using power-perturbation-based decision boundary analysis of deep learning models. It focused on addressing the challenges associated with fault diagnosis in deep learning models. The researchers proposed a method that combined power perturbation and decision boundary analysis to achieve robust and explainable fault diagnosis (Gwak, 2022). The study recommended the need for further research and optimization to ensure that the power perturbation and decision boundary analysis techniques can efficiently handle the computational demands and complexities associated with real-world, high-dimensional fault diagnosis scenarios.

These recent papers demonstrate the ongoing interest and research in investigating the impact of input perturbations on Uncertainty Estimation in deep learning models, as well as the development of new methods to improve the robustness and accuracy of Uncertainty Estimation. One study investigated the impact of input noise on the uncertainty estimates produced by Deep Learning Models (Lakshminarayanan, 2017). The study found that adding Gaussian noise to the input images improved the quality of the uncertainty estimates, especially for Out-Of-Distribution Samples. However, the study also found that the effect of noise on Uncertainty Estimation is highly dependent on the specific architecture and training procedure of the Deep Learning Model.

A research study done by Zhang S. C.(2021) investigated the impact of adversarial attacks on the Uncertainty Estimation produced by deep learning models. The study found that adversarial attacks can significantly reduce the reliability of uncertainty estimates, especially for models with low robustness to adversarial examples. The study also recommended that future studies design a method to improve the robustness of uncertainty estimates by augmenting the training data with adversarial examples. In addition to input noise and adversarial attacks, other types of input perturbations have also been studied in the context of Uncertainty Estimation. For example, a researcher investigated the effect of input perturbations on the uncertainty estimates produced by Bayesian Neural Networks (DeVries, 2018). The study found that corrupted inputs led to overconfident predictions and underestimation of uncertainty, especially for models with small training data sizes. However, the authors did not show how transferability challenges were handled in their study.

A research study focused on a Novel Framework for Performance-Based Seismic Assessments of Bridges Based on Uncertainty Estimates produced by Deep Learning models (Sajedi, 2022). The authors evaluated the robustness of their method under different levels of input noise. However, they did not explicitly investigate how different types and levels of input perturbations affect the accuracy and reliability of uncertainty estimates. The authors conducted a sensitivity analysis to evaluate the robustness of their method under different levels of input noise. They added different levels of white noise to the input signals and evaluated the performance of their method using several metrics, including mean absolute error (MAE) and coefficient of determination (R^2). The results showed that the proposed method was robust to different levels of input noise and outperformed other traditional machine learning methods, such as support vector regression (SVR) and random forest (RF) (Sajedi, 2022).

The study did not investigate the effect of other types of input perturbations, such as adversarial attacks or label noise; the analysis conducted by the authors provided insights into the potential robustness of the proposed method to input perturbations. Their method utilized filter banks and a hybrid deep learning architecture that combined a Convolutional Neural Network (CNN) and a Long Short-Term Memory (LSTM) network (Sajedi, 2022).

2.2.1 Weaknesses of Existing Power Demand Forecasting Models with Uncertainty Estimation Under Input Perturbations for Accurate Forecasting Models

Research question one required a review of the weaknesses of existing models in enhancing the accuracy and confidence in the adoption of forecasting results for key decision-making in the real world. This was achieved through a meta-analysis of the literature to identify the weaknesses of existing models. This research design provides a comprehensive, critical, and objective analysis of current research studies. Secondary data on the weaknesses in the performance of hybrid power demand forecasting models is presented in this thesis.

2.2.2 Methodology for Identification of the Weaknesses of Existing Deep Learning Forecasting Models

A meta-analysis literature review was carried out to identify the weaknesses of existing models in ensuring the accuracy of forecasting models. A meta-analysis literature review is a systematic approach that quantitatively synthesizes research findings from multiple studies to draw overall conclusions about a particular topic (Oblak, 2021). The research question was how the weaknesses of existing Power Demand Forecasting Models with Uncertainty Estimation under Input Perturbations could be assessed.

Clearly defined inclusion and exclusion criteria guided the selection of literature for this study to ensure rigor, relevance, and alignment with the research objectives. In line with recommended evidence-based review practices (Cooper, 2024), the review included studies published between 2020 and 2024, reflecting the most recent advancements in machine learning-driven power forecasting within the rapidly evolving field of energy analytics. Peer-reviewed journal articles and reputable conference publications were prioritized to maintain academic credibility and methodological quality. Literature was selected if it explicitly addressed power or energy demand forecasting using machine learning, deep learning, or hybrid modelling approaches and provided empirical validation through real-world datasets or experimental simulations. Studies that investigated uncertainty estimation, model robustness, or forecasting performance under data variability or perturbations were considered essential for inclusion, as they directly aligned with the core focus of this research.

Conversely, studies published prior to 2019 were excluded, as they did not adequately reflect contemporary methodological developments in artificial intelligence and hybrid forecasting models. Purely theoretical papers lacking empirical validation were omitted to uphold the study's emphasis on practical applicability and evidence-driven insights. Literature outside the energy or power systems domain, or that did not involve forecasting, was excluded due to misalignment with the scope of this research. In addition, non-peer-reviewed materials, sources lacking sufficient methodological transparency, and papers with unavailable full-text access were excluded to ensure the reliability and traceability of the reviewed evidence (David, 2025). Collectively, these criteria ensured that the reviewed literature was current, methodologically sound, contextually relevant, and contributed meaningfully to understanding hybrid power demand forecasting and uncertainty estimation under input perturbations. Thefindings from the literature were;

2.2.3 Weakness 1: Data Complexity and Sensitivity

Hybrid models often employ complex algorithms, such as artificial neural networks (ANNs), combined with optimization techniques like particle swarm optimization (PSO) (Anand, 2020). This complexity was seen as leading to overfitting, where the model performs well on the training data but fails to generalize effectively to new, unseen data. These models are susceptible to the quality of the input data; any variation in the data can significantly affect forecast accuracy, especially in models that incorporate seasonal factors and error correction (Tian &, 2021). Literature showed that many hybrid models do not effectively manage the complex non-linear interactions between electricity demand and influencing factors, resulting in reduced accuracy (Huang A. B., 2024).

2.2.4 Weakness 2: Lack of Flexibility

Hybrid models, such as those developed by the National Renewable Energy Laboratory (NREL), have significantly contributed to the advancement of renewable energy technologies through tools like the System Advisor Model. However, they have been criticized for their rigidity in system configuration and component distribution, which limits their applicability in a variety of hybrid power systems (Manwell, 2023).

2.2.5 Weakness 3: Performance Variability

Empirical studies demonstrate that hybrid models can outperform traditional methods; however, their performance tends to fluctuate depending on the specific context and data used, resulting in inconsistent results (Guo F. M., 2024). There is also limited consideration of factors influencing demand forecasting. Despite these challenges, hybrid models remain a promising area of study, with ongoing research aimed at enhancing their robustness and adaptability in various forecast scenarios.

2.2.6 Weakness 4: Risks of Overestimation

The research studies highlight that the Base model may overestimate the demand for products, especially if the selling period is near the end of the growth cycle. This overestimation can confound operational decisions, suggesting that the model may not always provide reliable forecasts during peak demand seasons (Yin, 2020).

2.2.7 Weakness 5: Inadequate Temporal Dynamics

Some existing models cannot capture the long-term dependencies and temporal dynamics necessary for reliable power demand forecasting (Cao, 2023). There is also a limitation due to the lack of historical data for forecasting new product demand.

2.2.8 Weakness 6: Model Complexity

Another weakness was the design of an overly complex model, which led to reduced model interpretability and required extensive computing resources for training and deployment (Fan, 2023). Incorporating uncertainty estimation techniques helped simplify the model by identifying the components that have the greatest impact, thereby reducing unnecessary complexity.

Table 1

Summary of Weaknesses of Existing Models

No.	References	Year	Weaknesses	Risk
1	(Huang A. B., 2024).	2024	Data complexity and sensitivity	Models fail to generalize effectively to new, unseen data affects forecast accuracy
2	(Manwell, 2023)	2023	Lack of flexibility	Low forecast accuracy
3	(Guo F. M., 2024)	2024	Performance variability	Leads to inconsistent forecast results Unreliable forecasts Unstable power supplies Increased risks of power outage. Investment uncertainty Suboptimal use of predictive maintenance.
4	(Yin, 2020)	2020	Risks of overestimation	Inaccurate forecast Wasted resources Power instability Financial losses Missed opportunities for flexibility and demand response.
5	(Cao, 2023)	2023	Inadequate temporal dynamics	Leads to inaccurate forecasts Inefficient power grid operation Financial risks
6	(Fan, 2023)	2023	Model complexity	Integration of multiple algorithms cause increased model complexity
7	(Fan, 2023)	2023	limited comparisons with other forecasting models	-Leads to low model results reliability

2.2.10 Design Recommendations to Address the Identified Weaknesses of Existing Power Demand Forecasting Models

The findings of the literature review highlighted several weaknesses in the existing power demand forecast models, particularly in terms of model rigidity, sensitivity to noise input, overfitting, limited adaptability, and inadequate handling of uncertainty under input perturbations. To resolve these challenges, key recommendations have been identified. These recommendations informed and guided the design of the proposed hybrid forecast models with uncertainty estimation, ensuring that they capitalize on the limitations of prior approaches, thereby improving model performance and reliability.

First, the literature suggests the use of a modular hybrid architecture that combines multiple base models, including Artificial Neural Networks (ANNs), Recurrent Neural Networks (RNNs), Support Vector Regression (SVR), XGBoost, Random Forest, LightGBM, and CatBoost. Such a modular setup promotes flexibility and adaptability, allowing components to be modified or replaced as data characteristics or forecasting requirements. This directly addresses the rigidity of a less adaptable single model framework.

Second, it is recommended to integrate uncertainty estimation techniques to increase flexibility in the face of data disturbances. Embedding these methods, especially within the RNN framework, enables the model to determine forecast variability and dynamically adjust predictions. This would improve consistency and reliability when working under variable or noisy conditions.

Third, to mitigate the risks of overestimation, the literature emphasizes the importance of using uncertainty thresholding. By leveraging the potential capacity of RNNs, this approach enables the model to detect overly optimistic forecasts resulting from input perturbations.

Fourth, literature recommends hybrid learning strategies to capture temporary mobility more effectively. ANNS modeling complexes are favorable for nonlinear dependencies, while SVR regression-based patterns recognition. Combining these Approaches enhances the ability to capture both global trends and localized demand patterns, thus improving forecast accuracy.

Fifth, the study emphasizes the role of regularization and model simplification in improving interpretability and reducing computational costs. Within the ANN and RNN framework, trimming less important features or neurons helps manage complications, reduces overfitting, and yields a more efficient yet accurate system.

Sixth, the importance of benchmarking and continuous refinement continues to emerge in the literature. Regularly identifies areas for improvement and promotes continuous growth to ensure competitiveness of hybrid models with both traditional statistical methods (e.g., ARIMA, ETS) and advanced machine learning approaches.

Seventh, complementing the integration of models is recommended to strategically address the strengths and weaknesses of the balance. For example, while ANS involves non-lectured interactions, simple models such as the K-Nearest Neighbors (K-NN) are effective for detecting local trends. Additionally, probabilistic methods such as Gaussian processes enrich uncertainty, equipping the hybrid framework with probabilistic forecasts that better support decision-making under uncertain conditions (Zeng, 2023).

Ultimately, the literature emphasizes the importance of context-specific deployment of model components. For example, ANNs can analyze extensive temporary patterns in daily cycles, while learners such as CatBoost and LightGBM refine interval-specific predictions. This layered application ensures the forecasts are both granular and contextually relevant, which addresses the short-term operating requirements, supporting the long-term strategic plan (Hajirahimi, 2023). These recommendations from the literature review were summarized as;

i. Apply a modular hybrid architecture for flexibility and adaptability.
ii. Integrate uncertainty to handle input perturbations.
iii. Apply uncertainty thresholding to reduce overestimation risks.
iv. Use hybrid learning (ANN + SVR+ tree-based learners) to catch complex temporary mobility and overfitting.
v. Include regularization and model simplification to reduce complexity and overfitting.
vi. Establish benchmarking cycles to maintain competitiveness and guide refinement.
vii. Employ the complementary model integration for better accuracy and uncertainty estimation.
viii. Strategically deploy models for broad patterns and interval-specific refinements.

2.2.11 Summary of Specific Objective One: Weaknesses of Existing Models

This section discusses the literature review on the weaknesses of existing models that affect the performance of hybrid forecasting models. The results from the literature review of existing power demand forecasting models highlighted several important weaknesses that compromise their effectiveness in a dynamic and uncertain environment. The most significant of these are issues of rigor, susceptibility to overfitting, inadequate mechanisms for determining uncertainty, and sensitivity to input perturbations. These limitations reduce the reliability of forecasts and affect decision-making in energy management and planning.

The synthesis of evidence from literature suggests that while artificial nerve networks (ANNS), Recurrent Neural Networks (RNNs), Support Vector Regression (SVR), and XGBoost, RandomForest, LightGBM, and CatBoost exhibit strong predictive capacities in specific contexts, no single approach offers a comprehensive solution to the multifaceted challenges of power demand forecasting. This recognition has increasingly directed attention toward hybridization strategies, which combine the complementary strengths of several models to improve future accuracy, strength, and adaptability.

Furthermore, uncertainty emerged as a major recommendation from the literature, which enabled the model to provide probabilistic rather than deterministic forecasts. This enhancement is crucial for capturing the underlying variability of electricity power demand and mitigating the risks associated with overconfident or erroneous predictions. Regularization and sorting techniques were also identified as an effective approach to simplify model architecture, reduce computational burden, and increase interpretability without compromising forecasting accuracy. Constant benchmarking against both traditional statistical methods and advanced machine learning models ensures that hybrid approaches remain competitive and contextually relevant.

Taken together, these insights formed the conceptual foundation for the hybrid power demand forecasting model, which includes an estimate of uncertainty under input perturbations. Literature confirms that a hybrid design equipped with uncertainty estimation offers a viable pathway for overcoming the weaknesses of existing models. Therefore, this summary marks a significant transition from diagnosing weaknesses in the existing studies to informing the design of our study. The next stage of this research involved operationalizing these insights into a systematic design that integrates adaptive benchmarking into diverse teaching strategies, uncertainty modeling, and adaptive benchmarking to create a coherent and scalable forecasting framework.

2.3 Review of Existing Hybrid Models Design

In recent years, there has been an increasing interest in developing new methods for Uncertainty Estimation designed to be robust to input perturbations. Various drawbacks of individual methods in load forecasting include increased error percentage, computational complexity, and reduced computing efficiency. Over the years, researchers have focused on developing hybrid methods and models that offer higher precision and accuracy while minimizing error rates.

Hybrid models combine two or more single methods to achieve improved efficiency and accuracy in load forecasting (Dai, 2022). In the hybrid model, single methods are chosen based on their specific contributions and suitability for load forecasting. SVM and ANN are two widely used and prominent methods. These techniques are combined with other single methods to achieve minimal error rates and create the most effective load forecasting model. SVM is particularly well-suited for handling semi-structured and unstructured data, yielding optimal output (Dai, 2022).

A research study by Xu (2020) focused on generating adversary examples using the Iterative Nesterov-Momentum Fast Gradient Method (Nesterov-Momentum I-FGSM). The aim of the study was to investigate the effectiveness of this iterative attack method and its potential to generate adversary examples with a high success rate. Nester momentum I-FGSM is a variant of the original FGSM that incorporates Nester momentum into the iteration process. This technique enables more efficient and effective generation of adverse examples by using momentum information to guide perturbation updates. By generating adversary examples using Nesterov-momentum I-FGSM, Xu sought to contribute to the understanding of iterative attack methods and their potential to exploit vulnerabilities in Deep Neural Networks. The study's findings have significant implications for proving the resilience and security of deep learning models against adversarial attacks (Xu, 2020).

From the literature, Deep Neural Networks (DNNs) are vulnerable to adversarial examples, which are deliberately perturbed samples created by adversaries. Fast gradient sign methods (FGSM) are commonly used for effective attacks on DNNs, but a fixed perturbation size may not be optimal for quickly finding adversary examples. The study on Adaptive Moment Iterative Fast Gradient Method (Adam-FGSM) involved moment estimations of the gradients. However, further investigation of the algorithmic efficiency and convergence of Adam-FGSM compared to traditional FGSM and other iterative white-box attacks is needed. There is also a need to evaluate the generalizability and robustness of Adam-FGSM across different architectures and datasets, as well as its portability to various models and datasets (Xu, 2020).

One approach that has not been fully used is Ensemble-Based Methods. It leverages multiple models trained on different subsets of data or with different architectures to improve the robustness of Uncertainty Estimation. Ensemble learning enhances classification performance by leveraging the strengths of multiple models, as individual machine learning models can exhibit high variability, high bias, and low accuracy. Studies have shown that ensemble models often achieve higher accuracy than individual models (Mienye, 2022).

The study titled "How Good is the Bayes Posterior in Deep Neural Networks Really?" identified important design gaps in the application of the Bayes posterior approximation for Uncertainty Estimation in Deep Neural Networks (DNNs) (Wenzel, 2020). One of the key design gaps highlighted by the authors is the lack of calibration in the Bayesian posterior. The study revealed that the predicted uncertainties from the Bayes Posterior often did not align well with the observed frequencies of events, indicating a lack of reliability in the uncertainty estimates. This calibration gap highlights the need for enhanced methods to calibrate uncertainty estimates in Bayesian deep learning models, as proposed in our study.

The Research study done on Variational Mean-Field Theory (VMFT) for training Restricted Boltzmann Machines (RBMs) with binary synapses provided valuable insights into training techniques for RBMs with discrete synapses (Huang, 2020). There are potential design gaps that could be further addressed in their study to assess its performance on more complex and diverse datasets, including those with high-dimensional input spaces, imbalanced classes, or noisy samples.

2.4 Review of Previous Study Model's Implementation Methods

This section discusses various similar research studies involved in implementing a hybridized robust technique for enhancing Uncertainty Estimation in deep learning models under input perturbations, including Data Preparation, Model Architecture, Training, Uncertainty Estimation, Validation and Evaluation, hyperparameter tuning, Testing, and Deployment.

Data preparation is employed to prepare the initial signal in a way that enhances the accuracy of the forecasting. Various methods can be utilized during the data preprocessing phase of electricity demand forecasting. It includes the division, decomposition, standardization, and normalization of data. The dataset is divided into a training set and a testing set. A researcher employed 10-fold cross-validation to increase the diversity of the training subset and assess performance evaluation in hybrid wind power forecasting (G. Wang, 2020).

The Random Forest algorithm begins with data sampling, where random subsets of the training data are selected with replacement. For each of these subsets, a decision tree is constructed using a randomly chosen subset of features. Once all trees are built, the model makes predictions through a process of voting or averaging. In classification tasks, each tree votes for a class, and the class with the majority of votes is selected as the final prediction. For regression tasks, the model calculates the average of the predictions made by all trees (Zhou Z. H., 2022).

From the Literature on the computation of input Perturbations, a researcher used a Fast Gradient Sign Method (FGSM). The Objective was to generate adversarial examples for testing model robustness. The Equation used;

$$X_{adv} = X + \epsilon \cdot sgn\left(\nabla_x J(xy)\right) \quad (6)$$

Where;

X is the original input,

Xadv is the adversarial example,

ϵ is a small perturbation magnitude,

J is the model's loss function, and

∇X represents the gradient with respect to the input.

Many perturbation methods focus on generating adversarial examples for a specific model but may lack transferability to other models. There is a need to assess the transferability of generated adversarial examples across different architectures and datasets (Sarkar, 2020).

Data decomposition is a technique commonly used in signal processing problems, where a high-dimensional dataset is broken down into multiple low-dimensional sub-datasets. Researchers have used a four-signal decomposing algorithm in the multi-step wind speed forecasting. An imputation and decomposition algorithm-based integrated approach with a bidirectional LSTM Neural Network for wind speed prediction was deployed. Various data decomposition strategies, including EMD, EEMD, and CEEMDAN, were used to separate datasets into their constituent components, aiming to minimize data errors. The application of these methods improved the accuracy of wind speed prediction in forecasting models (Sareen, 2023).

The Fast Gradient Sign Method (FGSM) is an algorithm for generating adversarial examples in Deep Learning models. The main idea of FGSM is to perturb the input data in a direction that maximizes the loss function's gain, with the goal of causing misclassification. A description of the Fast Transition Method (FGSM) algorithm (Chen Z. D., 2023):

i. Input:

X- Original input data

Y- The actual label corresponding to the original input.

J (θ,X,Y) - A loss function that measures the difference between the predicted output and the actual label.

ii. Gradient Computation

Calculate the loss gradient given the input data: XJ (θ,X,Y)

iii. Perturbation Calculation:

Compute the perturbation as the sign of the gradient: sign(∇XJ (θ,X,Y))

Sign (·) here returns the sign of each element in the gradient vector.

iv. Perturbed Input:

Generate an adversary example by adding a small deviation (ϵ) scaled by the sign of the gradient to the original input:

$X_{adv}=X+\epsilon\cdot\text{sign}(\nabla_X J(\theta,X,Y))$............................. ....(7)

where ϵ is a small constant representing the magnitude of the perturbation.

v. Adversarial Prediction:

Use the perturbed Xadv input to predict with the target model:

$\hat{Y}_{adv}=f(X_{adv};\theta)$

vi. Evaluation:

Evaluate the impact of the perturbation on the model's prediction and assess whether the adversary example is misclassified or causes the model to make an incorrect prediction.

Data normalization is particularly useful when features in the dataset have different scales or units, as it helps to eliminate biases that may arise due to these differences. By bringing the data to a consistent scale, normalization ensures that no single feature dominates the analysis or learning process due to its magnitude. In a research study, the normalization technique was employed to rescale the input measurements of wind energy within the range of 0 to 1. Additionally, data standardization was applied in deep learning models to ensure consistent data sizes, utilizing Z-Score values as a metric to evaluate the scales of the data (J. Manero, 2019).

Data filtering is a fundamental process in research that involves removing unwanted noise or artifacts from a dataset to enhance the quality and reliability of the data. It is especially important when dealing with measurements or signals that contain undesired variations, disturbances, or outliers. A research study applied the Kalman Filter (KF) to reduce the learning rate and increase performance in long-term wind power forecasting (J. M. Lima, 2017).

Fort (2019) focused on Deep Ensembles in the context of improving accuracy, Uncertainty Estimation, and out-of-distribution robustness in deep learning models. Deep Ensembles are a technique where multiple neural networks are trained independently and their predictions are combined to make a final decision. A researcher compared Deep Ensembles with Bayesian Neural Networks, which learn distributions over network parameters. The researcher noted that while Bayesian Neural Networks are theoretically motivated by Bayesian principles, in practice, they do not perform as well as deep ensembles, especially when faced with dataset shift.

To explain this performance gap, the researchers suggested that popular Scalable Variation Bayesian methods tend to focus on a single mode, while Deep Ensembles explore different modes in the functional space. To investigate this hypothesis, they analyzed the Loss Landscape of Neural Networks and measured the similarity of features in the prediction space. The researcher recommended that future research could explore alternative Variational Bayesian methods that can effectively explore and exploit multiple modes, potentially bridging the performance gap between Bayesian Neural Networks and Deep Ensembles (Fort S. H., 2019).

The loss landscape in deep learning refers to the space formed by the loss function with respect to the model parameters. Visualizing the loss landscape offers insight into the optimization process, decision boundaries, and challenges associated with training deep neural networks. Common visualization techniques include contour plots, 3D surface plots, heat maps, trajectories, PCA projections, slices, scatter plots, interactive tools, parallel coordinates, and density plots. It is important to note that the loss landscape is highly dimensional, so researchers often use a combination of these methods. These visualizations help to understand the optimization dynamics and identify potential challenges in training deep neural networks (Elhamod, 2023).

Data sampling is a common research technique used to select a subset of observations or data points from a larger population or dataset. It is employed when it is not feasible or practical to collect and analyze data from the entire population. In a research study, wind power forecasting was assessed by researchers using a data sample from a training dataset spanning from three months to two and a half years. The findings revealed that the forecasting results were notably consistent when using a dataset longer than one year. However, the accuracy of the forecasts decreased when using a dataset that was either longer than two years or shorter than one year.

Data downscaling is a technique used in research to refine coarse-scale data to a finer spatial or temporal resolution. This process enhances the accuracy and applicability of the data for specific research objectives. The downscaled data provides a more detailed representation of the phenomenon under study at a localized level. A research study applied data downscaling to refine satellite-based precipitation estimates to a higher spatial resolution using ground-based observations. The downscaling method employed in the study provided a more accurate representation of precipitation patterns at a local scale, enabling a more precise analysis of the impact of rainfall on ecosystem dynamics (Smith, 2019).

Feature Selection in Electricity Demand Forecasting refers to the process of identifying and selecting the most relevant and informative features or variables from a larger set of potential predictors. The goal is to identify the subset of features that contribute most to the accuracy and efficiency of demand forecasting models. It identifies a set of potential features that are likely to have an impact on electricity demand. These features can include historical demand data, weather variables (such as temperature, humidity, or wind speed), calendar-related variables (such as day of the week or holiday indicators), economic indicators, or any other factors that may influence electricity consumption. A research study was conducted to examine the impact of wind speeds, blade pitch angles, temperature, and nacelle orientation on the prediction of wind power at different heights and under wind shear conditions using a deep learning neural network. The findings demonstrated that the blade pitch angle played a crucial role in wind power generation, highlighting its significance in determining the amount of power produced from wind turbines (Z. Lin and X. Liu, 2020).

The Research study done on residential Electricity Consumption. The power demand of individual buildings studied exhibited fluctuations due to the connection of different loads during various seasons. The research studied a system specifically designed for residential houses in four seasons: Summer, Fall, Winter, and Spring. The system also considered the distinction between weekdays and weekends. The consumer’s response to sudden events was reflected in the varying power consumption patterns (Nadeem & Arshad, 2019).

Algorithm functions and hyperparameter adjustment played a crucial role in Uncertainty Estimation models. These aspects determined the behavior and performance of the model in quantifying and representing uncertainty. The choice of algorithm functions depends on the specific Uncertainty Estimation technique being employed. Hyperparameters are parameters that are not learned from the data but are set prior to training the Uncertainty Estimation model (P. A. Adedeji, 2020). Adjusting hyperparameters is a crucial step in optimizing the performance and behavior of the model. It involves tuning these parameters to enhance the model's ability to accurately estimate uncertainty. A researcher trained and testing a function using data collected from different time frames (Kilimci et al., 2019). These data were divided into a training and a testing dataset.

Activation functions play a crucial role in Uncertainty Estimation models for electricity demand forecasting. These functions introduce non-linearity into the model and allow it to learn complex relationships between input features and uncertainty estimates. The Linear Activation Function (KLF) simply passes the input through without any transformation. It is commonly used in the output layer of the Uncertainty Estimation model when the goal is to estimate point forecasts rather than uncertainty measures. Rectified Linear Unit (ReLU) is a widely used activation function that maps all negative input values to zero, while preserving positive values. It is simple, computationally efficient, and helps in learning Non-Linear patterns in the data (M. Qolipour, 2019). Sigmoid Activation Function maps the input to a value between 0 and 1. It is often used in Uncertainty Estimation models to transform the outputs into probabilities or confidence measures, as it can restrict the values to a specific range.

Softmax Activation Function is typically used in multi-class classification tasks, where it normalizes the outputs into a probability distribution across different classes. It ensures that the estimated probabilities sum up to one, facilitating interpretation and Uncertainty Estimation in classification problems. The Hyperbolic Tangent (tanh) function maps the input to a value between -1 and 1, providing a smooth curve with both positive and negative values. It has been commonly used as an activation function in recurrent neural networks (RNNs) or long short-term memory (LSTM) networks.

The Exponential Linear Unit (ELU) is an activation function that combines the benefits of ReLU with improvements that address some of its limitations. It allows negative values and has a smooth curve that helps alleviate the "dying ReLU" problem, where neurons become inactive for negative inputs. The training algorithms utilized by researchers include Support Vector Machine (SVM), Logistic Regression, Functional Margin, and Radial Basis KnelFunction (M. Liu, 2020).

The proper selection of hyperparameters significantly influences the accuracy of electricity demand forecasting in research. Incorrect combinations of hyperparameters can lead to inaccurate results in electricity demand forecasting. The traditional trial-and-error method for hyperparameter tuning is inefficient and time-consuming. Therefore, optimization techniques are utilized to identify the optimal values of hyperparameters.

Typically, forecasting error functions serve as objective functions in the optimization process a study by C. Wu (2020) applied a multi-objective grey wolf optimization to update the weight and threshold of the ELM model. A genetic algorithm (GA) was employed to optimize a Radial Basis Function Neural Network (RBFNN) designed specifically for wind power forecasting. The GA was used to find the optimal values for the center vector and spread of the Gaussian function within the RBFNN model (A. Zameer, 2017).

In the research study on “How Many Degrees of Freedom do we Need to Train Deep Networks A Loss Landscape Perspective,” The authors investigated the context of random subspaces by initially examining the probability of successfully reaching a specific sub-level set of training loss while training within a randomly chosen subspace of a given dimensionality (Larsen, 2021). It was observed that a distinct and abrupt transition in the success probability from 0 to 1 occurs as the training dimension exceeds a certain threshold. The threshold training dimension increased as the desired final loss became smaller, but decreased as the initial loss decreased.

The researchers provided a theoretical explanation for the occurrence of the phase transition and its relationship to initialization and the desired final loss. This explanation was based on the characteristics of the high-dimensional geometry of the loss landscape. The study employed Gordon's Escape Theorem (GET) to demonstrate that the combination of the training dimension and the Gaussian width of the desired loss sub-level set, projected onto a unit sphere surrounding the initialization point, must surpass the total number of parameters in order for the success probability to be significant (Larsen, 2021).

In the paper titled "The Loss Landscape Perspective. The researchers focused on understanding and analyzing the geometric properties of the loss landscape in deep neural networks. They introduced the key components of the loss landscape, including the weights, biases, and activations of a neural network (Fort, 2019). They researched how the loss function defines the landscape and influences the network's learning process. It is seen that as deep learning models grow in complexity, there is a need for techniques that can achieve high performance while using fewer parameters. Exploring parameter-efficient methods can help address this challenge. The researcher’s view was that future research should focus on understanding the role of random initialization, exploring methods that achieve higher diversity than deep ensembles, and developing parameter-efficient approaches. These directions hold potential for advancing the field of deep learning and improving model performance.

A comparative study involving data Clustering Techniques within a PSO-ANFIS hybrid model was done. The objective was to examine the impact of the clustering approach on both standalone ANFIS and optimized ANFIS models using wind turbine output data from a potential site before installing the wind turbine. The ANFIS program was executed using MATLAB (R2015a) on a desktop computer workstation with a 64-bit configuration, 32 GB RAM, and an Intel (R) Core (TM) i7 5960X processor (Adedeji, 2020).

Last layer dropout (LL-dropout), akin to other dropout variants, aimed to perform Bayesian inference on the model parameters by introducing a dropout layer on the activation before the final layer exclusively. During testing, the model was evaluated 2000 times. The model's performance during testing was heavily reliant on the chosen dropout rate "p". In accordance with prior conducted studies, the dropout probability was set to 0.2. The dropout mechanism was maintained throughout the evaluation, and Monte Carlo (MC) simulations were carried out to obtain a distribution of the predictions (Hirschfeld L, 2020).

According to research studies, the Bootstrap Technique has been employed to deduce the distributions of errors in measurements based on actual data. These errors were then propagated through the model to generate diverse instances of model parameters, challenging the assumptions of Gaussian and independent and identically distributed errors in the data. The ensemble of all modeled outcomes contributed to the construction of the overall outcome distribution, involving a total of 400 bootstrap samples (re-samplings) (Frutiger J, 2016). The initial purpose of the Ensemble Technique was to counteract theoverfitting problem and enhance the model's accuracy. This approach involves combining a sequence of independently trained similar "weak" learners, each consisting of "M" components, and merging their outputs, often through techniques such as model averaging, to generate a unified prediction.

The presence of random starting points and the complex, non-linear nature of Graphical Neural Network models, coupled with a suboptimal optimization process, all contribute to achieving a diverse array of outcomes within the ensemble. The spread of outcomes reflects the uncertainty embedded in the model's predictions. The fundamental idea is that inputs resembling the training data will yield similar outputs, provided that the model's weights are optimized in relation to these inputs, even if they are distinct. On the contrary, data that wasn't part of the training set yielded higher variance because it was more susceptible to varying sources of divergence (Scalia G, 2020).

Typically, the selection of the ensemble size "M" is restrained by the speed of training, especially when dealing with voluminous datasets that result in prolonged training periods. Through the repetition of the regression process for every fabricated dataset, a fresh set of model parameters is generated, highlighting the influence of measurement errors on the inherent uncertainty of the parameters. After fitting an initial reference model, a set of "N" bootstrap samples is established by merging the output of the reference model with the errors that are drawn with replacement. These errors follow a uniform distribution and have an equal probability of being realized (Scalia G, 2020). Differential privacy is a concept in data privacy and security that aims to provide strong guarantees about the privacy of individuals when their data is used for analysis or statistical purposes. It addresses the challenge of sharing useful insights from data while preventing the disclosure of sensitive information about any specific individual (Hesamifard, Takabi, Ghasemi, & Wright, 2018).

Differential Privacy is widely used in various domains, including social sciences, healthcare, finance, and machine learning research. However, implementing it effectively requires careful consideration of the privacy parameter, noise generation mechanisms, and the impact on the accuracy of the analysis. Striking the right balance between privacy and utility is crucial to ensure that meaningful insights can still be derived from the perturbed data while protecting individual privacy. Differential privacy (DP) has garnered significant interest due to its robust and measurable approach to safeguarding privacy, regardless of attackers' prior knowledge. As machine learning advances, DP has emerged as a dynamic field of research within the realm of privacy-focused machine learning (Wang, Zhang, Feng, & Yang, 2020).

In recent times, numerous investigations have focused on enhancing the model's effectiveness through data perturbation methods, yielding favorable outcomes, as seen in the workf the researcher (Duchi, Jordan, & Wainwright, 2018). The approach to designing data perturbation mechanisms aims to align the combined statistical outcomes with particular usability criteria by introducing positive and negative noise to individual values. Their primary focus is on enhancing error bounds while considering privacy constraints. Yet, these approaches have shortcomings in terms of limited communication efficiency and inadequate adaptability when dealing with intricate deep learning models, largely due to the extensive parameter range differences across various layers of neural networks.

In a study on Local Differential Privacy (LDP), every individual user independently introduces randomness to their data and forwards the randomized version exclusively to an aggregator responsible for conducting analyses. This practice safeguards the privacy of both users and the aggregator against the exposure of confidential information. Despite the growing research interest in LDP, most current efforts are not fully directed towards implementing LDP in scenarios involving intricate data and/or analytical assignments (Wang et al., 2019).

The design of the data perturbation mechanism aims to align the statistical outcomes of combined disturbances with particular usability criteria. This is achieved through the introduction of both positive and negative noise to individual values. The primary goal is to enhance asymptotic error limits while maintaining privacy. Nevertheless, these approaches exhibit drawbacks such as limited communication efficiency and limited adaptability when applied to intricate deep learning models. This is due to the substantial variation in the range of model parameters across different layers of neural networks. If a constant parameter range is assumed, considerable variance in estimates will arise, inevitably leading to a reduction in model accuracy (Sun, Qian, & Chen, 2020).

An alternative approach involves employing methods centered on perturbations. In this scenario, the primary challenge lies in identifying logical sets of elements within the input and applying perturbations to assess them within a reasonable timeframe, resulting in an approximate optimal solution.

A noteworthy benefit of Perturbation-Based Methods is their dynamic nature. While many other methods treat the model as a fixed entity, perturbations allow for repetitive inquiries. This enables the formulation and testing of hypotheses on the go. These methods are versatile in their applicability to various model architectures. Unlike model-dependent techniques (such as several Back Propagation-based methods) that necessitate access to internal model information for generating explanations, often yielding limited clarity, Perturbation-Based Methods avoid such limitations. As they operate as Model-Agnostic Approaches, they sidestep the requirement for internal model information, making them suitable for elucidating predictions across a wide array of Machine Learning models, including deep neural networks (DNNs) (Li W. H., 2020).

2.5 Review of Existing Evaluation Methods

AI-driven hybrid models for demand power forecasting undergo testing and validation using common performance evaluation metrics. Various research papers have presented diverse forecasting evaluation indicators to gauge the effectiveness of AI-based hybrid techniques. The commonly employed statistical error terms encompass Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE), Root Mean Square Error RMSE, and R-squared value (R2).

A researcher did a study titled "Can You Trust Your Model's Uncertainty? Evaluating Bayesian Neural Network Uncertainty Estimators on Image Classification Tasks". It provided valuable insights into the evaluation of Uncertainty Estimation methods for Bayesian Neural Networks (BNNs) on image classification tasks (Ovadia, 2019). The study primarily focused on image classification tasks, which limited the generalizability of the findings to other types of tasks or domains. The authors recommended that future research include a more diverse range of tasks, such as natural language processing, time series analysis, or Reinforcement Learning, to evaluate the performance of Uncertainty Estimation methods across different domains.

The Mean Absolute Error (MAE) is a common metric used to measure the accuracy of a prediction model. It calculates the Average Absolute Difference between the predicted values and the True Values(Mehmood, 2021). The formula for MAE is

MAE $(1 / n) * \Sigma |yi - \hat{y}i|$ .... (8)

Where:

MAE is the Mean Absolute Error.

n -is the number of data points (samples) in the dataset,

y_i-represents the true (actual) value of the i-th data point.

$\hat{y}_i$ represents the predicted value for the ith data point.

To calculate the MAE, you take the absolute difference between each predicted value and its corresponding true value, sum them up, and then divide by the total number of data points. The lower the MAE, the better the model's predictions align with the actual values.

The Mean Absolute Percentage Error (MAPE) is a commonly used metric in forecasting to assess the accuracy of prediction models. It calculates the average percentage difference between the predicted values and the actual values, providinga measure of how well the model predicts relative to the actual values (V. N. Sewdien, 2020). The formula for MAPE is as follows

MAPE = $(1 / n) * \Sigma | (y_i - \hat{y}_i) / y_i| * 100$...(9)

Where:

MAPE is the Mean Absolute Percentage Error.

n- is the number of data points (samples) in the dataset.

yi-represents the true (actual) value of the i-th data point.

yi - ŷirepresents the predicted value for the i-th data point.

MAPE expresses the error as a percentage of the actual value, allowing for a better interpretation of the model's accuracy, especially when dealing with datasets of varying scales. In Uncertainty Estimation models in electricity demand forecasting, MAPE is commonly used as one of the evaluation metrics to quantify the uncertainty in the predictions. Uncertainty Estimation models, such as Bayesian Neural Networks or Gaussian Processes, produce probabilistic predictions that include not only point estimates but also measures of uncertainty (e.g., predictive variance). MAPE helps to assess how well these models capture the uncertainty by providing a percentage measure of the prediction error (Lu, 2023).

The Mean Absolute Scaled Error (MASE) is a metric used to assess the accuracy of a forecasting method in predicting future values of a time series. It provides a scale-independent measure of forecast accuracy, making it useful for comparing the performance of different forecasting models across various time series (P. Scarabaggio, 2020).

The formula for MASE is given by:

$$\text{MASE}=\frac{1}{n}\sum_{n=1}^{n}\frac{1}{h}\frac{\sum_{t=n+1}^{n+h}|y(t)-y2n(t)|}{\sum_{t=n+1}^{n+h}|y(t)-y(t-m)|}. \quad \ldots\ldots(10)$$

where;

n- is the forecasting window index of the total N forecasting windows (with a rolling window stride σ = 1 hour),

h is the forecasting horizon (h = 24 hours in our case),

y(t) is the electrical load measurement at time t Ɛ T,

y(t-m) is the naive-m seasonal forecast at time t (with season m = 24 hours: the last known observation from 24 hours ago) and

yˆn(t) is the estimated forecast of the window n at time t.

MSE, which stands for Mean Squared Error, is a widely used metric for evaluating the performance of prediction models, including those used in electricity demand forecasting. It measures the average squared difference between the predicted values and the true values (H. Yin, 2017). The formula for MSE is as follows:

$MSE = (1 / n) * \Sigma (y_i - \hat{y}_i)^2$ ....(11)

Where:

MSE is the Mean Squared Error.

n is the number of data points (samples) in the dataset.

Yi represents the true (actual) value of the i-th data point.

ŷi represents the predicted value for the i-th data point.

MSE is popular due to its ability to penalize larger prediction errors more severely than smaller errors. It is non-negative, and a lower MSE indicates better model performance in minimizing prediction errors.

$R^{2,}$ also known as the Coefficient of Determination, is a commonly used metric to assess the goodness of fit of a regression model. It measures the proportion of the variance in thedependent variable (target) that is predictable from the independent variables features) usedin the model. R^2 is a value between 0 and 1, where 1 indicates a perfect fit and zero indicates that the model explains no variance in the target (C. Li, 2018).

The formula for R^2 is as follows:

$R^2 = 1 - (SSR / SST)$ ....(12)

Where:

R^2 is the Coefficient of Determination, which is the Sum of Squared Residuals, which measures the variation that the model could not explain, i.e., the sum of squared differences between the predicted values ($\hat{y}_i$) and the true values (y_i). SST is the Total

Sum of Squares, which measures the total variation in the target variable, i.e., the sum of squared differences between the true values (yi) and the mean of the true values ($\bar{y}$).

A different approach is to incorporate information about input perturbations directly into the Uncertainty Estimation process. For example, Li et al. (2021) investigated a method called Gradient-Adjusted Epistemic Uncertainty (GAEU), which utilizes gradient information to estimate the uncertainty of a model's predictions under input perturbations. However, the researcher did not conduct a comparative analysis with other state-of-the-art Uncertainty Estimation methods to assess the GAEU method's advantages and diffusion limitations. Comparisons with methods such as Bayesian neural networks, Monte Carlo Dropout, or Deep Ensembles would provide insights into the relative performance and effectiveness of the GAEU approach.

The approach involves leveraging information from external sources, such as natural language descriptions or the physical properties of the input data, to enhance the robustness of Uncertainty Estimation. For instance, a research study was done on a method called Scene-Graph Bayesian Deep Learning (SGBDL). The study employed a scene graph representation of the input data to enhance uncertainty estimation in object detection tasks (Sun, 2021).

In a research study on Uncertainty-Aware Scene Graph Generation, the authors introduced a Bayesian Classifier Parameterization and a Latent Relationship Distribution to address two aspects of uncertainty: Parameter and Label Uncertainty. Through experimental results on a popular benchmark dataset, the authors concluded that both strategies improved the generalization of different backbones, and the inclusion of additional labels based on the latent relationship distribution facilitated comprehensive model training, makinga significant contribution (Li, 2023). However, the authors did not show how the model would handle the input perturbation challenges.

In a research study by Fan, the study offered limited comparisons with other forecasting models. Although it claims that the proposed model achieves smaller errors than several commonly used methods, it does not engage in a comprehensive comparative analysis with a wider range of forecasting techniques (Fan, 2023). Incorporating an uncertainty estimate into this benchmarking analysis could provide deeper insight into model performance. By evaluating how well a model accounts for uncertainty relative to others, researchers can better understand its strengths and limitations, ultimately strengthening findings.

A research study was done on methods that leverage adversarial examples to improve the robustness of Uncertainty Estimation. For instance, Kaur (2021) studied a method called Robust Bayesian Neural Network (RBNN), which uses adversarial examples to train a Bayesian Neural Network with improved robustness and accuracy in Uncertainty Estimation. The challenge with this study is that it did not consider extending the evaluation of RBNN to include more challenging datasets with severe outliers, noise, or adversarial attacks, in order to further assess its robustness and generalizability. It also did not test the method on real-world datasets.

The research study titled "Can You Trust Your Model's Uncertainty? Evaluating Bayesian Neural Network Uncertainty Estimators on Image Classification Tasks" focused on evaluating different Uncertainty Estimation methods for Bayesian neural networks (BNNs) on image classification tasks (Ovadia, 2019). To evaluate the methods, the researchers conducted extensive experiments on various image classification datasets. They compared the estimated uncertainties with the actual correctness of predictions to determine the calibration and reliability of each method. They further assessed the coverage of the estimated uncertainties and investigated their correlation with model accuracy. However, the research study primarily focused on image classification tasks, and its findings did not directly generalize to other domains or tasks. Further investigations and evaluations of different types of datasets and tasks would be valuable for assessing the performance and reliability of Uncertainty Estimation methods in diverse settings (Ovadia, 2019).

A study on residential electricity load forecasting investigated a distinctive decision integration strategy, resulting in a significant reduction in the Mean Absolute Percentage Error (MAPE). The model was evaluated using a one-year dataset of Pakistan Residential Electricity Consumption (PRECON). The results demonstrated that the model achieved the best feature selection, yielding promising MAPE values of 1.70%, 1.77%, 1.80%, and 1.67% for the summer, fall, winter, and spring seasons, respectively. However, the study did not address Uncertainty Estimation in response to input perturbations (Yousaf, Asif, Shakir, Rehman, & S. Adrees, 2021).

A researcher conducted a comparative analysis of data clustering methods within a PSO-ANFIS model. The study aimed to assess how the choice of clustering algorithm influenced the performance of both a standalone ANFIS model and an ANFIS model optimized with particle swarm optimization (PSO). The research utilized synthetic wind turbine power output data from a potential site in the Eastern Cape, South Africa (Adedeji, 2020). The findings indicated that the standalone ANFIS model clustered with the SC technique yielded the best results among the standalone models, achieving a root mean square error (RMSE) of 0.132, a mean absolute percentage error (MAPE) of 30.94, a mean absolute deviation (MAD) of 0.077, a Relative Mean Bias Error (RMBE) of 0.190, and a variance accounted for (VAF) of 94.307. Furthermore, among the three hybrid models examined, the PSO-ANFIS model, which clustered with the SC technique, outperformed the others, with an RMSE of 0.127, a MAPE of 28.11, a MAD of 0.078, an RMBE of 0.190, and a VAF of 94.311 (Adedeji, 2020).

A study investigated the calibration of Neural Networks and proposes using temperature scaling to improve Uncertainty Estimation. To evaluate the calibrated models, (Guo (2017) used reliability diagrams, expected calibration error, and negative log-likelihood scores as evaluation metrics. These metrics provided insights into the reliability and accuracy of the calibrated neural networks. The findings of this study have practical implications for applications that require reliable uncertainty estimates, such as healthcare, finance, and autonomous systems. While the study demonstrated the effectiveness of temperature scaling and other calibration methods across various datasets and architectures, further research is needed to assess their performance in domain-specific scenarios, as well as in the context of diffusion challenges.

Research studies have been conducted on the AI-based hybrid method using Extreme Learning Machine (ELM), Stacked Autoencoder (SAE), and Empirical Mode Decomposition (EMD), which is a powerful approach for solving complex data analysis and prediction tasks (H.-F. Yang and Y.-P.-P. Chen, 2020). A research study on AI hybridization formed a model using Principal Component Analysis (PCA), K-Means clustering with Polynomial Expansion (KMPE), and Extreme Learning Machine (ELM). The method was a powerful approach for data analysis and prediction tasks. Each component of this hybrid method served a specific purpose in processing the data and improving the overall performance of the model (N. Li, 2019).

A study on Assessment of critical parameters for Artificial Neural Networks based Short-Term Wind Generation Forecasting. The authors designed a hybrid AI model that combines Random Forest (RF) and Deep Belief Network (DBN) to create a powerful and versatile system for data analysis and prediction tasks. Each component of the hybrid model played a specific role, allowing the model to handle complex data, learn hierarchical representations, and improve prediction accuracy (V. N. Sewdien, R. P., 2020).

A study on Artificial Intelligence-based hybrid forecasting approaches for wind power generation demonstrated that AI-based hybrid methods for predicting wind power generation demand significant datasets and high computational resources. Additionally, effective data pre-processing and error post-processing strategies are necessary to ensure high-quality wind data, which results in a substantial computational load. Therefore, further research is needed to reduce noise in the raw data and develop hybrid models with lower computational requirements. It is crucial to strike a balance between extensive computation and prediction accuracy for optimal performance (Lipu, 2021).

2.6 Perturbation Theory

Perturbation is a mathematical technique used to study and solve complex problems by breaking them down into simpler, more manageable parts. It is a powerful tool applied in various fields of science and engineering, such as quantum mechanics, classical mechanics, and fluid dynamics. The basic idea behind perturbation theory is to start with a known or "unperturbed" system and then introduce a small or "perturbative" change to the system. This change is typically represented by a small parameter, which is then used to construct a series of approximations that describe the behavior of the perturbed system (Patkowski, 2020). The perturbation approach follows the following steps;

i. Identify an unperturbed system,
ii. Perturbation Introduction,
iii. Constructing the series expansion,
iv. Solving the series expansion.

General Perturbation Equation. Let $H(x, \lambda) = 0$ .... (13)

Where:

H (x, λ) is the equation or the problem to be solved

x represents the variable(s) of interest

λ is the small perturbation parameter

The general perturbation theory approach involves the following steps:

i. Identify the unperturbed problem:

Consider the unperturbed problem, which is obtained by setting

λ=0 in the original equation:

H(x,0)=0

The unperturbed problem is solved to obtain the unperturbed solution $x0$

ii. Construct the perturbation expansion:

The solution $x$is expanded as a power series in the perturbation parameter

λ: x= $x0$ + $\lambda x1$ +$\lambda 2x$ 2+ ⋯+ λnxn +….. (14)

Substitute this expansion into the original equation H (x, λ) = 0 and collect terms of the same order in λ.

i. Solve the perturbation equations:

The resulting equations for the different orders of λ can be solved sequentially to obtain the perturbation terms x_1, x_2,…, x_n,…

The first-order perturbation term x_1 is typically the most important, as it represents the leading-order correction to the unperturbed solution x_0

ii. Construct the approximate solution.

The approximate solution is obtained by truncating the perturbation expansion at a desired order of λ: $x \approx x_0+\lambda x_1+\lambda^2 x_2+\cdots+\lambda^n x_n$

A key advantage of perturbation theory is that it enables the study of complex systems by breaking them down into simpler, more manageable parts. This can lead to significant simplifications and insights that would be difficult to obtain by other methods.

Perturbation models with tree-structured potential functions and edge-based randomization are a class of probabilistic graphical models widely used in various applications, including computer vision, natural language processing, and computational biology. These models are characterized by their tree structure and the way they incorporate randomness into potential functions (Gallavotti, 2022).

In a tree-structured potential function, the potential function is defined as the product of the individual potential functions associated with each edge in the tree. This structure allows efficient inference and learning algorithms because the model can be decomposed into simpler subproblems. The edge randomization component of these models introduces additional randomness into the potential functions. Specifically, each edge in the tree is associated with a random variable, and the potential function for that edge is a function of both the states of the nodes and the random variable. This randomization can help improve the model's ability to capture complex interactions and dependencies in the data (Zhang C. Q., 2024).

Mathematically, the joint probability distribution of a perturbation model with tree-structured potential functions and edge-based randomization can be expressed as:

$P_{(x)} = \frac{1}{z} \Pi \psi_{ij}\left(x_i, x_j \middle| \theta_{ij}\right)$ .... (15)

where:

$x = (x_1, x_2, \ldots, x_n)$ is a vector of node states

E is the set of edges in the tree

$\psi i(x_i, x_j, \theta_{ij}, \epsilon_{ij})$ is the potential function associated with the edge (i,j)

(i,j), which depends on the states of the node x_i and x_j , edge parameters $\theta_{ij,}$ and the random variable ϵ_{ij}

Z is the normalization constant (also known as the distribution function)

Edge randomization introduced by random variables ϵ_{ij} helps the model capture complex dependencies and patterns in the data, which can lead to improved performance in a variety of tasks.

2.7 Extreme Gradient Boost

XGBoost, also known as Extreme Gradient Boosting, is a highly efficient machine learning algorithm recognized for its exceptional performance, scalability, and accuracy. It has gained wide recognition in a variety of predictive tasks, including regression, classification, and evaluation. In the context of hybrid power demand forecasting models, XGBoost effectively addresses several challenges, including capturing nonlinear relationships, managing complex feature interactions, and accounting for uncertainties arising from input disturbances (T.Zhang., 2023).

2.7.1 Advantages of XGBoost

One of the most significant strengths of XGBoost lies in its accuracy and robustness. By using gradient boosting, the algorithm iteratively minimizes errors, allowing it to capture complex, non-linear relationships in the data. Additionally, it incorporates regularization techniques, including L1 and L2, to mitigate overfitting, even in high-dimensional datasets. In terms of efficiency and scalability, XGBoost supports parallel processing and distributed computing, making it adept at handling large datasets. It also has its own ability to handle missing data, a common problem in real-world applications. The algorithm excels in feature engineering, automatically calculating feature importance and helping to identify critical variables necessary for predicting energy consumption (Qinghe, 2022).

XGBoost demonstrates versatility by working efficiently with various data types and loss functions, making it well-suited for integration into hybrid models. Its interpretability is enhanced by feature importance metrics and tools such as SHAP (Shapley Additive exPlanations), which provide valuable insights into predictions and help build confidence in its outputs. Incorporating XGBoost into hybrid prediction models offers several advantages. It significantly improves accuracy by capturing complex element interactions and nonlinearities. Its robustness to noise is enhanced by regularization and perturbation-based training, increasing robustness under uncertain conditions. Additionally, XGBoost efficiently processes high-dimensional and large-scale datasets, making it a scalable solution for energy consumption prediction.

2.7.2 Weaknesses of XGBoost

Despite its many advantages, XGBoost has some weaknesses. One notable problem is its computational complexity. While this is faster than other support implementations, training can still be resource-intensive, especially when working with large datasets or during hyperparameter tuning. The algorithm is also sensitive to hyperparameters; careful tuning of parameters, such as the learning rate, tree depth, and number of guesses, is critical for optimal performance. Additionally, although tools like SHAP aid in interpretation, XGBoost remains essentially a black-box model, which makes it less transparent compared to simpler models, such as linear regression. There is also a risk of overfitting, especially with noisy or small datasets, if proper regularization or validation techniques are not employed. Additionally, XGBoost provides deterministic point estimates by default, which requires additional steps such as quantile regression or Bayesian adaptation to quantify uncertainty (Xue, 2021).

Hybrid models that combine different prediction methods use XGBoost as a robust predictive component. The design of these models begins with the integration of various input data that include historical energy consumption, weather forecasts, and socio-economic variables. Through feature engineering, the model creates lagged variables, time-based elements, and weather-derived indices to increase prediction accuracy.In terms of model components, XGBoost efficiently captures non-linear dependencies between inputs and outputsenergy demand,while complementary statistical models, such as ARIMA or deep learning models, such as LSTM, deal with seasonality, trends, or long-term dependencies (T.Zhang., 2023). The training and testing phases involve partitioning historical data for validation purposes, while perturbed datasets are used to simulate uncertainty, ensuring robust model performance.

Input perturbations, including noise or errors in weather and load forecasts, introduce significant uncertainty into forecasts. XGBoost contributes to mitigating these uncertainties in several ways.Data augmentation techniques create multiple perturbed datasets to simulate the real-world variability of input. By training on these variations, XGBoost learns to form strong relationships. Monte Carlo simulations further improve this process by running simulations with perturbed inputs to generate a distribution of predictions. The results are aggregated to derive confidence intervals for the predictions (Cao W. L., 2023). XGBoost can be modified to estimate prediction intervals directly using quantile regression, thereby quantifying uncertainty more efficiently. Combined with Bayesian frameworks, XGBoost can generate probabilistic predictions that account for parameter uncertainty.

2.8 Light Gradient Boosting Machine (LightGBM)

Light Gradient Boosting Machine (LightGBM) is a gradient boosting framework developed by Microsoft, specifically designed for efficiency, scalability, and the ability to process large datasets. It has gained prominence in various applications, including classification, regression, and evaluation, making it a popular choice in machine learning competitions and practical implementations. The model uses decision trees as base learners and uses gradient boosting to iteratively minimize the loss function, thereby enhancing its predictive capabilities (S.Chowdhury, 2020).

What sets LightGBM apart from others is its remarkable speed and memory efficiency, achieved through several innovative techniques. Among them, the histogram-based binning and tree leaf growth functions are particularly notable, allowing the model to process data more efficiently. One of the key innovations of LightGBM is its histogram-based feature merging. Instead of relying on raw feature values, this approach groups continuous data into discrete intervals (Wu, 2023). This transformation significantly reduces the computational load and memory requirements during training, enabling the model to operate more efficiently.Another significant advance is the leafy growth of trees. Unlike traditional gradient boosting methods that use level growth, LightGBM expands the leaf to offer maximum gain distribution. This strategy yields deeper and more efficient trees, which enhance accuracy, particularly when working with complex datasets.

In addition, LightGBM supports parallel learning, which enables parallel computation, distributed training, and GPU acceleration. This capability enables the framework to efficiently manage large datasets, making it well-suited for big data applications.

2.8.1 Advantages of LightGBM

LightGBM boasts several advantages, the most prominent of which are efficiency and speed. It consistently outperforms traditional gradient boosting models in terms of training speed and resource consumption due to histogram-based binning and optimized data structures. The model also achieves high accuracy, especially due to its leafy tree growth, which often leads to better performance compared to layered approaches, especially on large and complex datasets.Another significant advantage is its scalability. LightGBM is designed to work with large datasets and supports distributed training, making it an ideal choice for applications involving big data. Its flexibility further increases its appeal as it adapts to different loss functions and custom objectives, allowing it to be applied to various tasks, such as regression, classification, and evaluation. LightGBM provides insight into feature importance, enhances interpretability, and facilitates effective feature engineering.

2.8.2 Weaknesses of LightGBM

Despite its strengths, LightGBM is not without its weaknesses. One of the main problems is the possibility of overfitting. Leafy tree growth can lead to overfitting, especially for smaller or noisier datasets. Careful regularization and tuning of the hyperparameters are necessary to mitigate this problem. The complexity of tuning the model can present problems. With many hyperparameters to consider, the tuning process can be time-consuming and require considerable expertise. LightGBM is also sensitive to data preprocessing. It requires careful handling of categorical variables and missing data to avoid biasing model predictions. Although generally memory-efficient compared to some alternatives, the histogram-based approach can still be memory-intensive when working with datasets that contain a large number of features or exhibit very high cardinality. Finally, similar to XGBoost, LightGBM does not inherently model sequential dependencies, which means that significant feature engineering is required for efficient time series tasks (Chenxi Ni., 2024).

2.9 Categorical Boosting (CATBOOST)

CATBoost, which stands for "Categorical Boosting", is a machine learning algorithm developed by Yandex, aimed at solving various supervised learning problems, including regression, classification, and other tasks. This algorithm is built on the principle of gradient boosting and is specially optimized to efficiently process categorical data without extensive pre-processing. Its design is particularly advantageous for datasets that contain a combination of numerical and categorical variables, allowing it to provide robust performance in a variety of applications (Zhang W.(2022).

One of the key innovations introduced by CATBoost is its streamlined boost mechanism. This property effectively eliminates the target drift problem in gradient and Hessian calculations. By ensuring that predictions in any given iteration do not depend on future data, ordered boosting is crucial in scenarios such as time series forecasting, where the integrity of sequential data is paramount (Abed, 2024). This approach is in contrast to traditional gradient boosting methods, which often run the risk of incorporating future information and potentially biasing the results.

In addition to ordered boosting, CATBoost boasts efficient processing of categorical data. The algorithm natively supports categorical features and converts them into numerical representations using methods such as median coding, which is designed to avoid overfitting. This capability streamlines the preprocessing phase, enabling users to focus on model development rather than extensive data preparation. In addition, CATBoost includes robust mechanisms that protect against overfitting, using advanced regularization techniques and noise-resistant criteria for splitting trees. Thanks to this, the model is particularly suitable for smaller or noisier data sets where there is frequent switching (Abed, 2024).

CATBoost is also optimized for GPU acceleration, which significantly increases training time, especially when working with large datasets. This optimization enables users to leverage GPU computing power, resulting in faster modeling and increased efficiency. The mathematical design of CATBoost mirrors that of other gradient boosting algorithms, where decision trees are iteratively trained to minimize the loss function. The model is expressed as a sum of trees, where each tree contributes to the final prediction (Du, 2024).

The overall objective function in CATBoost combines a loss function with a regularization term, enabling a balance between fitting the data and maintaining a simple model. A commonly used loss function is the mean squared error for regression tasks. An innovative ordered boost technique further enhances this process by calculating residuals based solely on previous data points, thereby increasing the robustness of the algorithm in sequential tasks.For categorical features, CATBoost uses a goal-based encoding method that computes encoded values based on previous observations and includes a regularization parameter to mitigate overfitting. This transformation occurs internally during training and is seamlessly integrated into the strengthening process, facilitating the effectiveness of the model (Zhang W. (., 2022).

At each iteration, CATBoost fits the new tree to the residuals of the previous predictions and updates the model output accordingly. The tree splitting criterion is designed to maximize the gain, which is calculated based on the gradients and Hessians of the nodes. This careful approach to tree construction contributes to the overall efficiency of the algorithm.In addition, CATBoost can be modified to estimate the uncertainty in its predictions, which is particularly valuable for tasks such as energy consumption forecasting. Two common methods for estimating uncertainty include Monte Carlo simulation and quantile regression. Monte Carlo simulation involves adding noise to the input functions and generating multiple predictions to calculate the mean and variance. In contrast, quantile regression adjusts the loss function to predict specific quantiles, allowing the model to efficiently estimate prediction intervals (Rongquan. Z, 2022).

Despite its advantages, CATBoost is not without its weaknesses. The algorithm can be computationally intensive compared to simpler models and may exhibit longer training times than non-boosting algorithms, especially for large datasets. Additionally, the high memory consumption associated with processing large categorical variables can pose challenges for users with limited resources.In the context of forecasting daily energy demand, CATBoost demonstrates its ability to model relationships between lagged energy consumption, time-based features (such as hour and day of the week), and external variables (such as weather conditions). Its ability to handle uncertainty through simulation or quantile regression makes CATBoost a strong candidate for predictive models that require robust and reliable predictions. CATBoost's unique features and design make it a powerful machine learning tool, especially for tasks involving complex data structures (Zhifan, 2024).

2.10 K-Nearest Neighbors Modeling Concept

The K-Nearest Neighbors (K-NN) algorithm is a basic machine learning method known for its simplicity and efficiency in solving both classification and regression problems. As a non-parametric instance-based learning algorithm, K-NN does not assume a specific form of the underlying data distribution. Instead, it operates on the principle of similarity, predicting results based on the closest data points in the feature space. This feature makes K-NN particularly suitable for applications where the relationships between features and the target variable are complex and non-linear, enabling robust performance in a variety of real-world scenarios (Marzbani, 2023).

In regression tasks, K-NN predicts continuous target values by aggregating the outputs of the k-nearest neighbors from the training data. The K-NN regression process involves three primary steps: calculating distances, selecting neighbors, and making predictions through aggregation. Together, these steps allow the model to estimate a target value for a new query point based on existing patterns in the data. The first step is calculating the distance, which involves determining the distance between the query point and all points in the training set (Sreekumar, 2024). This distance serves as a measure of similarity in feature space. Several distance metrics can be used, including:

Euclidean distance: Commonly used for numerical elements and is defined as

$$d(x_1, x_2) = \sqrt{\sum_{i=1}^{n} (x_{1i} - x_{2i})^2} \quad \text{.......} \quad (16)$$

which is commonly used for numerical values.

Manhattan Distance: This metric is particularly useful for high-dimensional or sparse data

$$d(x_1, x_2) = \sum_{i=1}^{n} |x_{1i} - x_{2i}|$$

Minkowski distance: A generalization of both Euclidean and Manhattan distance, defined as $d(x_1, x_2) = \left(\sum_{i=1}^{n} |x_{1i} - x_{2i}|^{\frac{1}{p}} \right)$ .... (17)

where p is a parameter that allows flexibility in distance measurement.

Once the distances are calculated, the next step involves finding the k-nearest neighbors. The k-closest query points are identified as most relevant for predicting the target value based on the assumption that they share similar properties with the query point.

The last step is aggregation for prediction, where the predicted value is typically the average of the k-nearest neighbor target values. Mathematically, this is expressed as

$$\hat{y}_q = \frac{1}{k}\sum_{i=1}^{k} y_i \quad (18)$$

Whereyi represents the k-nearest neighbor target values. Alternatively, a weighted averaging approach can be used, which gives the nearest neighbors a higher influence in the prediction, represented as

$$\hat{y}_q = \frac{\sum_{i-1}^{k} \omega_i y_i}{\sum_{i-1}^{k} w_i} \quad (19)$$

The weights (w_i) are often inversely proportional to the distance

K-NN offers several advantages for regression tasks. Its simplicity is among its most significant strengths, as the algorithm is easy to understand and implement and requires minimal assumptions about the underlying data. This non-parametric nature allows K-NN to adapt to different data distributions, which is particularly advantageous for capturing complex, non-linear relationships. Additionally, K-NN is flexible in handling multidimensional data, working efficiently with numeric, categorical, or mixed feature types, provided that appropriate distance metrics are chosen. The algorithm is also robust to small variations in the data because predictions are based on aggregated information from multiple neighbors, rather than a single model parameter (Alam, 2023).

Despite its strengths, K-NN has several limitations. One of the most significant disadvantages is its computational cost, especially for large datasets. The need to calculate the distances between the query point and all the training points can become prohibitive as the dataset grows—this computational burden, combined with the algorithm's sensitivity to noise and outliers.A single outlier among the nearest neighbors skews the forecasted values, especially when k is small. The choice of k itself is important, as a small k can lead to overfitting, while a large k can result in underfitting of the model (Mukilan, 2023). KNN also assumes uniform importance of features, which is a problem if certain features are more influential than others.

In the field of energy demand forecasting, especially for electric vehicle (EV) charging stations, K-NN has been effectively applied in various studies. For example, research conducted at UCLA involved forecasting charging demand at 20 EV stations using data collected over several years. The study aimed to provide hour-ahead forecasts and used both Euclidean distance and time-weighted dot product (TWDP) as similarity measures (Q. Huang, 2017). The findings showed that K-NN provided more accurate predictions when TWDP was used, demonstrating the algorithm's adaptability to different data characteristics and temporal patterns.

Further evaluation of the effectiveness of K-NN compared it with other forecasting methods, such as ARIMA and Pattern Sequence-based Forecasting (PSF). The results showed that K-NN outperformed these methods in terms of forecasting accuracy, highlighting its reliability in forecasting energy demand. K-NN has been integrated into hierarchical clustering frameworks to predict energy consumption profiles across multiple EVs. This method involved treating each data point as a separate cluster and merging the closest clusters to minimize within-cluster differences, illustrating the ability of K-NN to identify distinct patterns in energy consumption.

The implementation of K-NN in a hybrid energy demand forecasting model is justified by its non-parametric nature, simplicity, robustness to input perturbations, flexibility in distance metrics, increased prediction accuracy, adaptability to clustering, and efficiency with small datasets (Mukilan P. B., 2023). These attributes make K-NN a compelling choice for accurately predicting EV charging requirements while addressing uncertainties in the input data. K-NN stands out as a critical component in developing effective forecasting solutions in a dynamic energy demand management environment.

2.11 Artificial Neural Network (ANN) Concept

An Artificial Neural Network (ANN) is a mathematical regression model for predicting energy consumption. Artificial neural networks (ANNs) have emerged as a powerful class of machine learning models that draw inspiration from the biological neural networks of the human brain. Historically, their effectiveness in capturing complex, non-linear relationships in data makes them an ideal choice for predicting power demand (Khan A. S., 2020). This is especially true when ANNs are integrated into a hybrid model designed to account for uncertainty estimation. Such a design not only clarifies the operations of ANNs but also presents a compelling argument regarding their strengths and weaknesses in the context of hybrid forecasting models that use time series data for predictions, all while accounting for input perturbations.

i. ANN model design

The ANN architecture starts with an input layer that serves as a gateway for features derived from the dataset. In the field of energy demand forecasting, these features typically include historical energy consumption data, temperature data, prevailing weather conditions, and temporal elements such as time of day or day of the week (Althubiti, 2023). Each input is represented as a feature vector, denoted as:

$$X_t = [x_1, x_2, \dots, x_n] \quad \text{......(20)}$$

where *n* denotes the number of elements and t denotes a specific time step in the time series.

ii. Hidden layers

As input data flows through the network, it is processed by one or more hidden layers. In these layers, neurons perform weighted sums and apply non-linear activation functions that are necessary to capture complex patterns in the data. Exit $z_j^{[l]})$from neuron j in the layer can be expressed mathematically as:

$$z_j^{[l]} = \sum_{L=1}^{n} w_{ij}^{(l)(l-1)} {a_i} + b_j^{(l)} \quad \text{.......} (21)$$

Where;

w_{ij}: represents the weight connecting neuron i from the previous layer to neuron j in the current layer.

$B_j^{(l)}$: is the bias term associated with the neuron

$a_i^{(l-1)}$: indicates the activation output from the previous layer.

Various activation functions, such as the Rectified Linear Unit (ReLU), sigmoid, and hyperbolic tangent (tanh), are used to introduce nonlinearity into the model. These features enable ANNs to learn and generalize from complex data patterns efficiently.

iii. Output layer

The output layer of the ANN is tasked with predicting the target variable, specifically energy consumption for the next time step. If the output neuron uses a linear activation function, the predicted energy consumption y^t at time t can be calculated as follows:

$$zj(l) = \sum_{i=1}^{n} wij(l)ai(l-1) + bj(l) \quad \text{.......}(22)$$

Where m represents the number of neurons in the last hidden layer and ai(L) denotes their activation outputs.

iv. Loss function.

A loss function is used to evaluate the accuracy of the predictions generated by the regression model. The mean squared error (MSE) is a widely accepted choice for this purpose:

$$L(y, \hat{y}) = \frac{1}{N} \sum_{t=1}^{N} (y_t - \hat{y}_t)^2 \quad \text{.......}(23)$$

Where;

y_t denotes the actual power demand,

$\hat{y}_t$ represents the predicted value,

N is the total number of samples.

v. Training through back propagation

The training process for ANNs involves backpropagation, a method used to optimize the weights and biases of the network. During this process, the gradients of the loss function with respect to the weights are calculated using the chain rule (Dampfhoffer, 2023). These gradients make it easy to update the parameters using optimization algorithms such as Stochastic Gradient Descent (SGD) or Adam, as illustrated by the following equation:

$$w \leftarrow w - \eta \frac{\partial L}{\partial w} \quad \text{...(24)}$$

Where η represents the learning rate, a hyperparameter that determines the size of the steps taken during the optimization process, equation 24 is an update equation used in Gradient Descent,a fundamental optimization algorithm for training machine learning and deep learning models.

2.11.1 Advantages of ANNs

One of the most significant advantages of ANNs is their ability to capture non-linearity, for example, in power consumption data. The interplay of various factors, such as temperature fluctuations, peak usage times, and holiday effects, often manifests in complex, non-linear patterns. ANNs excel at modeling these relationships, making them invaluable tools for forecasting tasks characterized by irregular and fluctuating demand patterns.ANNs also have a remarkable capacity for feature representation. Deep networks autonomously learn hierarchical representations of data and uncover latent features that may not be readily apparent using conventional feature engineering techniques (Bu, 2023). This inherent capability not only minimizes the need for manual intervention but also enhances the model's adaptability to various datasets.

Additionally, ANNs can effectively incorporate uncertainty estimation under conditions of input disturbance. By employing techniques such as Monte Carlo Dropoutwhere a dropout is applied during both training and prediction,ANNs can generate probabilistic outputs that provide estimates of prediction uncertainty. This property is especially crucial for robust decision-making in the energy sector, where understanding the degree of uncertainty can significantly influence operational strategies and resource management (Ahmed, 2023).Integrating ANNs into hybrid energy demand forecasting models is a transformative approach that leverages their strengths in nonlinearity, feature representation, and uncertainty estimation, ultimately leading to more accurate and reliable forecasting results.

2.11.2 Weaknesses of ANN and Their Role in Hybrid Model

Although powerful machine learning tools, artificial neural networks (ANNs) have revealed several notable weaknesses in various applications, particularly in the context of power demand forecasting, one of the most significant challenges associated with ANNs has been their tendency to overfit the training data. This problem was particularly pronounced in deep neural networks, which, when trained on small or noisy data sets, often learned to memorize training examples rather than generalize from them (Al-Sharafi, 2023). Although regularization techniques such as dropout and L2 Norm regularization were used to mitigate this problem, incorporating simpler and more interpretable models such as K-Nearest Neighbors (KNN) or decision trees into the hybrid model provided a more robust solution. These simpler models demonstrated superior generalizability, improving the overall predictive performance of the hybrid framework.

Another weakness of ANNs is their interpretability. Often described as “black boxes,” ANNs offered limited transparency about how they arrived at specific predictions. This lack of interpretability presented significant challenges in critical applications such as energy forecasting, where it was essential for stakeholders to understand the underlying factors influencing forecasts (Bilgiç, 2023). To address this limitation, hybrid models were developed that included interpretable components such as linear regression or KNN. These components not only complemented the comprehensive forecasts generated by ANNs but also provided valuable insights into the key factors influencing energy demand, promoting greater trust and understanding among users.

The high computational complexity associated with training and implementing ANNs, particularly when working with large datasets, proved to be another significant drawback (Uzor, 2024). The extensive resources required for processing could hinder the effectiveness of forecasting operations. However, by delegating simpler tasks to other models in a hybrid setup, computational resources could be managed more efficiently. This strategic division of labor enabled a more efficient allocation of computing power and allowed the hybrid model to operate at scale without incurring disproportionate costs.

2.12 Bayesian Neural Network Concept

Bayesian Neural Networks (BNNs) have become a crucial component in the field of machine learning, particularly in scenarios that require not only accurate predictions but also a comprehensive understanding of uncertainty. In areas such as Power demand forecasting, where external factors and fluctuating inputs introduce significant variability, BNNs present unique advantages that traditional models often lack. Incorporating a probabilistic framework, BNNs enable practitioners not only to forecast demand but also to quantify the uncertainty surrounding those forecasts. This is critical in environments where decision-making depends on understanding risk and thereby improving operational strategies. This study delves into the training trends, strengths, and weaknesses of BNNs while exploring their synergistic potential in combination with other machine learning models such as Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), K-Nearest Neighbors (KNN), XGBoost, CatBoost, and Extreme Gradient Boosting (XGB) to create a robust hybrid forecasting framework.

At its core, a Bayesian neural network extends the traditional neural network architecture by treating the model weights as distributions rather than fixed constants. This innovative approach enables BNN to capture both aleatoric uncertainty, stemming from inherent data noise, and epistemic uncertainty that arises from limited training data or model complexity. By leveraging the principles of Bayesian statistics, BNNs offer a more nuanced understanding of the underlying data, enabling them to generate not only point estimatesbut also confidence intervals and predictive distributions (Tziolis, 2023). This capability is particularly beneficial in high-stakes areas, such as energy demand forecasting, where understanding the range of possible outcomes can lead to more informed decision-making. The probabilistic nature of BNN fosters greater confidence in model outputs, particularly in scenarios where data may be incomplete or subject to significant variance.

Mathematically, the posterior distribution $P(W|X,y)$ of weights W is derived using Bayes' theorem, which encapsulates the relationship between prior beliefs about the weights, the likelihood of the observed data given those weights, and the evidence that normalizes the results. This formulation is crucial because it allows BNN to quantify uncertainty in a mathematically precise way (Tziolis G. L.-L., 2024). The ability to express uncertainty not only aids in risk assessment but also improves model interpretability, enabling stakeholders to make informed decisions based on a comprehensive understanding of potential risks and rewards.

Training methodologies for BNNs have seen significant progress, characterized by several key trends that increase their usability and performance. One notable method is variational inference (VI), which approximates the posterior distribution by optimizing a simpler distribution.

q(W) to minimize the Kullback-Leibler divergence (Devadas, 2023). This approach has evolved with the introduction of expressive approximations such as normalization flows, which allow for more complex posterior shapes, and stochastic VIs, which facilitate scalability in large datasets. As computational resources have become more readily available, these techniques have opened up new avenues for the effective training of BNNs in real-world applications.

Another application is Monte Carlo Dropout, which applies dropout during both the training and inference phases. This technique effectively samples the approximate posterior, providing a lightweight and computationally efficient mechanism for uncertainty estimation. Its suitability for large datasets makes it particularly attractive to practitioners who wish to implement BNNs without incurring prohibitive computational costs. Moreover, Markov Chain Monte Carlo (MCMC) methods, while providing accurate posterior sampling, are often computationally intensive (Sabri, 2024). Recent innovations have sought to hybridize MCMC with gradient-based methods, achieving a balance between the accuracy of posterior estimation and the computational efficiency required for practical applications.

Ensemble methods have gained traction in the BNN landscape. By training multiple neural networks and averaging their predictions, these methods can simulate Bayesian behavior, allowing a more comprehensive approximation of posterior distributions in high-dimensional spaces (Bhotto, 2021). This ensemble approach not only enhances the accuracy of predictions but also fosters a more robust understanding of uncertainty, making it a valuable tool in the arsenal of machine learning practitioners.

One of the outstanding features of BNNs is their unique ability to quantify uncertainty, which is critical in scenarios characterized by noisy, incomplete, or perturbed input data. In power demand forecasting, where accurate forecasts are critical to operational efficiency and resource allocation, the ability to estimate uncertainty can significantly influence decision-making processes. BNNs excel in this regard by providing not only point estimates but also probabilistic outputs that reflect the confidence associated with each prediction (Brusaferri, 2022). This feature is particularly beneficial in dynamic environments where demand fluctuates due to various external factors, such as weather changes, economic conditions, or unexpected events.

The uncertainty quantification provided by BNN enables stakeholders to engage in risk-informed decision-making. By understanding the range of possible outcomes and the associated probabilities, organizations can design strategies that mitigate risks while capitalizing on potential opportunities. For example, utilities can better prepare for peak demand scenarios by recognizing the likelihood of different demand levels and thus optimize their resource allocation and operational strategy. This proactive approach, based on probabilistic BNN outputs, increases the overall resilience of energy systems, ultimately benefiting both consumers and providers (Yaseen, 2023).

In addition to quantifying uncertainties, BNNs are known for their robustness in making predictions, especially in the presence of outliers and noisy data. Traditional deterministic models often struggle to maintain accuracy when faced with such challenges, leading to overconfident predictions that can result in poor decision-making.

In contrast, BNNs take into account distributions over weights, allowing them to adaptively account for variability in the data (Franchi, 2024). This inherent flexibility enables BNNs to provide more reliable predictions, as they are less susceptible to being affected by anomalous data points or noise.

Another significant strength of BNNs is their ability to generate confidence intervals around predictions, thereby providing interpretable outputs that aid in risk assessment. These confidence intervals serve as a valuable tool for stakeholders to gauge the reliability of forecasts. In high-stakes power demand forecasting, a clear understanding of the potential demand scale is crucial for forffective resource planning and management (Kaiyan, 2022). By offering insight into forecast variability, BNN enables decision-makers to consider not only the most likely outcomes but also less likely scenarios that could impact operations.

The interpretability of BNN outputs is particularly beneficial in industries where it is essential to understand the rationale behind the predictions. Stakeholders can evaluate the likelihood of different demand scenarios, facilitating more detailed discussions on risk management and operational strategies. This transparency promotes greater confidence in the decision-making process, as organizations can align their strategies with a comprehensive understanding of potential risks and rewards. Ultimately, the confidence intervals provided by BNN contribute to a more informed and proactive approach to energy demand forecasting.

The flexibility of BNNs is another notable attribute that increases their applicability in various domains, especially in the context of hybrid modeling. BNNs can adapt to different levels of uncertainty, making them suitable for integration into hybrid models that combine multiple machine learning approaches. This adaptability allows professionals to tailor solutions that meet specific forecasting needs and adapt to unique data characteristics and application requirements. For example, when forecasting energy demand, BNNs can be seamlessly integrated with other models to enhance forecast accuracy while providing valuable uncertainty estimates (Mahajan, 2024).

In hybrid models, BNNs complement other machine learning techniques by adding a probabilistic layer that enriches the overall framework. This enables the hybrid model to leverage the strengths of each component while mitigating its respective weaknesses. For example, while ANNs excel at capturing complex, non-linear relationships, they often lack an inherent estimate of uncertainty. By incorporating BNN, the hybrid model can achieve robust predictions while quantifying uncertainty, ultimately leading to more reliable predictions. This flexibility and adaptability position BNNs as a critical element in the evolving machine learning landscape, particularly in applications that require a fine grained understanding and decision-making.

Despite their numerous advantages, BNNs present significant challenges, particularly in terms of computational complexity. BNN training often involves significant computational overhead, especially when using methods such as Markov Chain Monte Carlo (MCMC) or Variational Inference(VI). This complexity can be a barrier to widespread adoption, especially in resource-constrained environments where computing resources are limited. The complex nature of Bayesian inference necessitates careful consideration of model architecture and training methodologies, leading to longer training times compared to traditional neural networks. As a result, practitioners may face difficulties in deploying BNNs in real-time applications where speed is crucial (Anum, 2024).

Moreover, the computational demands of BNN can escalate as the dataset size increases. In scenarios involving large datasets or high-dimensional feature spaces, the training process can become prohibitively expensive, limiting the practical utility of BNNs in certain contexts. This challenge necessitates ongoing research into efficient training techniques and optimization algorithms that can minimize the computational burden while maintaining the integrity of the probabilistic framework (Naqash, 2022). Addressing these computational challenges is crucial to ensuring that BNNs can be effectively utilized in a wide range of applications, particularly in industries where timely decision-making is essential.

Scalability is another significant issue when it comes to BNN. While they excel in smaller datasets and controlled environments, scaling BNNs to accommodate larger datasets or high-dimensional spaces presents inherent challenges. As the amount of data increases, the training time and computational resources required can grow exponentially, potentially making BNNs impractical for large-scale applications. This limitation may hinder their adoption in industries where huge amounts of data are generated daily, such as energy management, finance, and healthcare (Anum, 2024).

The complexity of choosing appropriate priors for weight distribution adds another layer of difficulty to scaling BNNs. In high-dimensional spaces, the choice of priors can significantly impact model performance, and finding meaningful priors becomes increasingly challenging. As practitioners grapple with these challenges, there is an urgent need for research to develop scalable BNN architectures and efficient training methodologies that can handle the demands of large datasets without degrading performance. By addressing scalability issues, BNNs can unleash their full potential and make them viable options for a wider range of applications.

The selection of priorities for weight distribution in BNN is a critical aspect that can significantly affect model performance. Choosing meaningful priorities is not only non-trivial but can also be a source of frustration for practitioners. Poorly chosen priors can lead to suboptimal results, undermining the advantages that BNNs offer in terms of uncertainty quantification and robust predictions (Zulfiqar, 2024). This challenge underscores the significance of domain knowledge and expertise in the model-building process, as the effectiveness of BNNs may depend on the appropriateness of the priors used.

Additionally, the complexities associated with prior selection may deter practitioners from fully adopting BNN in their modeling efforts. In cases where prior knowledge is limited or uncertain, the risk of mis-specifying priors can pose significant problems, leading to unreliable predictions and reduced confidence in model outputs (Mahajan A. D., 2024). To alleviate these concerns, further research is needed to develop adaptive preselection techniques that can dynamically adapt to data and problem contexts. By strengthening the pre-selection process, BNNs can become more accessible and effective tools for practitioners in various fields.

2.13 Random Forest Concept

Random Forest is a robust machine learning method widely employed for both classification and regression tasks. It functions by constructing a multitude of decision trees during the training phase and outputs the mode of the classes for classification or the mean prediction for regression tasks. This model leverages the strengths of multiple decision trees to enhance overall predictive performance while minimizing the risk of overfitting (Wang Y. Z., 2023). One of the defining characteristics of Random Forest is its use of Bootstrap Aggregating (Bagging). This technique involves training each decision tree on a random subset of the training data, which is sampled with replacement. This sampling method helps to reduce variance and improve model stability. Random Forest introduces feature randomness; at each split in the tree, only a random subset of features is considered. This approach promotes diversity among the trees, which is crucial for the effectiveness of the ensemble method. Strengths and Weaknesses of Using Random Forest for Time Series Prediction

2.13.1 Advantages of Random Forest

Random Forest is less prone to overfitting compared to individual decision trees due to its ensemble nature. By aggregating predictions from multiple trees, it can generalize more effectively to unseen data (Wang Y. Z., 2023). This characteristic is particularly beneficial in time series forecasting, where overfitting can significantly degrade model performance.

The model effectively captures non-linear relationships in the data, which are common in time series datasets. Random Forest can model complex interactions between variables without requiring extensive preprocessing or transformation (Ahmad A. R., 2025). Random Forest provides insights into feature importance, enabling practitioners to identify which variables have a significant influence on predictions. This is particularly useful in time series analysis, where understanding the impact of different factors can lead to more informed decision-making (García-Alvarado, 2025). The model is versatile and can be applied to various types of time series data, whether univariate or multivariate. It can handle different data distributions and structures, making it suitable for diverse forecasting tasks (Rajawat, 2023).

2.13.2 Weaknesses of Random Forest

While Random Forest provides feature importance metrics, the model itself is often considered a "black box." This lack of interpretability can be a drawback in fields where understanding the decision-making process is crucial. Users may find it challenging to explain predictions to stakeholders.Training a Random Forest model can be computationally intensive, especially with large datasets and a high number of trees. This may lead to longer training times and increased resource consumption, which can be a concern in time-sensitive applications (Zhou Z. Y., 2025).

Random Forest models typically struggle with extrapolation beyond the range of the training data. In time series forecasting, this limitation can be problematic when predicting future values that lie outside the historical data range (Spiliotis, 2022). As a result, careful consideration is needed when applying Random Forest to long-term forecasting tasks .

2.14 Hybridization Concept in Machine Learning

Hybridization in machine learning refers to the combination of multiple models or techniques to leverage their strengths and mitigate their weaknesses. This approach aims to enhance predictive performance, improve robustness, and achieve better generalization across various tasks. Hybrid models often integrate different algorithms, including traditional statistical methods, machine learning techniques, and optimization strategies, to create a more effective solution (Wang X. W., 2023).

There are several types of Hybridization, such as Model Hybridization, which involves combining different machine learning algorithms. For instance, integrating decision trees with neural networks can harness the interpretability of trees and the powerful feature extraction capabilities of neural networks. By combining different approaches, hybrid models can be more resilient to changes in data distribution or noise, leading to improved generalization (Wang X. M., 2022). The Feature Hybridizationapproach combines features from different sources or types of data. For example, integrating time series data with categorical variables can provide a more comprehensive view for prediction tasks. Hybridization enables the adaptation of models to various types of data and tasks, making it a versatile approach in machine learning applications (Wellens, 2023). Algorithmic Hybridization: This involves merging different optimization techniques. For instance, combining genetic algorithms with local search methods can enhance the optimization process in training machine learning models (Xiao, 2022).

2.15 Bootstrap Aggregating Technique

Bagging, short for Bootstrap Aggregating, involves training multiple instances of the same algorithm on different subsets of the data, each subset being sampled with replacement. This method aims to create a diverse set of models that can collectively improve forecast accuracy. Once individual models are trained, their predictions are aggregated typically through averaging for regression tasks or majority voting for classification tasks. The basic principle of bagging is to reduce variance by averaging the errors of individual models, which leads to more stable and reliable predictions (Dube, 2023).

2.15.1 Advantages of Bagging

One of the main advantages of bagging is its ability to significantly reduce variance, which is particularly beneficial for high-variance models such as decision trees. By training on different subsets of the data, bagging reduces the impact of noise and outliers, resulting in a more robust model. This stability is key when forecasting power demand, where data can be noisy and unpredictable. A striking example of bagging in action is the Random Forest algorithm, which constructs a large number of decision trees during training and creates a way to predict them. The ensemble nature of Random Forest not only increases accuracy but also provides insight into the importance of features, making it a powerful tool for understanding complex datasets (Taieb, 2021).

2.15.2 Boosting Method

Boosting is an ensemble technique that focuses on sequential training models, where each successive model aims to correct the mistakes made by its predecessor. This iterative process involves assigning weights to predictions based on the confidence of each model, allowing the ensemble to focus more on hard-to-predict cases (Shah, 2022). The idea is that by combining weak learner models that perform slightly better than random guessing, boosting can create a strong overall model that excels in accuracy.

2.15.3 Advantages of Boosting

The primary benefit of boosting lies in its ability to improve accuracy by highlighting difficult cases that previous models have misclassified. This targeted approach allows the boosting algorithms to learn from mistakes, resulting in refined predictive power adaptively. Popular implementations of boosting include XGBoost, LightGBM, and CatBoost, each designed to efficiently process structured data while capturing complex feature interactions. These algorithms are particularly effective in energy demand forecasting, where accurate forecasting of peak demand scenarios can lead to better resource allocation and operational efficiency. By focusing on the most difficult examples, boosting not only increases accuracy but also contributes to a deeper understanding of underlying data patterns (Cardo-Miota, 2024).

2.15.4 Stacking Method

Stacking is an ensemble technique that combines predictions from multiple base models using a meta-model or higher-level model to produce a final prediction. In this approach, different models are trained independently, and their predictions serve as inputs to the meta-model. This hierarchical structure allows the suite to leverage the strengths of various modeling techniques and efficiently integrate both linear and nonlinear relationships present in the data (Banik, 2024).

2.15.5 Advantages of Stacking

One of the key advantages of stacking is its ability to leverage model diversity to enhance overall predictive performance. By utilizing various algorithms such as ANN, SVM, KNN, and gradient boosting models, a stacked file can capture a wide range of data characteristics. This diversity is particularly beneficial in energy demand forecasting, where different models can excel at different aspects of the data. For example, while ANNs are adept at capturing complex nonlinear relationships, SVMs can provide robust decision boundaries in high-dimensional spaces (Alghamdi, 2024). The meta-model learns to optimize the combined predictions, resulting in higher accuracy and more reliable predictions.

2.15.6 Voting Method

Voting is a straightforward ensemble technique that aggregates predictions from multiple models by majority voting for classification tasks or averaging for regression tasks. This method is straightforward to implement and does not require complex training procedures, making it a suitable option for many applications. In the voting ensemble, each model contributes equally to the final prediction, and a collective decision is made based on the aggregated outputs (Conte, 2024).

2.15.7 Advantages of Voting

The simplicity of the voting method is one of its primary advantages, as it offers a balanced solution that is both easy to understand and implement. By combining predictions from different models, voting helps mitigate the risk of overfitting that can occur in individual models, leading to more generalized predictions. This technique is particularly useful in energy demand forecasting, where the integration of different models can capture a range of trends and behaviors in the data. The collaborative nature of voting ensures that the ensemble remains robust to noise and variability, ultimately increasing prediction accuracy (Aurangzeb M. W., 2024).

2.15.8 Blending Method

Blending is a simplified form of compounding that involves keeping part of the data for training the meta-model, while using the rest to train the underlying models. This approach simplifies the model integration process and makes it easier to implement compared to traditional stacking methods. The key idea behind interpolation is to leverage the strengths of different models while minimizing the risk of overfitting that can arise when using the entire dataset to train both the base and metamodel (Siddiqui, 2022).

2.15.9 Advantages of Blending

One of the main advantages of blending is its ease of implementation, which allows professionals to create and evaluate file models quickly. By leaving out some of the data, blending reduces the likelihood of overfitting and ensures that the meta-model generalizes well to unseen data. This approach is particularly effective in forecasting energy demand, where rapid adaptation to new data can significantly influence operational decisions (Hasan, 2024). By combining forecasts from different models—such as ANN, SVM, KNN, XGBoost, CatBoost, and BNN blending increases the overall robustness of forecasts and provides a comprehensive view of potential demand scenarios.

The ensemble methodology of combining different machine learning models, such as ANN, SVM, KNN, XGBoost, CatBoost, and BNN, increases the accuracy and robustness of the hybrid energy consumption prediction model. By leveraging the unique strengths of each algorithmand integrating their predictions, the ensemble benefits from their individual strengths while mitigating their respective weaknesses (Sheikh, 2024). This approach is particularly effective for the challenges of estimating uncertainty and input perturbations in time series energy consumption data.

2.16 Research Gap

Table 2

Research Gap

Existing Research Papers	Description & citation	Research Gap
Deep ensembles method	(Ren, 2022) Deep ensembles method	The study did not evaluate the proposed method in real-world applications, so it is unclear how effective it would be in practice. Additionally, the study did not assess the computational cost of using deep ensembles, which could be a limiting factor in certain applications.
Bayesian approach	(Jeon, 2022) a Bayesian approach for Uncertainty Estimation that improved the robustness of neural networks against adversarial attacks	The study did not investigate the computational cost of using the Bayesian approach, which could be a limiting factor in some applications. It did not factor in the transferability and generalization challenges.
RBFNN, NSFM, TS hybrid approach	(D. Karamichailidou, 2021) Wind turbine power curve modeling using radial basis function neural networks and tabu search	The proposed model can be utilized to handle fault detection and diagnostics. It did not show how it factored in the transferability and generalization challenges.
AUniversal adversarial perturbation approach	(Liu, 2019) A Universal adversarial perturbation via prior-driven uncertainty approximation	Existing unsupervised methods fail to leverage model uncertainty to produce robust perturbations.
Artificial Neural network.	(Zhang J. Z., 2022) Diffusion kernel attention network for brain disorder classification	Diffusion problems can make neural networks sensitive to out-of-distribution data, causing unreliable predictions
Certified Robust Control under Adversarial Perturbations approach.	(Yang, 2023) A study on Certified Robust Control under Adversarial Perturbations	The weakness noted in this study is that Certified robust control designs often make assumptions about the behavior of the adversary, such as the type and strength of the perturbation they will introduce. However, these assumptions may not hold in all scenarios, which can limit the effectiveness of the control system.

A Learning to Learn Transferable Attack (LLTA) method for creating adversarial perturbations	(Fang, 2022)	The study did not evaluate the LLTA method on a more diverse set of datasets and models to further confirm its generalization capabilities.
"Optimism-Based Adaptive Control of Linear-Quadratic Systems"	(Faradonbeh, 2020)	The researchers focused on the development of an adaptive control framework for linear-quadratic systems. Further research is needed to analyze the robustness of systems or models to input perturbations.
Uncertainty Estimationfor Deep Regression via Perturbation-Based Ensembles	A research study by (Chandak, 2023)	The research study acknowledged that there is potential for improvement in terms of scalability and computational efficiency.
Impact of adversarial attacks on the uncertainty estimates produced by deep learning models	study done by (Zhang S. C., 2021) .	The study also recommended that future studies design a method to improve the robustness of uncertainty estimates by augmenting the training data with adversarial examples.
Adaptive Moment Iterative Fast Gradient Method (Adam-FGSM)	The study involved moment estimations of the gradients (Xu, 2020).	There is also a need to evaluate the generalizability and robustness of Adam-FGSM across different architectures and datasets, as well as its portability to various models and datasets.
Local Differential Privacy (LDP),	This practice safeguards the privacy of both users and the aggregator against the exposure of confidential information. (Wang, et al., 2019)	Despite the growing research interest in LDP, most current efforts are not directed towards implementing LDP in scenarios involving intricate data and/or analytical assignments.
Robust Bayesian Neural Network (RBNN)	(Kaur, 2021)	The challenge with this study is that it did not consider extending the evaluation of RBNN to include more challenging datasets with severe outliers, noise, or adversarial attacks to further assess its generalizability.
Deep Learning Models approach using adversarial perturbations	(Zhang, 2021) Study on Uncertainty Estimation in deep learning models using adversarial perturbations	The study evaluated the quality of their uncertainty estimates using metrics such as the expected calibration error and the maximum calibration error. However, there is still a need to evaluate the usefulness of these uncertainty estimates for downstream tasks, such as decision-making under uncertainty.

[

2.17 Research Gap

Recent research in power demand forecasting has predominantly focused on improving predictive accuracy, particularly through deep learning models. However, a notable gap exists in studies that systematically integrate uncertainty estimation within forecasting models, especially under conditions of input perturbations such as demand fluctuations, measurement noise, and weather-related variability. While several studies have explored techniques to enhance the robustness of forecasting models, many do not incorporate mechanisms that quantify how reliable the predictions are when exposed to real-world data disturbances.

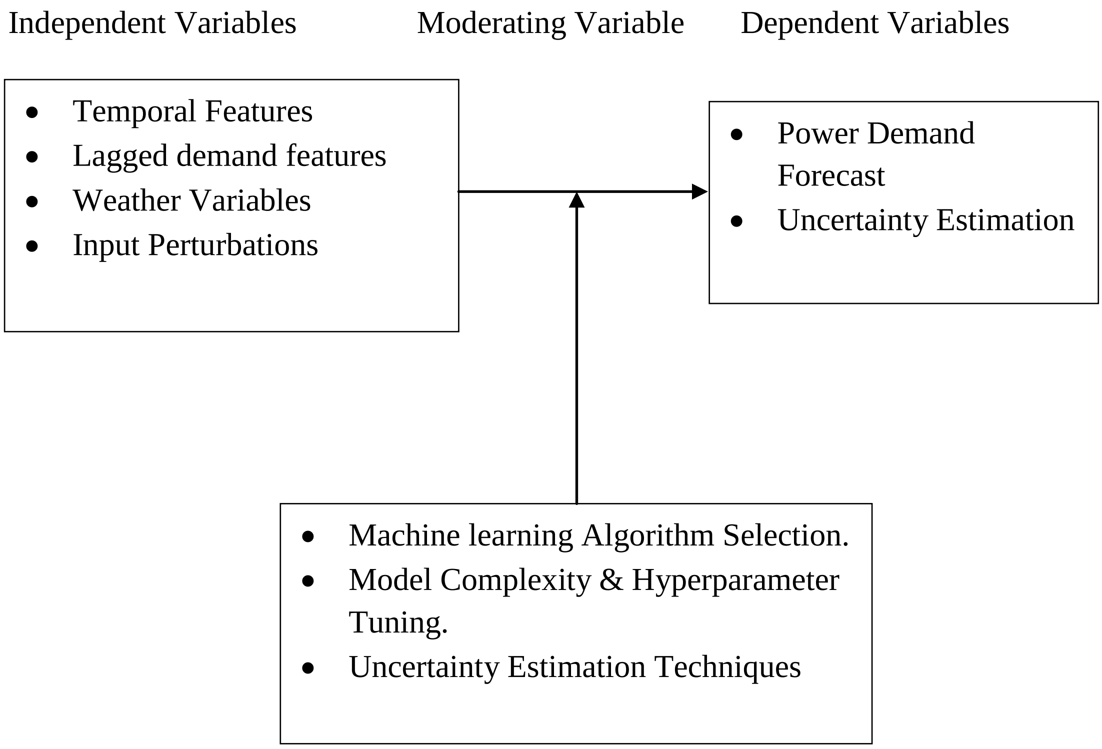
Furthermore, existing models often demonstrate limitations in generalizability, as they tend to perform well on clean experimental datasets but experience performance degradation in operational environments where data irregularities are common. Limited research has also been conducted in the context of developing economies, where grid instability, data inconsistencies, and environmental volatility heighten the need for uncertainty-aware forecasting approaches.

This gap highlights the need for a hybrid power demand forecasting model that not only enhances prediction accuracy but also integrates uncertainty estimation to improve reliability under input perturbations. Addressing this gap, the present study proposes the development of a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations (HPDEF-MUIP), using XGBoost, CatBoost, and Random Forest. The model aims to offer a practical, reliable, and context-appropriate forecasting solution suitable for real-world smart-grid environments.

2.18 Proposed Conceptual Framework

Figure I

Conceptual Framework



Source: Author, (2025)

The conceptual framework for this research study involved the examination of the impact of input perturbations on Uncertainty Estimation in deep learning models for Electricity Demand Forecasting.

CHAPTER THREE

RESEARCH METHODOLOGY

3.1 Introduction

This chapter discusses the methodologies used in collecting both primary and secondary data for the assessment, design, implementation, and evaluation of a Hybrid Power Demand Model with Uncertainty Estimation Under Input Perturbations for the Enhancement of Accuracy in Power Demand Models.

3.2 Research Philosophy

The ontology philosophy was used in this study because it provides a valuable lens through which to examine the hybridization methodology in machine learning. By focusing on the nature of entities, their relationships, classifications, and the dynamic construction of knowledge, ontology enriches our understanding of how diverse models can be effectively combined to enhance predictive accuracy and robustness. The philosophy of machine learning models delves into fundamental questions about how these systems work, what they can and cannot do, and the profound implications they have for our understanding of knowledge, intelligence, and the world around us. In the context of a hybrid methodology for forecasting power demand, this philosophy offers a nuanced lens for evaluating and constructing forecasting models, integrating interdisciplinary insights from computer science, statistics, and philosophy to address real-world challenges (Esnaola-Gonzalez, 2021).

At the heart of the machine learning ontology philosophy is the "no free lunch theorem," which emphasizes that no model is universally optimal for all problems. This notion forces researchers and practitioners to consider a variety of approaches tailored to specific problem areas. When forecasting power demand, the inherent variability and complexity of time series data require a hybrid approach (Zhao H. L., 2024). By combining models such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), K-Nearest Neighbors (KNN), XGBoost, CatBoost, and Bayesian Neural Networks (BNN), the hybrid model aligns with this philosophy, ensuring robustness and adaptability. This diversity enables the system to navigate changing conditions, ranging from linear trends to nonlinear anomalies, underscoring that intelligence is inherently situational and context-dependent.

The hybrid approach also resonates with the concept of the "Wisdom of Crowds," a principle rooted in the idea that collective decision-making often trumps individual judgment. Applied to machine learning, this principle suggests that aggregating predictions from different models can mitigate individual bias and increase overall accuracy (Porcello, 2022). Philosophically, it reflects the interdisciplinary nature of knowledge, where different perspectives converge to create a more complete understanding. In the hybrid model, the Artificial Neural Network (ANN) provides insight into nonlinear patterns, the Support Vector Machine (SVM) stabilizes decision boundaries, and the Bayesian Neural Network (BNN) quantifies uncertainty. Together, they form a collective intelligence system capable of navigating uncertainties and perturbations in energy consumption data.

Uncertainty is a fundamental aspect of the real world, and addressing it is crucial to the philosophy of machine learning. Probabilistic interpretivism advocates models that not only provide deterministic outputs but also quantify uncertainty (Rieder, 2024). This is consistent with the inclusion of Bayesian Neural Networks (BNNs) in hybrid models. BNN not only forecasts power demand but also provides probability intervals and offers insights into the reliability of these predictions. Philosophically, this reflects a Bayesian view of knowledge where uncertainty is not a limitation but a feature of how we understand the world. Incorporating an uncertainty estimate increases the reliability of the hybrid models and provides useful information about the input perturbations.

The hybrid philosophy in machine learning emphasizes the integration of different methods to achieve greater system capabilities. This reflects the interdisciplinary collaboration that defines the philosophy of science, where different disciplines contribute unique perspectives to solving complex problems. The hybrid methodology illustrates this by integrating the ability of ANNs to model complex, nonlinear relationships, the robustness of SVMs in high-dimensional spaces, the local adaptability of KNNs, and the capacity of boosting algorithms, such as Random Forest, XGBoost, and CatBoost, for feature interaction. Philosophically, this approach promotes the belief that intelligence and knowledge arise from collaboration and synthesis rather than isolation.

Philosophy often combines theory and practice, and empirical pragmatism is a reflection of this in machine learning. The hybrid model for forecasting power demand is deeply rooted in empirical evidence, ensuring its practicality and applicability to real-world problems. Using historical data, the model adapts to patterns and anomalies, demonstrating a pragmatic approach to knowledge application. This philosophy emphasizes that machine learning systems are tools designed to serve specific purposes, with their value measured by their effectiveness in solving tangible challenges such as managing power grids under uncertainty.

The hybrid methodology employed in this study raises broader philosophical questions about the nature of knowledge and intelligence. It suggests that intelligence, whether human or artificial, thrives on diversity, adaptability, and the ability to reason under uncertainty. The hybrid model embodies these principles and shows that machine learning systems are not static algorithms, but dynamic entities capable of synthesizing different insights. By embracing the complexities of predicting energy demand, the hybrid model highlights that understanding and navigating uncertainty are hallmarks of true intelligence.

The philosophy of machine learning models provides a basic framework for understanding and designing a hybrid methodology for forecasting Power Demand. It integrates principles such as the "No Free Lunch Theorem" (Zhao H. L., 2024), "The Wisdom of the Crowd" (Porcello, 2022), Probabilistic Interpretivism and Hybridization Philosophy, based on empirical pragmatism. Together, these philosophies emphasize the interplay between theoretical rigor and practical application, offering a blueprint for creating intelligent systems that are not only accurate but also robust and insightful. Through this lens, the hybrid methodology transcends its technical implementation. It becomes evidence of the evolving relationship between machine learning and our understanding of intelligence and uncertainty in the real world.

3.3 Quantitative Research Design

This study applied the Quantitative Research methodology.The research approach focused on explaining the phenomenon of power demand uncertainty through the collection and analysis of numericaldata using statistical methods(Mohajan, 2020). Quantitative research utilizes numbers, measurement quantities, and patterns to quantify information. Quantitative Research was applicable to this study, as it facilitated an organized inquiry into power demand forecasting with uncertainty estimation through the collection of numerical data and the execution of computational techniques(Kumari, 2023).

3.4 Meta-analysis Literature Review Research Design

This study employed a meta-analysis literature review design to achieve specific objective four by comparing the accuracy of results with those of previous studies. A meta-analysis researchdesign is a study that integrates and combines the results of multiple studies into a single comprehensive analysis. The majority of the meta-analysis literature focuses on data analysis and statistical discussions(Paul, 2022). Existing literature and secondary data were reviewed to examine and identify the weaknesses of the Power Demand forecasting model (Mathew, 2021). The weaknesses identified in the literature review section were used to inform the design of this study'sHybrid Power Demand Forecasting Model with Uncertainty Estimation Under Input Perturbations.

The scope of this section was to review existing hybrid models for power demand forecasting. The reviewed literature consisted of only studies published within the last five years from the current date. The literature review was conducted on academic papers, specifically peer-reviewed journals. The first research question was addressed by assessing the weaknesses of existing hybrid power demand forecasting models with uncertainty estimation under input perturbations. During the literature review, conference papers and presentations were excluded from the study. Only research papers and review papers were included and considered in the study.

We achieved a high-quality literature review by evaluating the paper quality, publishers, and journal quality. The study investigated the papers found in the directory of open-access journals, checked for a good Scopus quartile ranking, ensured the paper was only published by credible publishers, and verified that the paper was from a reputable journal and not a predatory publisher. The study reviewed the existing papers' abstracts and conclusions to determine their relevance. We checked against inclusion and exclusion criteria. Conference papers and presentations were excluded from the review. Only published research papers and peer-reviewed papers in power demand forecasting were considered in reviewing the quality of the literature.

A detailed assessment of selected and relevant research papers was carried out. The study identified key themes, findings, methodologies, and gaps. This enabled the creation of a review matrix that was instrumental in organizing and comparing the information.

3.5 Deep Learning Experimental Research Design

The studyappliedan experimental quantitative research approach where there is manipulation of independent variables to establish a cause-and-effect relationship between independent and dependent variables(Gaur, 2009)(Jain, 2023)(Kumatongo, 2021).An experimental research design was used to address the second research question: How will the HPDEF-MUIP model be designed? This was because Deep learning models are known to have good learning abilities that enable the attainment of good model performance (González-Díaz, 2021). Deep Learning modelsutilize artificial Intelligence Techniques to enhance the accuracy of models (Cho, Shin, Kim, Jeong, & In, 2020).This Experimental research was important in providing a rigorous and systematic way to train and evaluate model performance, thereby enhancing the accuracy of the forecasting models.

Input perturbation analysis was performed, which introduced variations into the input data, helping to train and evaluate the model's performance. Calculation of prediction intervals was done during the experiments. The experimental results indicated the range within which the actual power demandwas likely to fall.

In the experiments, an ablation study was done to analyze the effect of individual components of the hybrid model on prediction accuracy. By removing or modifying specific elements, the contribution of each component was assessed in the model. A comparative analysis of the separate performance of the techniques integrated into the hybrid model further helped to evaluate its effectiveness. Statistical and error analysis were conducted using statistical tests, such as MSE, to compare the performance of the hybrid model with the baseline models and determine the significance of observed differences. Analysis of residuals helped identify systematic bias and ensured that they were randomly distributed. Outlier detection and analysis were performed to provide insights into the limitations of model experiments by examining inconsistencies in predictions.

A Comprehensive documentation of the experimental setup wascreated, including data sources, preprocessing steps, and model configurations, which was essential for reproducibility. Version control for datasets and code was done to ensure that changes are tracked and reproducibility is maintained. The process followed the sequence of steps as shown in Figure 2 below.

Figure 2

Experimental Research Design

Figure 2 illustrates the deep learning architectural experimental research design, beginning with research study data collection and preprocessing, then proceeding through validation, experimental design, feature selection, evaluation, model training, and, based on the model evaluation, hybrid model design and evaluation.

3.5.1 Data Collection and Pre-processing

The dataset for this study was collected from a central smart meter for the Kenya Power and Lighting Company's Central Rift Region, and weather data was collected from the Nakuru Meteorological Station. The Central Region was preferred for data collection because it had both small and large power consumers, as well as experienced dynamic weather conditions thatwereuseful in carrying out the study. The dataset collected consisted of 14,094 records, all stored in an Excel file. The Power demand data for the six feeder lines serving six regions is supplied from the KPLC Lanet Substation. The variables wereLine Voltages, four 11 kV feeder transformers, five 32 kV feeder transformers, and power demand in Megawatts.

Hourly temperature data were collected using Excel sheets from Nakuru Meteorological Station. Both Primary and Secondary data for the period between January 2018 and October 2024 wereused in the study. The collected dataset was pre-processed to ensure the data was good for use. Before entering the data into the model, the collected data wascleaned, missing values were handled, outliers and anomalies were identified, and extreme values were mitigated. Data augmentation techniques were applied to time series data to increase their variability and improve the generalizability of machine learning models. These techniques included seasonality, noise embedding, dropout, smoothing, feature scaling and normalization, as well as dimension reduction (Salehin, 2023).

Normalization was done byscaling the input data to a standard range (e.g., between 0 and 1). This helped mitigate the impact of large variations in the data. Data normalization was used in conjunction with input perturbation techniques to enhance the accuracy and generalization of machine learning models (Huang L. Q., 2023). The normalization method used in the study was Minimum and Maximum scaling (function scaling). This function scaled the element values to the range [0, 1] based on the minimum and maximum values in the dataset. Historical power data was collected for the previous two years. This data was thensplit into a 60% training set, a 20% validation set, and a20% testing set.Logarithmic transformation was applied to stabilize the data and make it more symmetrical. This process enhanced model accuracy.

3.5.2 Feature Selection Engineering

A feature selection method was employed toselect a set of relevant features that were less prone to failure. This process involved selecting elements with stable patterns over time. Feature engineering of time series data in our study involved creating meaningful and relevant features that captured temporal patterns, trends, and dependencies. An embedded method together with feature ranking was used in the study. Robust feature engineering for time series data was critical in building a model thatgeneralizes well across time periods and conditions(Vinayakumar, 2019). This was achieved by utilizing delay features, which created delay functions by incorporating past values of the target variables. This was necessary to capture trends and autocorrelation in a time seriesof data.

3.5.3 Model Training

Training the base models, including the Artificial Neural Network model, Bayesian Neural model, and Support Vector Machine model, was done using the preprocessed data. Techniques such as cross-validation and hyperparameter tuning were utilized to optimize model performance. Predictions from individual models wereaggregated using averaging or voting methods. This was seen to help mitigate the impact of failures.

3.5.4 Bayesian Neural Network Training

Bayesian Neural Networks were trained by employing Bayesian principles, which estimate weight uncertainties and adapt to input perturbations. The Monte Carlo method was used because it addresses the challenges of recursive procedures not terminating due to multiple paths between network nodes(Izmailov, 2021). A clustering method was also considered in training the model. Temporal Attention Mechanisms for Time Series Clustering was incorporated into deep learning architectures. This is because itallows the model to attend to different temporal segments during the clustering process.

A detailed algorithm for designing a Bayesian Neural Network (BNN) for a hybrid energy demand forecasting model with input disturbance uncertainty estimation was developed. This BNN training focused on using time series data for intraday forecasting and efficient uncertainty quantification.

Algorithm for BNN Mathematical Model Design

i. Definition of inputs

X: Input function (time series data for energy consumption, weather conditions, etc.).

y: Target variable (actual power consumption).

P(W): Preliminary distribution of model weights.

P(y|X,W): Probability of the data given the weights.

ϵ: Input perturbation noise, representing real-world uncertainties.

ii. Bayesian framework

Defined rear weight distribution W using Bayes' theorem:

$P(W \mid X, y) = \frac{P(y|X,W)P(W)}{P(y|X)}$ .... (25)

Where:

P(W) -Encoded prior knowledge about model weights.

P(y|X,W)-Probability representing how well the model fits the data.

P(y|X)-Proof term, acted as a normalization constant.

The numerator $P(y \mid X, W)P(W)$ expresses the balance between model fit (likelihood) and prior belief (regularization). The denominator $P(y \mid X)$ was computed by integrating over all possible W, which is intractable for neural networks. Since $P(y \mid X)$Intractable variational inference or Monte Carlo approximation is used to estimate the posterior during Bayesian Neural Network training.

iii. Formulation of the model:

BNN is considered a neural network with stochastic weights

$\hat{y} = f(x; w)$

Where:

f- is a network function parameterized by W.

W is sampled from a learned posterior distribution rather than optimized to a single point estimate.

Perturbation $\epsilon$is introduced to the inputsas $X^{\text{perturbed}} = X + \epsilon$

Variational derivation for posterior approximation. Approximation of the posterior P(W|X,y) using the variational distribution q(W). Min KL(q(W)‖P(W|X,y).The term KL-divergence is extended as:KL(q(W)‖P(W))−E q(W) [logP(y|X,W)]. The lower bound of evidence (ELBO) is used as a target. L=E q(W)[logP(y|X,W)]−KL(q(W)‖P(W)). $L$is maximized during training.

v. Monte Carlo approximation:

The approximate expectation is over q(W) using Monte Carlo sampling:

$L \approx \frac{1}{S}\sum_{s=1}^{S} logP(y \mid X, W\ ) - KL(q(W) \parallel P(W))$ ....(26)

Where:

L is the functionoptimized when training a Bayesian Neural Network using variational Inference. Its goal is to find an approximate posterior $q(W)$. That is as close as possible to the true (but intractable) Posterior P(W|X,y).

$^{(s)}$ is the S^{th} sample of the model weights drawn from the variational posterior

$\sim$ means “is sampled from” or “follows the distribution”

(W) = the variational posterior distribution over weights

S = the number of Monte Carlo samples used to approximate the expectation

$logP(y \mid X, W\ )$ = log likelihood of the data given sampled weights

$KL(q(W) \parallel P(W)$ = Kullback-Leibler divergence between approximate posterior and prior.It prevents overfitting, keeps weights close to the posterior.

L is the cost (or score) we want to maximize so that the model learns useful weight distributions that fit the data well but do not overfit

vi. Prediction with uncertainty:

a. For a given disturbed input $X^{\text{perturbed}}$, obtain the predictive distribution by marginalizing over the posterior. $P(y^* | X^{\text{perturbed}}) = \int P(y^* | X^{\text{perturbed}}, W) P(W|X,y) dW$

b. Approximate use of Monte Carlo samples was given as:

$\hat{y} \approx \frac{1}{s}\sum_{s-1}^{s} f\left(x_{P_j}, w^{(s)}\right)$ .... (27)

where:

$Ws \sim P(W|X,y)$.

X is perturbed (e.g., noise, fluctuations, or uncertainty in power demand drivers like temperature, load history, or industrial activity).

WS: This covers uncertainties in the model's parameters, such as the weights in the artificial neural network (ANN) or how the boosting trees are split.

y^*:This gives a strong forecast of future power demand that takes into account both model and input uncertainty.It helps grid operators make better decisions.

c. Get an estimate of the uncertainty as the variance of the predictions

Uncertainty =1$S\sum s$=1(($X^{\text{perturbed}}$; $Ws)^2$

vii. Integration into the hybrid model:

BNN was used to estimateAleatory uncertainty (from input noise ϵ) and Epistemic uncertainty (from model weight distributions W). BNN predictions and uncertainties were combined with other models (ANN, SVM, KNN, XGBoost, CatBoost) in a hybrid modeling:

$y^{\wedge hybrid} = \sum_{i=1}^{N} \alpha i y^{\wedge} i \ y^{\wedge} i$ ....(28)

Where:

αi-was the weight of each model,

$y^{\wedge}{}_{i}$-was the output of the model i

viii. Output.

Power demand forecast for the day $y^{\wedge}$hybrid with associated uncertainty bounds.

i. $y^{\wedge}$lower = $y^{\wedge}$hybrid−2 x Uncertainty

ii. y^upper = y^ hybrid +2x Uncertainty

The BNN algorithm defined the mathematical foundations of a BNN model designed for power demand forecasting. It integrated uncertainty estimation and input bias, making it a robust part of a hybrid forecasting model.

3.5.5 Support Vector Machine (SVM) Optimization

Support Vector Machines were optimized through margin maximization and kernel selection, thereby enhancing their performance in perturbed data scenarios. Stochastic Gradient descent optimization was used to find the optimal decision boundary. The performance of the perturbed SVM model wasevaluatedon unseen test data using appropriate metrics (e.g., precision, accuracy, recall, F1 score).

The model performance of the perturbed model was compared with that of the original SVM base model to determine the effectiveness of the perturbation technique. The SVM optimization process was based on minimizing the deviation of the predictions from the actual values within a specified tolerance.

i. SVM for Regression (SVR)

In SVR, the goal was to find a function (x) which approximates the relationship between the input x and output y, while balancing complexity and error within tolerance ϵ.

Function f(x) was defined as:

$(x) = w^{\top}(x) + b$ .... (29)

Where:

(x) -Mapping function to transform the input x into a multidimensional feature space.

w -was the Weight vector.

B – was the Bias term.

ii. The goal of optimization

The goal of optimization is the Loss function.The optimization problem minimizes structural risk:

$min/w, b, \xi, \xi * \frac{1}{2} \|w\| 2 \sum_{i=1}^{n} (\xi i \ + \xi i \, *)$.. (30)

Subject

$yi - (w \top \phi(xi) + b) \leq \epsilon + \xi i,$

(w ⊤ϕ(x)+b) -yi ≤ϵ+ξ

≤ϵ+ξ ≥0.

Where:

ϵ - Tolerance margin within which predictions are not penalized.

$\xi_i, \xi_i{}^*$- Slack variables capturing forecast errors outside ϵ margin

C - Regularization parameter controlling the trade-off between margin width and error tolerance.

iii. Kernel Trick for Nonlinearity

To handle nonlinear relationships, SVMs used a kernel trick that implicitly computed the dot product in a high-dimensional feature space:

$(x_i, xj) = (xi) \top \phi(xj)$ .... (31)

Common cores included:

a. Linear Kernel

   $K(xi,xj) = x_i{}^{\top} x_j$

b. Polynomial Kernel:

   $K(x_i,x_j) = (x_i{}^{\top} x_j + c)^d$

c. Radial Basis Function (RBF/ Gaussian) $K(x_i, x_j) = exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right)$

   Where;σ = kernel width (controls spread)

Exp- exponential function

iv. Optimization with a dual problem

The SVM optimization problem was solved in its dual form for efficiency as:

$$\max_{\alpha_i \alpha_i^*} \left[-\frac{1}{2} \sum_{l-1}^{n} \sum_{J-1}^{n} (\alpha \cdot i - \alpha_j^*)(\alpha_j - \alpha_j^*) K(X_i, x_j) + \sum_{i=1}^{n} (\alpha_i + \alpha_i^{0^*}) \right] \text{(32a)}$$

subject: $0 \leq \alpha i, \alpha i * \leq C, \sum_{i=1}^{n} (\alpha i - \alpha i *) = 0$ ....(32b)

Here:

α_i, $\alpha i*$: Lagrange multipliers corresponding to the constraints.

$\alpha_i$represents error on the upper side of the ϵ- tube. In the ϵ Support VectorRegression (ϵ-SVR) formulation $\alpha i*$ - (read as alpha-i-star) is a Lagrange multiplier associated with the upper constraint of the ϵ- intensive loss function.

α, represents how much the model overshoots the target $\alpha i*$

3.5.6 Artificial Neural Network (ANN) Learning

The Artificial Neural Networks base model was trained using backpropagation, adapting to intricate non-linear patterns in the data, including those arising from perturbations.The resulting ANN model trained with the perturbed architecture was evaluated on both clean and perturbed test data to assess its accuracy. The learning process for ANNsinvolved training the network to approximate a target function or make predictions from input data. A sigmoid activation function was used in the hidden layers because it is useful in binary classification tasks (Ballesteros, 2020).ANN base model Algorithm steps followed were:

i. Defined the input data

Let $X = \{xt\}_{t=1}{}^{T}$ represent the input properties of the time series:

$x_t = [x_1, x_2 \cdot \cdots x_n] t \in \{1, 2, \dots, T\}$

Where:

n -was the number of features (e.g., temperature, time of day, and historical power demand).

ii. Adding perturbation to estimate the uncertainty

Introduction of Gaussian noise $\epsilon \sim N(0, \sigma\epsilon 2)$ to simulate variability:

$$x_t perturbed = x_t + \epsilon \text{ .... (33)}$$

iii. Forward propagation of ANN

Input layer

The input vector $x_t^{\text{perturbed}}$ is fed to the ANN.

Hidden layers

For each hidden layer l, we calculated the weighted sum $z^{(l)}$ and activation $a_j^{(l)}$:

$$z_j(l) = \sum_{i=1}^{n} wij(l)ai(l-1) + bj(l) \text{ ....(34)}$$

$^{(l)}=(z_j^{(l)})$, $f(\cdot)$ is the activation function (eg, ReLU or sigmoid).

Output layer

To Predict Power Demand

$$\hat{y}^{(t+1)} = g\left(\sum_{j-1}^{m} w_j^{(0)} a_j^{(L)} + b^{(0)}\right) \text{ .... (35)}$$

Where:

L -was the number of hidden layers, and$m$was the number of neurons in the last hidden layer.

$\hat{y}^{(t+1)}$ -is the forecasted power demand at time t+1.

Summation from j=1 to m.This is a sum over all the neurons in the last hidden layer, called L.The index j starts from i (often 1) and goes up to m, which is the total number of neurons in that layer.This means the output is a combination of the contributions from all the neurons in the last hidden layer.

$w_j^{(0)}$-This is the weight connecting the j-th neuron in the final hidden layer L to the output neuron. It is learned during training. This weight determines the influence the j-th neuron has on the final prediction.

$a_j^{(L)}$-This is the activation of the j-th neuron in the last hidden layer L. It is calculated after applying an activation function (like ReLU or tanh) to the weighted inputs of that neuron.This activation captures a nonlinear transformation of the input features, such as historical load, weather, or calendar effects.

$b^{(0)}$-This is the bias term in the output layer. It shifts the prediction up or down, allowing the model more flexibility. It's essential that the activations are zero or balanced, which helps the model predict non-zero demand.

$\hat{y}^{(t+1)}$ -is the forecasted power demand at time t+1.

m- number of units in the last hidden layer

iv. Monte Carlo dropout due to uncertainty estimation

Dropout was applied during inference to the sample *K*

To predict;

$\{\hat{y}_{t+1}^{k}\}_{k=1}^{k}$ ....(36)

To calculate the average forecast

$\mu_{t+1} = \frac{1}{k}\sum_{k=1}^{k} \hat{y}_{t+1}^{(k)}$... (37a)

To forecast uncertainty forecast

$\sigma_{t+1} = \frac{1}{k}\sum_{k=1}^{k}(\hat{y}_{t+1} - u_{t+1})^2$ ....(37b)

v. Integration of the hybrid model

Artificial Neural Network (ANN) base model predictions are combined with predictions from an auxiliary model (e.g., KNN)

$\hat{y}_{t+1}^{hyb} = \alpha_{t+1}^{\hat{y}ANN} + (1-\alpha)\hat{y}^{aUx}$..(38)

Where:

α-is the weighting factor optimized during validation.

$\hat{y}_{t+1}^{hyb}$ = Hybrid model prediction at time t+1

$\alpha_{t+1}^{\hat{y}ANN}$ = Auxiliary weighting factor optimized at time t+1

$\hat{y}^{aUx}$= Auxiliary model predicted value (e.g., KNN)

$\alpha \in [0,1]$= Weighting coefficient tuned during model validation to optimize performance. The weighting factor α ensures optimal contribution from each model, improving accuracy and generalization through complementary strengths.

vi. Objective of training

The ANN was trained using backpropagation so as to minimize the mean squared error (MSE):$L = \frac{1}{2}\sum_{t=1}^{N}(y_t - \hat{y}_t)^2$..(39)

Where;

L = loss value (MSE)

3.5.7 Extreme Gradient Boosting (XGBOOST)

Extreme Gradient Boosting is a powerful and efficient machine learning algorithm that excels in predicting energy demand one day ahead(Champahom, 2025). Its ability to model complex interactions between different input features,such as historical energy consumption, weather data, and time-based variables,makes it particularly valuable. In order to effectively deal with forecasting uncertainties caused by input perturbations, methodologies such as Monte Carlo simulation and Quantile Regression were seamlessly integrated into the forecasting process.

XGBoost wasdesigned as a time series algorithm; it was suitably adapted for forecasting tasks by engineering strategic features and input transformations.To improve the predictive capabilities of the model, lagged variables were created using historical data. For example, when forecasting demand y_t for the next day, inputs includedy_{t-1}, y_{t-7}, the temperature at t−1, and indicators for weekdays or weekends. In addition, exogenous variables such as weather forecasts were incorporated to enrich the dataset(Shabbir, 2024). Time-based features, such as day of the week and month,were also included to capture regular demand patterns.

The model forecasted a single aggregate value for the next day, representing the total daily demand, or provided hourly forecasts for the entire day, which were summed to yield the total demand.A mathematical framework for XGBoost followed(Harikrishnan, 2025). The core of XGBoost lies in its objective function that optimizes the regularized loss function:

$$L^{(t)} = \sum_{L=1}^{n} L\left(y_L, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \quad \text{(40a)}$$

Where:

y_i -True demand for data point

y_i = true target value (actual demand) for data point i

$\hat{y}_i(t-1)$= model prediction from previous iteration $(t-1)$

$f_t(x_i)$ = Forecast from the newly added tth regression tree

$l(.)$ = differentiable loss function (e.g., Mean Squared Error for regression)

$y_i^{(t-1)}$ -Prediction from the previous iteration

$f_t(x_i)$ -Prediction from the *t-th* decision tree,

$l(y_i, y_i)$ -Loss function (e.g., Mean Squared Error for regression)$\Omega(ft)$

$\Omega(ft)$ -Regularization term that penalizes model complexity.

XGBoost applies regularization to prevent overfitting and promote generalization. The regularization term is defined as;

$$\Omega(f_t) = rT + \frac{1}{2}\lambda \sum_{+1}^{T} w_j^2 \quad \text{.......(40b)}$$

Where:

T - is the number of leaves in the tree,

w_j- is the weight of leaf j,

r – is the penalty for adding a leaf node

λ = L2 regularization parameter on leaf weights.

The first term ensures prediction accuracy, while the second term controls complexity, preventing overfitting. This balance enabled XGBoost to achieve high predictive power with robustness, making it superior to traditional boosting.

a. Gradient boost update

Predictions are iteratively refined using the formula:

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + \eta f_t(x_i) \quad \text{.......(41)}$$

Where;

η-is the learning rate

$\hat{y}_i^{(t)}$-This is the predicted value at step t for a specific sample. It's the total prediction after combining all the previous weak learners. It shows the updated power demand forecast after step t.

$\hat{y}_i^{(t-1)}$ -This is the prediction from the step before. At the very start (t = 0), it is typically a simple initial guess, such as the average of the target values in a regression problem.

$f_t(x_i)$ -This is the new weak learner added at step t. It is usually a decision tree. It is trained to identify and correct mistakes made by previous predictions. It focuses on the errors the previous model made and attempts to correct them.

$\eta \in [0,1)$ – This is the learning rate controlling the contribution of the new base learner.

The model began with an initial forecast of the mean target values for regression. At each iteration, a new weak learner f_t was trained to correct the errors (residuals) made by prior forecasts. The learning rate η ensured that updates were gradual and stable, preventing aggressive overfitting. Over successive iterations, the hybrid model forecast became more accurate as the model learned from its mistakes by gradually reducing residual errors and improving model generalization.

b. Gradient and HessianFactor

To facilitate and enhance optimization, transitions (g_i) and Hessians (h_i) were calculated for each data point:

$$g_i = \frac{\partial_l(y_i, y_i)}{\partial \hat{y}_i}$$

$$h_i = \frac{\partial^2 l(y_i, y_i)}{\partial \hat{y}_i^2}$$

Where;

g_i = first-order gradient (measures direction and magnitude of change needed)

h_i = second-order derivative (Hessian) (measures curvature to refine update steps

These values determined the split points and leaf weights in the decision trees.

c. Uncertainty estimation with input perturbations

Forecasting Power demand involves uncertainties arising from errors in input variables, such as weather forecasts or historical data. The Monte Carlo simulation technique was used to increase XGBoost's ability to estimate these uncertainties.Perturbing the input data by adding random noise simulates changes in the inputs. XGBoost was trained on these perturbed datasets, and predictions from multiple runs were aggregated to produce probabilistic predictions. The average forecast was calculated across N perturbed runs and is given as follows:

$$\bar{y} = \frac{1}{N}\sum_{J=1}^{N}\hat{y}_j \quad \text{...} \quad (42)$$

Where:

$\hat{y}_j$ = Forecast from the jth perturbed model run

$\bar{y}$ = aggregated (mean) forecast

Prediction intervals, such as a 95% confidence interval, are derived from prediction percentiles: $PI_{95\%}$ =[$P_{2.5}$, $P_{97.5}$], where Pk is the kth percentile of theforecast.

d.Quantile regression

The XGBoost target was modified to estimate quantiles directly, allowing a finer understanding of prediction intervals. The loss function for quantile regression was defined as:

$$l \propto (y_i, \hat{y}_i) = \begin{cases} \alpha(y_i - \hat{y}_i)' if y_i \geq \hat{y}_i \\ (1-a)(\hat{y}_i - y_L), if y_i < \hat{y}_i \end{cases} otherwise \text{ ....} (43)$$

where α and represented the quantile (e.g, 0.05 for the lower bound and 0.95 for the upper bound).

3.6 Design of LightGBM Model

LightGBM is a gradient boosting framework that sequentially builds decision trees to optimize a specific objective function(Munir, 2024). A modeldesign for time series forecasting is detailed as follows:

3.6.1 Objective Function

The objective function combined the loss term (ℓ) and the regularization term (Ω) to balance model accuracy and complexity:

$L = \sum_{i-1}^{n} l(y_i, \hat{y}_i) + \Omega(f)$.. (44)

Where:

$l(y_i, \hat{y}_i)$ - loss function measuring the prediction error and actual values.

Ω(f) - Regulating term to control overfitting.

3.6.2 LightGBM Loss Function

The Mean Squared Error (MSE),commonly used for regression problems,was represented as:

$l(y_i, \hat{y}) = \frac{1}{n} \sum_{i-1}^{n} (y_i - \hat{y}_i)^2$ ....(45)

Where:

y_i - Actual observed value,

$\hat{y}_i$ - the value predicted by the model.

3.6.3 Regularization Term

The term regularization,used to penalize the complexity of the model,was represented using the equation:

$\Omega(f) = \gamma^T + \frac{\lambda}{2} \cdot \sum_{j-1}^{T} w_j^2$ .... (46)

Where;

T - Number of leaves in the decision tree,

w_j - Weight of the j-thleaf,

γ - Penalty per leaf (checks the number of leaves),

λ - L2 regularization term for leaf weights.

3.6.4 Prediction in Gradient Boosting

LightGBM constructs decision trees sequentially, where each tree corrects the residuals of previous trees. Prediction at iteration m is shown as,

$$\hat{y}_i^{(m)} = \hat{y}_i^{(m-1)} + \eta \cdot f_m(x_i) \quad \text{.......(47)}$$

Where;

$\hat{y}_i^{(m))}$ - Represented prediction from the previous iteration at iteration $(m-1)$

$f_m(x_i)$ -Represented the output from m-th decision tree from the sample (x_i)

η-Represented the learning rate (controlled the contribution of each tree)

3.6.5 Splitting Criteria for Decision Trees

Each decision tree in LightGBM splits the nodes to minimize the objective function. Shared Gain was calculated as follows:

$$Gain = 1/2(\frac{(\sum(g)2}{\sum h+\lambda} - \frac{(\sum(gL)2}{\sum hL+\lambda} - \frac{(\sum(gR)2}{\sum hR+\lambda}) - \gamma \quad \text{.......(48)}$$

where:

$\sum g$, $\sum gL$, $\sum gR$ - Gradient sums for the parent, left, and right nodes,

$\sum h$, $\sum hL$, $\sum hR$ - Hessian sums for the parent, left, and right nodes,

λ - Regularization term for leaf weights,

γ - Minimum profit threshold for a split.

3.6.6 Gradient and Hessian Calculations

The gradient and Hessian values were derived from the loss function as described in the gradient and Hessian equations(Singh, 2024):

Gradient (gi):

$$gi = \frac{\partial(yi, y\hat{}i)}{\partial y\hat{}i} \quad \text{.......(49)}$$

Hessian (hi):

$hi \frac{\partial 2l(yi,y\hat{}i)}{\partial y\hat{}i2}$ ....(50)

For MSE

$gi = 2(y^i - yi)$, $h_i=2$

3.6.7 Feature Importance in LightGBM

The importance of features was evaluated using the split gains across all trees in the model:

$Feature\ Importance(xk) = \sum_{splits\ on(xk} Gain.$ .... (51)

This helped identify and rank the most influential predictors in the model.

3.6.8 Adapting LightGBM for Time Series

For time series forecasting, input features $x_t$included:

i. Lagged Features as;

$\{yt - 1, yt - 2, \dots, yt - p\}. \{y_{t-1}, y_{t-2}, \dots, y_{t-p}\}$

Where:

y_{t-p} represents the demand value p time steps before time t

P is the number of lagged terms included.

ii. Moving Statistics: Moving Averages or Standard Deviations:

$MA_t = \frac{1}{k} \sum_{L=t-k}^{t-k} y_i$.. (52a)

Rolling Variance

$(V_{ar_t} = \frac{1}{k} \sum_{i=t-k}^{t-1} (y_i - MA_t)^2$ ....(52b)

iii. Exogenous properties:

External data variables such as weather conditions or time indicators:

xt = {temperature, time of day, dayoftheweek}.

3.7 Estimation of Uncertainty

3.7.2 Monte Carlo Simulation

To estimate the uncertainty, perturbation of the input functions (x_t) was done multiple times by adding random noise:

$xt(j) = xt + \epsilon j, \epsilon j \sim N(0, \sigma 2)$ ....(53)

Generate multiple predictions:

$y\hat{}t + 1(j) = f(xt(j); \Theta).$ .... (54)

Computed the mean prediction and confidence intervals as:

Mean forecast:

$y^t + 1 = 1/N \sum_{j=1}^{N} y\hat{}(j)t + 1$ ....(55)

Forecast intervals: $[P_{2.5}, P_{97.5}]$, Pq = Quantile q $(\{y\hat{}t{+}1\ (j)\}_{j=1}{}^{N})$.

3.7.3 Quantile Regression

LightGBM was used to predict specific quantiles by adjusting the loss function:

$$\ell(yi, yi^{\alpha}) = \begin{cases} \alpha \cdot (\mathrm{yi} - y\hat{}i\alpha), & \text{if yi} > y\hat{}i\alpha, \\ (1-\alpha) \cdot \left(y^{i\alpha} - yi\right), & otherwise. \end{cases} \quad (56)$$

3.8 CATBOOST Model Design

CATBoost (Categorical Boosting), a gradient boosting algorithm,was designed to handle sequential data and categorical features efficiently. Its model design framework extended traditional gradient boosting by introducing ordered boosting and native handling of categorical features, which were particularly relevant for time series forecasting tasks (Abed, 2024).

3.8.1 Model Representation

The CATBoost model forecasted the power demand.$\hat{y}_{t+1}$for a future time step *t+1* based on input features x_t derived from historical data. The forecastwas expressed as a weighted sum of decision trees(Yang X. &., 2021):

$$\hat{y}_{t+1} = F(x_t) = \sum_{n=}^{M} \eta_{f_m}(x_t) \quad \text{.....(57)}$$

Where:

$\hat{y}_{t+1}$ - Predicted power demand at time *t*,

$f(x_t)$ - The m-th decision tree,

M - Total number of trees in the CATBoost model,

η - Learning rate controlling the contribution of each tree.

3.8.2 Objective Function

The objective function consists of two parts: the loss function ℓ and a regularization term Ω to control model complexity:

$$L = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{m=1}^{M} \Omega(f_m) \quad \text{..... (58)}$$

where:

y_i - True power demand for the *i-th* training sample,

$\hat{y}_i$- Predicted value,

ℓ -Loss function (e.g., Mean Squared Error, ℓ (*yi*, *y*^*i*) = (*yi*−*y*^*i*)^2),

Ω(fm) - Regularization term to penalize model complexity.

3.8.3 Ordered Boosting

In traditional gradient boosting, residuals are computed using all available data, which risks data leakage in time series forecasting (Yang X. &., 2021). CATBoost in this study addressed this issue by employing ordered boosting, where the data was permuted into multiple random orders.

For each data point i, only earlier observations in the permutation were used to compute residuals:

$$gi = \frac{\partial \ell(yi, y^\wedge i)}{y^\wedge i} \quad \text{.........(59a)}$$

$$hi = \frac{\partial 2\ell(yi, y^\wedge i)}{\partial y^\wedge 2i} \quad \text{.........(59b)}$$

Gradients and Hessians for each point were estimated from a subset of data:

$$g^\wedge iordered = \frac{\sum j \in P, igj}{|Pi|} \quad \text{.........(60a)}$$

$$h^\wedge iordered = \frac{\sum_{j \in P_i} hj}{|Pi|} \quad \text{.........(60b)}$$

where $P_i$represented the subset of prior observations.This ensuredthat no future data were used in training, maintaining temporal causality.

3.8.4 Decision Tree Splitting Criterion

Each tree $f(xt)$was constructed by recursively splitting the data to maximize the information gain at each node. The gain from splitting a node into left (L) and right (R) branches is shown as:

$$Gain = 1/2\left[\frac{(\sum g)2}{\sum h+\lambda} - \frac{(\sum gL)2}{\sum hL+\lambda} - \frac{(\sum gR)2}{\sum hR+\lambda}\right] - \gamma, \quad \text{.........(61)}$$

where:

g,h - Gradients and Hessians of the loss function for the entire node,

g_L,h_L - Gradients and Hessians for the left split,

g_R,h_R- Gradients and Hessians for the right split,

λ - Regularization parameter for smoothing,

γ - Minimum gain threshold to determine whether to split a node.

The splitting criterion ensured that splits significantly reduced errors and improved= demand forecasts.

3.8.5 Handling of Categorical Features

CATBoost natively supports categorical data without requiring manual preprocessing (Du, 2024). The model used target-based encoding during training as follows:

$EncodedCategoryi = \frac{\sum j \in Diyj}{|Di| + \lambda}$ ....(62)

where:

D_i - Subset of prior observations for the category,

y_j - Target values corresponding to Di

λ- Regularization parameter to prevent overfitting.

This encoding was integrated into the boosting process, ensuring efficient and reliable feature handling.

3.8.6 Prediction Update Process

After each iteration, the model updates predictions by adding the contribution of the newly constructed tree:

$\hat{y}^{(m)} = \hat{y}^{(m-1)} + \eta f_m(x_i)$(63a)

where:

$\hat{y}^{(m)}$ - Updated prediction after m-th tree,

η- Learning rate,

$f f_m(x_i)$)- Prediction from the m-th tree.

3.8.7 Time Series Design Methodology

For day-ahead power demand forecasting, the input feature vector x_t at time t included:

Lagged Power Demand -Historical demand values,

$\{yt - 1, yt - 2, \dots, yt - p\}. \{y_{t-1}, y_{t-2}, \dots, y_{t-p}\}$

i. Temporal Features - Hour of the day,

ii. Day of the week indicators,

iii. Weather Data -External variables like temperature.

The model predicted as:

$$\hat{y}_{t+1} = F(x_t) \quad (63b)$$

Where;

x_t Included all relevant historical and external features.

$\hat{y}_{t+1}$ -is the forecasted power demand for the next step

3.8.8 Uncertainty Estimation methodology in CATBoost

Uncertainty in CATBoost predictions was estimated using Monte Carlo simulation and Quantile Regression Techniques. The applied techniques are described as follows:

Monte Carlo Simulation

To assess model sensitivity to input noise, Monte Carlo perturbations were applied to the input features. For each simulation run j, a random noise term ϵ_j was sampled and added to the original input. Introduction of random perturbations ϵ to input features was done,and multiple predictions using the formula:

$$x_t^{(j)} = x_\tau + \epsilon_j{}' \in j \sim N(o_1 \sigma^2) \quad (64)$$

The forecast for each perturbed input

$$\hat{y}_{t+1}^{(j)} = F\left(x_t^{(j)}\right) \quad (65)$$

The predicted mean and variance used are:

Mean as

$$\hat{y}_{t+1} = \frac{1}{N} \sum_{j=1}^{N} \hat{y}_{t+1}^{(j)} \quad (66)$$

Uncertainty Estimation as;

$$\sigma^2 = \frac{1}{N}\sum_{j=1}^{N}\left(\hat{y}_{t+1}^{(j)} - \hat{y}_{t+1}\right)^2 \quad (67)$$

Quantile Regression

To directly predict prediction intervals, CATBoost was able to optimize a quantile loss function as shown in the equation:

$$\ell\alpha(y, y^{\hat{}}) = \begin{cases} \alpha(y - y^{\wedge}) & if\ y > y^{\wedge}, \\ (1-\alpha)(y^{-y}), & if\ y \leq y^{\wedge}. \end{cases} \quad (68)$$

Separate models for lower and upper quantiles (e.g., α=0.05,0.95) estimated confidence intervals for the forecast.

3.9 K-Nearest Neighbors (K-NN)Base Model Design Process

The task was to design a K-NN-based modelused to forecast power demand for the next 24 hours (H = 24) based on time series data from historical observations. The base model integrated K-Nearest Neighbors (KNN) regression with uncertainty estimation techniques to handle input perturbations (e.g., sensor noise, measurement errors). The model was designed to provide robust predictions and quantify uncertainty, enabling better decision-making.

3.9.1 Design of a K-NN Base Model

Representation of input data for a time series dataset was represented as:

$$D = \{(t_1, x_1, y_1), (t_2, X_2, y_2), \cdots 1(t_N, x_N, y_N)\} \quad (69)$$

where:

t_1 is the timestamp for the ith observation.

$xi=[x_i1,x_i2,\ldots,xid] \in Rdx$: Principal vector for ith observation, including variables such as temperature, voltage, current, and previous power requirements.

$y_i \in R_y$- Observed Power Demand over time t_i

N - Total number of observations.

a. Feature Engineering

Generated lagged functions to capture time series dependencies as:

$$x_i(Lagged) = [y_{i-1}, y_{i-21} \dots, y_{i-l}, x_{i_1}, x_{i_2}, \dots, x_i d] \quad (70)$$

where;

l is the lag window size. For the forecast for the day$l$=H=24

b. Incorporation of Perturbation

Input perturbations are modeled as follows:

$$x_i^{(\rho_e rturbe_d)} = x_i + \Delta_i \quad (71)$$

Where:

$\Delta_i \sim N(0,\Sigma)$ is a multivariate Gaussian noise with covariance matrix Σ, estimated from the historical variability of the data.

3.9.2 K-NN Regression with Time Distance Metric

To account for temporal and feature space similarities, a hybrid distance metric was defined as:

$$d(q, x_i) = \sqrt{\sum_{j-1}^{d} \omega_j \left(\frac{Lq_j - x_{iJ}}{\sigma j^z}\right)^2 + \lambda |t_q - t_i|} \quad (72)$$

Where:

q-Query point for forecast for the day.

w_j-Important weight for the j-th feature.

σ_j^2-Variance of the j-th feature perturbation.

λ-Temporal weighting factor to emphasize recent observations.

$/tq-ti/$- Temporal difference between query and training points.

K-Nearest Neighbor selection

Selected k-nearest neighbors $N(q)$ based on the temporal distance metric:

$$N_k(q) = \{(x_{i_1}, y_{i_1}), (x \cdot \lfloor 2 \rceil^y i_2)_{J^{7117}} (x_{ik}, y_{ik})\} \quad (73)$$

Where $d(q, x_i) \leq d(d, x_{i2}) \leq \cdots \leq H(qx_{ik})$

3.9.3 Hybrid Prediction Mechanism

Aggregation of forecasts was done by predicting power demand $y^{\wedge}{}_q$ at time t_q using the weighted average of neighboring outputs:

$$y^q = \frac{\sum_{i \in Nk(q)} \omega i \cdot yi}{\sum_{i \in Nk(q)} \omega i} \quad \text{...(74)}$$

where:

y^q represents the predicted value or overall forecast at a specific point called q. It is a smoothed estimate made by averaging nearby outputs, yi, but each output is given a different importance based on its weight, wi. In the context of predicting power demand, y^q refers to the forecasted power demand at timetq.

Nk(q) is the group of nearby points around the query point q. wi is the importance given to each neighbor i.

yi is the actual or predicted output from neighbor i. Each yi is used by the model to help estimate y^q.

Uncertainty estimation

Uncertainty estimation was used to quantify the prediction uncertainty as the variance between neighboring outputs:

$$\frac{\sigma_q^2 = \Sigma_i \in N_k (q v \Gamma v_i \mathbb{C} - y_i - \hat{y}^q)^2}{\Sigma_i \sqrt{\in} N_k (q) W} \quad \text{...}$$

(75)

Where;

$\sigma q2$ -Estimated prediction uncertainty (variance) at the query point q

Nk,, -Set of close neighbors to query Point q

ω_i - The weight of the neighbor i

y_i-Neighbor's output value iI (celebrated or predicted).

y_q, Predicted value at query point q

$(y_i - y^{\wedge q})2$ -The squared deviation of neighbor ii's output from the predicted mean

3.10 Hybrid Model Methodology

The Hybrid Model Methodology used in this study involved sequentially designing, training, and integrating multiple machine learning models XGBoost, CatBoost, and Random Forest into a single hybrid forecasting model, enhanced with an uncertainty-estimation mechanism to improve prediction accuracy, robustness, and reliability under input perturbations.

The hybridized model architecture wasdesignedby combining models such as XGBoost, CATBoost, and RandomForest into ahybrid Methodology. This embodied a unique solution for accurate Power Demand Forecasting under input perturbations. Through the harmonious convergence of these methodologies, the architecture was evaluated to determine the accuracy of the forecasts.Creating a hybrid of the perturbed model was achieved by repeating the steps in Model selection and Model training multiple times with different input perturbations until the desired optimum accuracy was reached.

The study employed a hybridization technique in Machine Learning by combining the strengths of multiple models to improve overall performance and generalization. Since time series data present specific challenges that require specialized modeling techniques, a hybrid approach was used in power demand forecasts, combining multiple models to improve forecast accuracy and robustness. Methods used include bagging and boosting time series data, combining exponential smoothing state-space models, and training multiple regression models on different subsets of the time series data. The choice of hybrid method was dependent on the characteristics of the time series data and the variability challenges as presented by the data (Tama, 2021).

3.11 Perturbation Methodology

The study employed an Input Perturbation-Based Uncertainty Estimation Method, where controlled Gaussian noise was injected into selected input features to assess forecast sensitivity and quantify prediction uncertainty.This study used historical data to forecast future power demand, with the input data intentionally perturbed to simulate uncertainty. The perturbations, either random or systematic variations applied to each element of the input data, mimicked the unpredictability of real-world scenarios and their impact on model predictions. The process involved generating multiple perturbed data sets, training the model, forecasting demand for each one, and estimating uncertainty using theresult analysis of the forecasted values. Common measures such as standard deviation or prediction intervals were used to quantify uncertainty. Subsequently, forecasts from different perturbed datasets were aggregated into a single forecast using methods such as Bayesian model averaging or voting were considered(Lee S. Y., 2024).

Evaluation using performance metrics was conducted to measure model accuracy, compare model forecasts to actual consumption, and assess the accuracy of uncertainty estimates. Training and retraining of the model and failure parameters were conducted, guided by the evaluation results, to increase forecast accuracy and confidence in uncertainty. Perturbation encoding was used to protect sensitive data while enabling useful analysis. Instead of directly using raw data, small perturbations were added to the data to protect privacy while maintaining the usefulness of the data for analysis. Differential privacy is one such technique that uses perturbation encoding to ensure that statistical data analysis does not reveal too much information about any single data point. A web-based system was designed to simulate the forecasting and perturbation scenarios.

3.12 Model Implementation

The specific objective three in the study was the implementation of HPDEF-MUIP. Theimplementation involved the use of a hybrid model designed to estimate the uncertainty in future 24-hour ahead forecasts under input perturbations in smart meter electricity data. The Python programming language and deep learning frameworks, such as TensorFlow, Keras, and PyTorch,were used to implement the hybrid model. Adversarial sampling techniques, such as the Fast Gradient Sign Method (FGSM) or Projected Gradient Descent (PGD), were used to generate adversarial examples for training and evaluating deep learning models. Cross-validation was employed to assess the performance of the models and tune the hyperparameters. The study was conducted on a high-performance computing platformto support the computational requirements of the deep learning models and the adversarial sampling techniques.

3.13 Model Evaluation

To achieve specific objective four, Model evaluation was done. The HPDEF-MUIP model was evaluated using 20% validation data to ensure it meets the performance criteria. Fine-tuning included adjusting the hyperparameters and retraining the underlying models. Evaluating the model's performance was done on a holdout dataset, using metrics such as accuracy, precision, recall, and F1-score. The performance of the developed hybridized model was evaluated using computational research methods, including statistical analysis, Machine Learning, and Data Visualization. This evaluation wasconducted using real-world datasets to validate the model's accuracy. The model was evaluated on metrics such as the Mean Absolute Scaled Error (MASE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE), Root Mean Square Error RMSE, and R-squared value (R^2). The model forecasts werefurther compared with the actual power demand to determine the accuracy of the HPDEF-MUIP model.

3.14 Model Testing and Interpretation

The performance of the designed model was compared with the actual demand and results from the literature for previous methods for Uncertainty Estimation in deep learning models using the time series smart meter electricity dataset. The performance of the developed hybrid model for power demand forecastingwas compared with the actual demand and with previous methods from the literature for Uncertainty Estimation in deep learning models, using the time series smart meter electricity dataset. This comparison was conducted using computational research methods,which highlighted the strengths and weaknesses of each method and identified areas for improvement when dealing with the smart meter electricity dataset.

3.15 Ethical Considerations

This study utilized time series smart meter electricity data and meteorological weather data. Data were collected from the Kenya Power and Lighting Company's central smart meter at Lanet substation, designated for the Central Rift Region. Weather datawere collected from the Nakuru Meteorological Station.The study ensured that no private information about power consumption patterns was exposed, as the data were figures collected from a centralized metering unit.

The study ensured that data from the eleven supply points were anonymized and handled securely to prevent unauthorized access to or misuse of the data. Only the required voltage data and demand data were collected using an Excel spreadsheet, which was thenencrypted and stored securely in a database. No unauthorized access to the collected data was allowed; only the researcher had access to the data for the sole purpose of the academic study, where the data was used for model training. Data safety was ensured by protecting the data during transit and at rest through the application of data encryption methods, such as the Advanced Encryption Standard (AES), and the use of passwords. This ensured data minimization, hence further reducing the vulnerability of data exposure during the collection process.

The dataset collected was stored and used during the period of this study. After the research is completed, the data will be disposed of through paper shredding, and all physical copies will be destroyed. Additionally, all soft copies will be deleted from all storage devices used. All relevant notes and records that could lead to the identification of collected data were destroyed after the study's completion.

Measures were taken to minimize the potential risk that the data used to train the model might be biased. Such a bias could lead to demand forecasts that perpetuate existing inequalities. This would lead to an inequitable distribution of power and misallocation of resources, adversely affecting marginalized communities in particular. To mitigate the bias, thorough data audits were performed to identify and correct any bias in the dataset. This involved ensuring that the data was representative of different demographics and energy consumption patterns. Techniques such as stratified sampling were used to ensure that all relevant groups were adequately represented in the training data.

The deployment of Artificial Intelligence, deep learning, and advanced optimization techniques in power demand forecasting raises ethical questions regarding the use of these technologies. This posed a risk that such technologies could prioritize profit over the public good, resulting in decisions that are out of step with broader societal needs. In this regard the study adhered to guidelines for the ethical use of technology, emphasizing the prioritization of public good and sustainability over profit.Only the required data was collected and stored securely. No unauthorized access to the collected data was allowed. Data safety was ensured through the use of encryption and passwords. After use, the data will be disposed of by deleting it from all devices used, in accordance with legal provisions. All relevant notes and records that could lead to the identification of collected data will be destroyed after the study is completed.

The study obtained ethical Clearance from Kabarak University Research Ethics Committee (KUREC). A recommendation letter tothe National Commission for Science, Technology, and Innovation (NACOSTI) was obtained from IPGS. A formal request was made seeking a research permit. A research permit was issued by NACOSTI after consideration of the submitted proposal and the recommendation letter from IPGS. Research data collection authorization was sought and issued by the Nakuru County Director of Education, the Nakuru County Governor, and the Nakuru County Commissioner. Informed consent wasobtained from the data owners or the relevant authorities before using the data for research purposes. The study adhered to ethical guidelines for research involving human subjects, as outlined by institutional review boards or other relevant bodies.

CHAPTER FOUR

DATA ANALYSIS, PRESENTATION AND DISCUSSION

4.1 Introduction

This chapter focuses on data analysis, presentation, and discussion of findings related to the Hybrid Power Demand Forecasting with Uncertainty Estimation Under Input Perturbations. As power systems become increasingly complex with the integration of renewable energy sources, accurate power demand forecasting is crucial for enhancing power system reliability and efficiency. The Hybrid Forecasting Model combined traditional statistical methods with deep learning methods to enhance the model accuracy. Input perturbations were applied to account for uncertainties in the data, providing a more robust estimation of future power demand. The following sections present the study results on thedesign of the HPDEF-MUIP model, its implementation process, evaluation of HPDEF-MUIP, and discuss the implications of the findings in the context of power generation, supply planning, and maintainingpower grid stability.

4.2 Design of a Hybrid Power Demand Forecasting Model with Uncertainty *Estimation under Input Perturbations*

This section presents the results of the second research question of this study. The research question was how a hybrid power demand forecasting model with uncertainty estimation under input perturbationscould be designed.This study focused on the innovative design of HPDF-MUIP, taking cognizance of the design recommendations from the weaknesses identified in the literature review. The recommendations i-viiwere taken into account in the design of this study's hybrid model. The model design followed the steps as;

4.2.1 Data Pre-Processing

i. Loading dataset

The process began by reading a dataset from an Excel file called pilotdata.xlsx using the pd. read_excel function. This method was part of the panda's library, which was particularly suitable for working with structured data that contained labels. Using pandas made it possible to easily analyze the dataset and prepare it for further processing. The code line is;

Line 58 # 1. Data Loading and Initial Processing

Line 59 def load_and_prepare_data(file_path):

ii. Handling of missing values

Data.fillna(data.mean(), inplace=True) was used to handle missing values in the dataset. This function fills in the missing items in each column with the median value of that column. This approach proved advantageous because it preserved the overall size of the dataset while minimizing the risk of bias that could be introduced by more arbitrary imputation methods, such as zero-padding or dropping rows. The code line is implemented.

86 # Fill missing values with rolling mean

87 for col in numerical_cols:

88 df[col] = df[col]. fillna(df[col]. rolling (window=24, min_periods=1). mean ())

iii. Removal of infinite values

The next step involved filtering out any rows that contained infinite values, which may have occurred during preprocessing or feature engineering processes, such as when division by zero resulted in an error. This was achieved using the data [np. isfinite(data). all (1)] command, which ensured that only finite values were retained in the dataset.

Clean data was essential for training machine learning models as it helped avoid errors and ensured reliable results. Effective handling of missing and infinite values mitigated the risk of model failure and prevented the introduction of bias into predictions. This careful data preparation served as a fundamental step in building robust machine learning systems

4.2.2 Feature Selection

The code includes several features that help capture different aspects of power demand (i)Temporal Features: Hour, day of week, month, and weekend indicators help capture daily and seasonal patterns that might be affected by perturbations. (ii) Lag Features, such as Power demand from previous hours, days, and weeks, helped the model understand trends and persistent patterns through disturbances. (iii) Weather Data: Temperature features allowed the model to adjust for weather-related changes in power demand

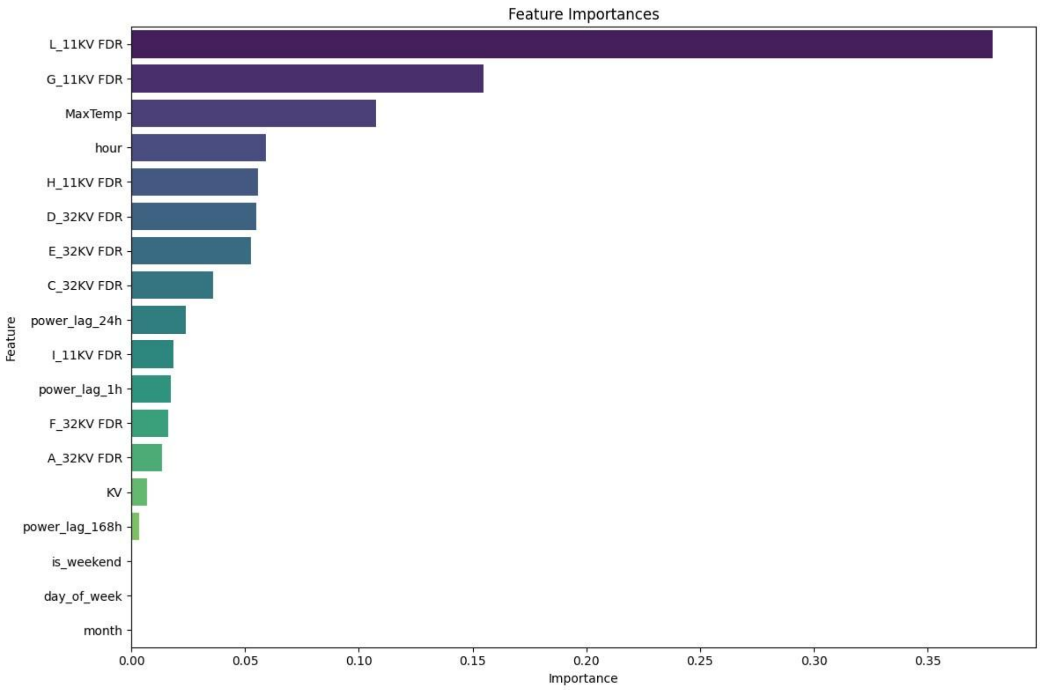
All columns in the data set, except the Power Demand column, were treated as independent variables. This decision meant that each of these columns represented a different characteristic that could potentially affect the target variable. By isolating these features, the analysis focused on identifying patterns and relationships that could be useful for forecasting modeling.

The Power Demand column was identified as the dependent variable that was the primary focus of the Forecasting task. This column represented the outcome that the models were supposed to predict based on the independent variables. By clearly defining the target variable, the modeling process focused on understanding how changes in functionality affected Power Demand.

A clear separation of features and target variables was performedto ensure that the data was properly prepared for training machine learning models. Models typically require distinct inputs (features) and outputs (goals) in order to learn effectively. This delineation facilitated a structured approach to the modeling process, enabling more accurate forecasts and a deeper understanding of the relationships within the data. By creating this framework, the foundations were laid for successful model training, which ultimately increased the reliability of the demand forecasts made. The feature importance results were as follows.

Figure 1

Feature Importance



4.2.3 Interpretation of Feature Importance Results

The feature importance plot highlights the relative contribution of input variables to the model's forecasting power.

Dominant Features

L_11KV FDR emerges as the most influential predictor, contributing nearly 40% of the total importance. G_11KV FDR follows as the second most important variable, accounting for approximately 15%, while MaxTemp is the third most significant at around 12%. These findings suggest that feeder-level consumption data (particularly from the 11KV network) and temperature conditions strongly influence short-term power demand.

Moderately Important Features

Hour of the day, H_11KV FDR, D_32KV FDR, E_32KV FDR, and C_32KV FDR all show moderate levels of importance, with each contributing between 4–7%. This indicates that time-related variables (hourly cycles) and additional feeder inputs from both 11 kV and 32 kV networks provide secondary but meaningful contributions to demand variability.

Lag Features

power_lag_24h (previous day demand) and power_lag_1h (previous hour demand) provide modest contributions, showing that temporal autocorrelation is present but less dominant compared to feeder and temperature variables. Interestingly, power_lag_168h (weekly lag) has very low importance, suggesting that weekly seasonality is less critical in this dataset compared to daily and hourly trends.

Low-Importance Features

Features such as KV level, is weekend, day_of_week, and month contribute minimally. This implies that while calendar-based and categorical features capture some variability, they are not strong predictors relative to feeder data and environmental conditions.

Implications for Model Design

The dominance of feeder-level data (especially L_11KV FDR and G_11KV FDR) emphasizes the need to prioritize infrastructure-level monitoring in forecasting frameworks.The importance of temperature (MaxTemp) validates the role of exogenous environmental variables in explaining load fluctuations.The relatively low importance of calendar effects suggests that the model is less dependent on seasonal/categorical structures, likely because feeder-level demand patterns already encode much of this temporal variation.

The results indicate that network-specific feeder data and environmental variables (temperature) are the most reliable drivers of power demand forecasts. Lagged demand values play a supportive role, while categorical and calendar-based features provide minimal additional benefit. This distribution of importance suggests that hybrid forecasting models should emphasize feeder monitoring, weather integration, and short-term lag features, while giving less weight to purely calendar-based predictors

4.2.4 Splitting the Data

The dataset was initially divided into two primary subsets to facilitate the modeling process. Sixty percent (60%) of the data was allocated for training purposes, creating the training datasets, denoted as X_train and y_train. This substantial portion of the dataset was essential for training the model as it provided a robust foundation for the algorithms to learn from the various patterns and relationships present in the data. The remaining fortypercent (40%) of the data set was set aside as temporary data labeled as X_temp and y_temp, which was further divided into validation and test data sets. This strategic division was critical to ensure that the model could be effectively trained while allowing subsequent evaluation of its performance.

The temporal data were then subjected to further partitioning to create two additional subsets: a validation dataset and a test dataset. Twenty percent (20%) of the temporary data was allocated for validation purposes, resulting in the creation of X_val and y_val. This validation dataset played a crucial role in tuning the hyperparameters and evaluating the model during the training phase, enabling adjustments to be made to optimize its performance. The remaining twenty percent (20%) of the temporal data was designated as the test dataset, labeled X_test and y_test. This dataset was reserved for final evaluation after the training process was completed, providing an unbiased assessment of the model's predictive capabilities.

The approach of partitioning the data into training, validation, and test sets was based on best practices for developing machine learning models. A training set comprising sixty percent of the data ensured that the model had ample opportunity to effectively learn from a large portion of the dataset and capture the underlying patterns necessary for accurate predictions. The validation set, which is twenty percent of the data, helped fine-tune the hyperparameters and prevent overfitting that could occur if the model were overfit to the training data. The test set, which also represents twenty percent of the data, provided a crucial, unbiased evaluation of the final model's performance, allowing for a clear understanding of how well the model would generalize to new, unseen data. To increase the reliability and reproducibility of these distributions, the random_state parameter was used to ensure that the random selection of data points remains consistent across different analysis runs. This careful approach to data partitioning laid a solid foundation for subsequent modeling and ultimately contributed to the development of a hybrid machine learning model.

4.3 Model Design Results

During the model development phase, several algorithms were used to train the predictive models, namely XGBRegressor, CatBoostRegressor, and Artificial Neural Network (ANN). XGBRegressor was chosen for its efficiency and effectiveness when working with large datasets, particularly its gradient boosting framework, which optimizes performance. Known for its ability to process categorical features without extensive preprocessing, CatBoostRegressor has also been leveraged to exploit its strengths when working with different types of data. Additionally, an ANN was implemented to explore the potential of deep learning in capturing complex relationships in data. Each model was trained and retrained on previously defined training datasets, X_train and y_train, allowing them to learn from the underlying patterns present in the data.

The study results represent seven different base models that were designed, trained, and evaluated as follows: Artificial Neural Networks (ANN), Recurrent Neural Networks (RNN), Support Vector Regression (SVR), XGBoost, Random Forest, LightGBM, and CatBoost. Their performances were analyzed using standard regression metrics. Coefficient of Determination (R^2), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The goal was to select the best individual base model and use it to develop a Hybrid Model with excellent forecasting accuracy by capitalizing on the strengths of individual models, thereby enhancing demand forecasting performance.

Table 2

Table of Model Design Results

Rank	Model	R^2	RMSE	MAE	MAPE	Ranking Observations
i.	XGBoost	0.9534	1.7217	1.2233	3.1143%	Best individual model overall
ii.	CatBoost	0.9445	1.8792	1.3857	3.4942%	Excellent accuracy
iii.	Random Forest	0.9330	2.0640	1.4226	3.5465%	Solid performance
iv.	LightGBM	0.9279	2.1408	1.4834	3.6805%	Very good
v.	SVR	0.9001	2.5200	1.8797	4.7460%	Good performance
vi.	RNN	0.5829	5.1497	3.5702	9.2606%	Slightly worse than ANN
vii.	ANN	0.7433	4.0404	3.2057	8.2576%	Weak performance

4.4 Model Design Performance Results Discussion

After model training and evaluation, the top three models with the best forecasting accuracy were selected using the following code snippet.The code listing line is available at the appendix.

456 top_models = sorted(trained_models.items(), # Select top 3 models based on RMSE.

While line 460 prints ("\nTop 3 models selected for hybrid model:").

4.4.1 XGBoost Base Model Design Results Analysis

The XGBoost model demonstrated the highest predictive performance in this study, achieving an R^2 value of 0.9534, an RMSE of 1.7217, and a MAPE of 3.11%. These metrics reflect exceptional regression accuracy and minimal prediction error, positioning XGBoost as a leading candidate for short-term power demand forecasting. Recent studies corroborate these results. For example, Liang (2023)employed a hybrid

XGBoost–LSTM model for load prediction, achieving impressively low MAPE values between 0.98% and 1.20%, underscoring the model's robustness when integrated with temporal deep learning techniques. Similarly, Chen et al. (2024) introduced an XGBoost–BiGRU framework for ultra-short-term load forecasting, emphasizing its capacity to adapt to rapidly shifting demand patterns (Zhang L. &., 2024). These contributions highlight the scalability, generalization power, and computational efficiency of XGBoost, especially in handling high-frequency, structured energy data (Wang Y. &., 2023). The consistent success of XGBoost across varied configurations and hybridizations affirms its strategic value in the development of high-performance predictive systems within the energy sector.

4.4.2 CatBoost Base Model Design Results Analysis

CatBoost also demonstrated strong future capabilities, achieving an R2 of 0.9445 and anMAPE of 3.49%. Its standout feature is the native handling of the category variable without the need for broad pre-processing, especially an advantage for smart grid forecast works, where mixed data types are common. Recent studies strengthen the growing utility of CatBoosts in energy prediction models. Zhang introduced a Hybrid CatBoost -XGBoost framework, carried forward by an arithmetic adaptation algorithm, with accurate load forecasts while maintaining computational efficiency (Zhang J. &., 2024).

Similarly, Wu et al. (2023) highlighted the better performance of the CatBoost over traditional ensemble methods in a dynamic load environment, attributing this to its gradient-based decision trees holding its shield-based oblivious decision trees and ordered boosting techniques (Wu Z. H., 2023). These studies confirmCatBoost's robustness in creating coherent and reliable outputs in various datasets, establishing it as an important contributor in the development of hybrid forecast systems and highlighting its emerging importance in intelligent energy management (Liu X. &., 2023). Its balance between interpretation, speed, and accuracy makes it a rapid preferred option in both academic and operational contexts.

4.4.3 Random Forest Base Model Design Results Analysis

The Random Forest model has demonstrated a solid performance in predicting the demand for electricity, as indicated by the following results. The R-Squared value of 0.9330 suggests that the Random Forest model may explain 93.30% of the variance in the target variable. It is an exceptionally high value, suggesting that the model has an excellent fit for the data and can capture the underlying pattern very well. Root Mean Squared Error (RMSE): 2.0640 MW. The RMSE represents a relatively low error in model predictions, indicating that the forecasts are closely aligned with actual power demand values.

Meaning Absolute Error (MAE): 1.4226 MW. The MAE further confirms the high accuracy of the model, with an average full error of only 1.4226 MW. Mean Absolute Percentage Error (MAPE): A MAPE of 3.5465% represents a low percentage error, indicating that model predictions are highly accurate, with an average percentage deviation of 3.55% from the actual values. The Random Forest Model ranked 3rd in individual models, demonstrating strong performance compared to other techniques. The impressive performance of the Random Forest model can be attributed to its ability to handle the complexity inherent in power demand data and effectively handle non-linearity. Random Forest combines multiple decision trees, which can capture intricate relationships and patterns in data.

The ranking of the Random Forest model in third position, behind the top-performing models XGBoost and CatBoost, suggests that it is a highly competitive and versatile algorithm that was considered a strong candidate for inclusion in a hybrid modeling approach. Its solid performance across various evaluation metrics makes it a valuable addition to the overall modeling strategy.

The performance of the Random Forest model in power demand forecasting has been well-documented in recent scholarly literature. Panapakidis and Dagoumas (2022) conducted a study comparing the effectiveness of various ensemble learning techniques, including Random Forest, for electric power demand forecasting (Panapakidis, 2022). The authors found that the Random Forest model achieved an R^2 of 0.92 and a MAPE of 3.8%, which is consistent with the results observed in the previous analysis (Rank 3, R^2 = 0.9330, RMSE = 2.0640, MAE = 1.4226, MAPE = 3.5465%).

The results reveal a clear stratification in the model's performance. XGBoost achieved the best overall accuracy with an R^2 of 0.9534 and the lowest RMSE of 1.7217. This suggests that XGBoost not only effectively captures data trends but also significantly minimizes prediction errors. CatBoost and Random Forest also performed admirably, offering R^2 values over 0.93, which indicates robust learning and generalization capabilities. Tree-based models, especially gradient boosting techniques such as XGBoost and CatBoost, tend to outperform neural networks and core methods in structured data tasks (Liu W. D., 2020). Neural models, such as ANN and RNN, although they are strong in sequential and unstructured contexts, have shown weaker performance.

To leverage the strengths of the most effective models, a hybrid model was created by combining XGBoost, CatBoost, and Random Forest using a weighted approach. To design anHPDF-MUIP, the code line 464-467 was implemented as;

Create a hybrid model

hybrid_model= HybridPowerDemandForecaster([model for name, model in top_models]   )

The model optimization procedure included weight adjustment to minimize RMSE and maximize R^2. The resulting optimized weight was:

Table 3

Top Three Accurately Performing Base Models Design Weights

S/No	Model	Weight
1.	XGBoost	67.7%
2.	CatBoost	18.2 %
3.	Random Forest	14.3%

From Table 4, it was observed that the fact that random forest was allocated 14.3% weight suggests that although it was strong in itself, its contributions to the hybrid model wereredundant or less additive in combination with the other two models. XGBoost, receiving the highest weight, clearly drove the hybrid model's predictive capability, with CatBoost showingsignificant improvements.

```
#Hybrid Model Class
class HybridPowerDemandForecaster:
  def __init__(self, models, weights=None):
    self.models = models
self.weights = weights if weights is not None else [1 / len(models)] * len(models)
  def fit(self, X, y):
    for model in self.models:
model.fit(X, y)
```

4.4.4 Design Results Discussion

The careful training and testing of seven main forecasting models yielded clear results that informed the design of the hybrid model. As shown in Table 3, the comparison between the models revealed clear differences in their performance, with gradient boosting models outperforming kernel-based and neural network methods in all regression metrics.

The group of models, called gradient boosting, which includes XGBoost, CatBoost, and LightGBM, had the best forecasting performance. These models had R^2 values over 0.92 and MAPE values under 4%. Among them, XGBoost was the top performer with an R^2 of 0.9534 and a MAPE of 3.11%, making it the best choice for the hybrid model. CatBoost came next with an R^2 of 0.9445 and a MAPE of 3.49%. It handles categorical data well and reduces bias by using ordered boosting. LightGBM performed slightly worse than the others, but it still demonstrated that boosting methods are reliable for understanding the complex patterns in the energy demand data.

RandomForest also did well, with an R^2 of 0.9330 and a MAPE of 3.55%.It's effective at reducing variability and works well with noisy data, although it doesn't match the accuracy of boosting models. However, its ability to handle noise makes it a valuable addition to the hybrid model, particularly in addressing inconsistencies in the input data. This aligns with previous research, which suggests that bagging methods are less accurate than boosting but more stable and reliable when forecasting in uncertain conditions.

Support Vector Regression (SVR) was only average, with an R^2 of 0.9001 and a MAPE of 4.74%. This suggests that it wasn't sufficient in handling the high-dimensional, nonlinear relationships in energy demand data. Neural networks, such as ANN and RNN, performed significantly worse, with R^2 values of 0.7433 and 0.5829, respectively. Despite their theoretical strength in capturing temporal dependencies, their poor generalization indicated overfitting, instability under noisy signals, and excessive computational complexity, as observed in prior studies (Hairani, 2023). These issues show that, in isolation, traditional neural models are not reliable for forecasting under dynamic and perturbed input conditions.

The empirical findings from Table 3, therefore, validate the hybrid design methodology by strategically selecting XGBoost, CatBoost, and Random Forest, which are the most effective and complementary models. The hybrid architecture is designed to strike a balance between high forecasting accuracy, effective control of variability, and adaptability in the face of input perturbations. The hybridization process is thus grounded in both experimental results and theoretical justification, ensuring that the model design addresses the weaknesses identified in both the literature review and experimental benchmarks.

This discussion provided a solid foundation for moving to the next step of the Hybrid Architecture Implementation, where the selected models were fused within the hybrid structure framework. This integration, augmented with uncertainty estimation mechanisms, ensured that not only high forecasting accuracy but also operational reliability in practical deployment scenarios, such as grid load management and short-term power demand planning

Figure 4

Hybrid Model Design Architecture

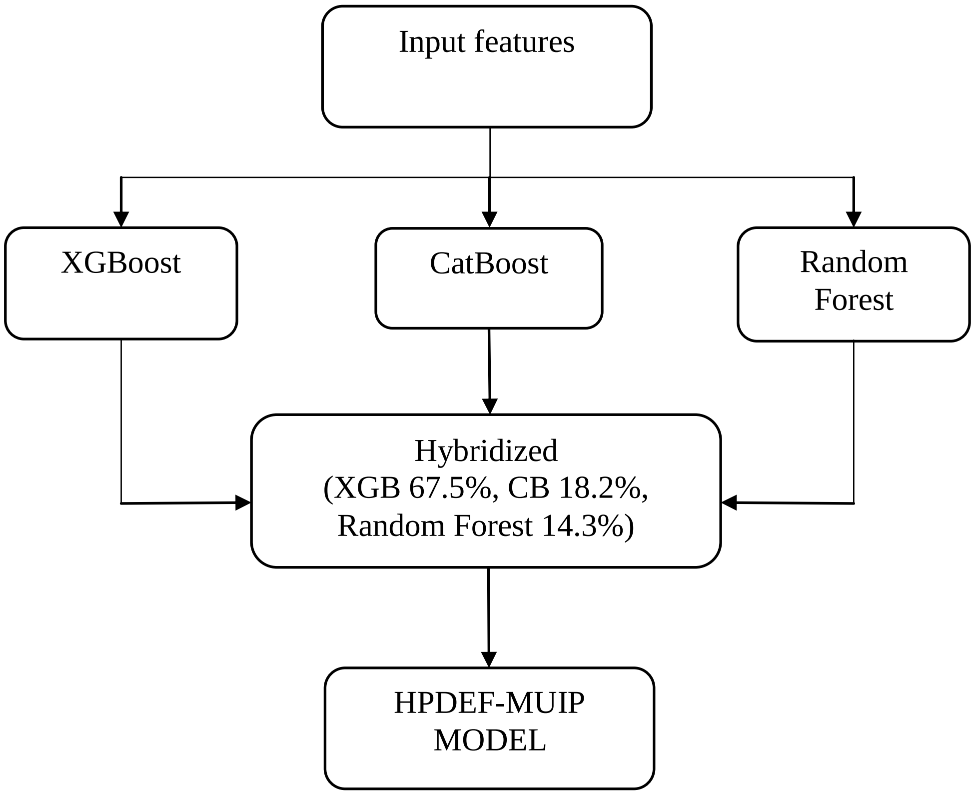


Figure 4 illustrates the schematic architecture of the hybrid model used in the study. The input features represent the input data fed into the model. The data was processed through the three algorithms as XGBoost, CatBoost, and RandomForest. The data was used to train the hybrid model. The Hybrid model leveraged the strengths of the individual base model weights of XGBoost (67.5%), CatBoost (14.3%), and Random Forest (18.2%). This method of hybrid modeling has been validated in recent literature. For example, Yidong et al. (Yidong, 2022) have shown that a weighted hybrid of boosting algorithms yields lower generalization errors than any single base model in energy forecasting models. The XGBOOST, as a dominant learner, aligns with numerous empirical findings across fields such as(Nti, 2023)

4.5 Model Implementation

The model was implemented by combining the best-performing models to create a hybrid model that improves prediction outcomes.

4.5.1Model Architecture

The following input parameters are required for making predictions:

i. Hour (0-23): The hour of the day for the prediction.
ii. Day of Week (0-6): The day of the week (0=Monday, 6=Sunday).
iii. Line Voltages (KV): Voltage measurements
iv. 11KV Transformer Feeder Currents (FDR A,C,D,E,F)
v. 32KV Transformer Feeder Currents(FDR GH,I,L)
vi. Power Demand in Mega watts (MW)
vii. Temperature (MaxTemp)

4.5.2 Hybrid Architecture Implementation

Building on the comparative evaluation of base models, the hybrid architecture was designed to strategically integrate the predictive accuracy of gradient boosting algorithms with the variance stability of bagging methods. The core guiding principle was to fuse the complementary strengths of top-performing models, such as XGBoost, CatBoost, and RandomForest, while systematically addressing weaknesses identified in both the literature review and experimental evaluation.

4.5.3 Hybrid Fusion Methodology

The hybrid model employed a stacked hybrid framework strategy organized into two layers:;

i.level-1.

The base models, XGBoost, CatBoost, and Random Forest, served as the primary forecasters. Each model generated independent predictions for the 24-hour demand horizon. These predictions reflect distinct inductive biases, where XGBoost optimized gradient-boosted regression trees with strong handling of complex nonlinearities, CatBoost excelled in managing categorical interactions and mitigating overfitting through ordered boosting, and Random Forest stabilized the ensemble by reducing variance across perturbed inputs.

ii. Level 2

Meta-Learner: Forecasts from the base models were combined through a weighted regression meta-learner, where weights were adaptively assigned based on the historical performance of each base model across validation folds. Weights were dynamically updated using an optimization criterion that minimizes the Mean Absolute Error (MAE) while constraining variance across input perturbation scenarios. This ensured that no single version dominates under unstable input conditions, thereby increasing model resilience.

4.5.4 Uncertainty Estimation under Input Perturbations

To explicitly model uncertainty, the hybrid model architecture integrated a dual mechanism of error propagation and perturbation simulation as follows;

i. Monte Carlo Dropout and Bootstrapping.

Within the RandomForest component, bootstrap sampling provided uncertainty bounds. For boosting models, Monte Carlo dropout is applied during inference to approximate posterior distributions of forecasts, generating probabilistic intervals around each forecasted value.

ii. Input Perturbation Simulation

Key exogenous variables (e.g., temperature, weekday/weekend effect) are perturbed within realistic ranges during training and inference. The ensuing distribution of predictions is quantified through variance metrics, yielding confidence intervals that capture the model's sensitivity to real-world volatility. This process operationalized the concept of uncertainty-aware forecasting, allowing operators to account for forecast reliability in their planning.

iii. Uncertainty-aware Loss function

A modified loss function is implemented at the hybrid model fusion stage, combining the minimization of forecasting errors (MAE, RMSE) with penalization for overly narrow confidence intervals. This prevents the hybrid model from generating misleading prices and in accurate forecasts, ensuring probabilistic reliability.

4.5.5 Adaptive Learning and Real-Time Feedback

The hybrid model incorporated a rolling update mechanism, wherein forecast residuals from real-time deployment are fed back into the model. Residual-based error correction allowed the system to dynamically recalibrate weights across base models, ensuring robust performance in evolving demand environments. Over time, this adaptive system strengthens the hybrid model's ability to cope with structural shifts (e.g., seasonal consumption spikes, business shutdowns, and renewable power penetration).

4.5.6 Interpretability and Decision Support

Recognizing the importance of interpretability in power system operations, the hybrid architecture integrated SHAP (Shapley Additive exPlanations) and feature importance analysis. These methods allow stakeholders to trace how input perturbations (e.g., unexpected temperature surges or calendar anomalies) contribute to forecast uncertainty. The interpretability module provided a visual dashboard displaying:

i. Forecast trajectory with shaded uncertainty bands.
ii. Feature contribution ranking under perturbed vs. unperturbed scenarios.
iii. Real-time confidence scores alongside traditional performance metrics (R^2, MAE, MAPE, RMSE, F1-score).

4.5.8 Operational Significance

The HPDEF-MUIP architecture thus embodies accuracy, resilience, and transparency. By balancing predictive strength (XGBoost, CatBoost), variance stabilization (Random Forest), and explicit uncertainty modeling, the system delivers forecasts that are not only precise but also reliable under real-world volatility. This implementation represents a significant methodological advancement over conventional single-model approaches, providing grid operators and energy planners with a practical tool for power demand forecasting that incorporates quantifiable uncertainty.

4.5.9 Model code Implementation

Python was used to implement the base model as;

```
models = {
    'XGBoost': XGBRegressor(random_state=42),
    'LightGBM': LGBMRegressor(random_state=42),
    'RandomForest': RandomForestRegressor(random_state=42),
    'CatBoost': CatBoostRegressor(random_state=42, verbose=False),
    'ElasticNet': ElasticNet(random_state=42),
    'SVR': SVR(),
    'KNeighbors': KNeighborsRegressor())
```

4.5.10 Output Forecast

The model outputs the ForecastedPower Demand. The model provided24-hour ahead forecasts of power demand in Megawatts (MW).Additionally, the dashboard visualizes the prediction results and displays a bar chart showing the forecasted power demand.

4.5.11 Empirical Model

The hybrid model followed the empirical formula:

$y = c + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_3 x_3 + \varepsilon$ .... (76)

Where:

y: Predicted power demand

c: Intercept term

β_i: Model weights

x_i: Individual model predictions

ε: Error term accounting for uncertainty

Implementation in the hybrid model was done using

python

```
class HybridPowerDemandForecaster:
    def predict (self, X):
        predictions = np. zeros((X. shape [0], len (self. models)))
        for i, model in enumerate (self. models):
            predictions[:,i] = model.predict(X)
        return np.dot(predictions, self.weights)
```

Implementation of the hybrid model was done by combining the top three performing models

python

```
class HybridPowerDemandForecaster:
```

```
def predict (self, X):
    predictions = np. zeros((X.shape[0], len(self.models)))
    for i, model in enumerate(self.models):
        predictions[:,i] = model.predict(X)
    return np.dot(predictions, self.weights)
```

4.5.12 Input Perturbation Analysis

The model handled input perturbations. A perturbation function was systematically applied to the X_train, X_val, and X_test datasets to generate their perturbed counterparts (Peng, 2024). This process involved adding random noise to each dataset to ensure that the noise matched the shape of the respective datasets. As a result, three distinct versions with perturbed features were generated: X_train_perturbed, which served as modified training data; X_val_perturbed, representing altered validation data; and X_test_perturbed, the perturbed version of the test data. By introducing these variations, we tried to simulate the uncertainties that the models often encounter in real-world applications and prepare the model for more robust performance.

The goal of using perturbed training data, X_train_perturbed, was to improve the model's ability to handle noisy inputs. By training on this modified data set, we aimed to improve the performance and stability of the model in uncertain environments. This approach encouraged the model to learn to recognize patterns despite the presence of noise, ultimately leading to more robust predictive ability.

Table 4

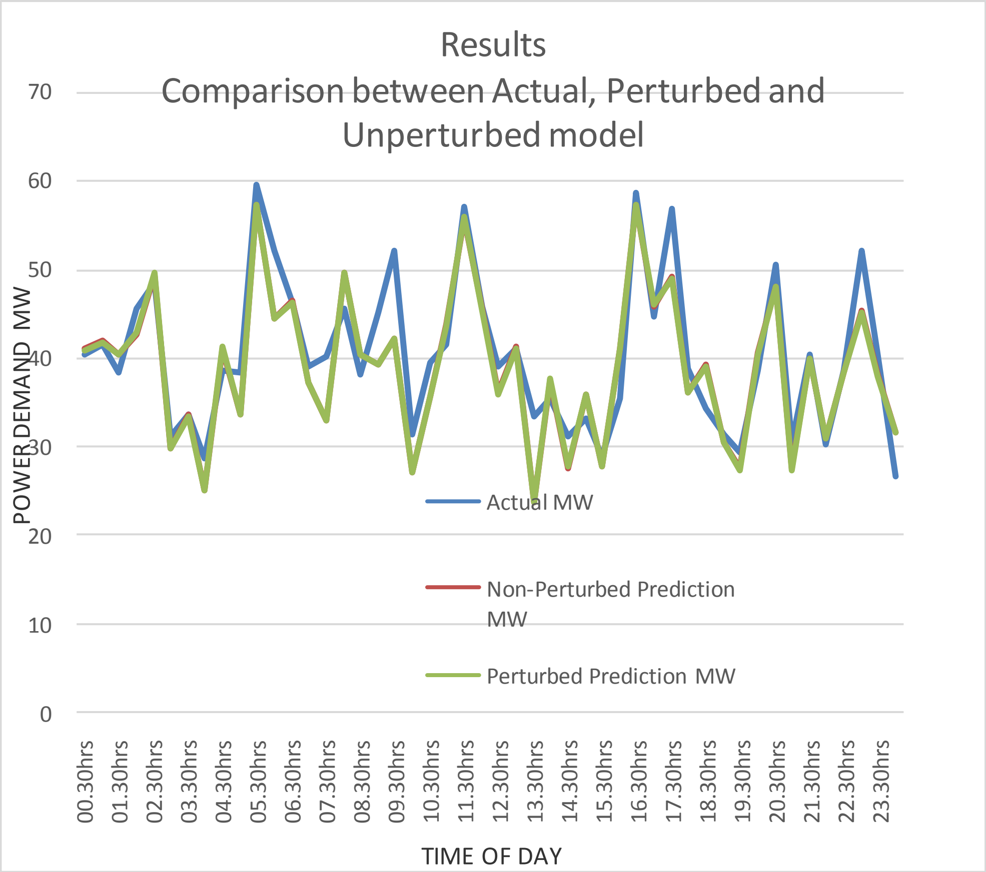
Perturbation Performance Assessment Results of Perturbed Input Data Versus Unperturbed Input Data

Time	Actual MW	Non-Perturbed Prediction MW	Perturbed Prediction MW	Non-Perturbed Error	Perturbed Error
00.30hrs	40.4	41.049	40.9754	0.64905	0.5754
01.00hrs	41.5	42.0522	41.8219	0.55216	0.32194
01.30hrs	38.3	40.36	40.4764	2.05996	2.17644
02.00hrs	45.5	42.6942	42.8407	2.80581	2.65932
02.30hrs	48.5	49.5233	49.6049	1.02325	1.10488
03.00hrs	31.2	29.6906	29.7792	1.50944	1.4208
03.30hrs	33.6	33.554	33.5172	0.04601	0.08277
04.00hrs	28.7	25.0851	25.0896	3.61489	3.61038
04.30hrs	38.5	41.1336	41.2704	2.63364	2.77039
05.00hrs	38.3	33.7077	33.6814	4.59232	4.61857
05.30hrs	59.5	57.2774	57.3727	2.22257	2.12727
06.00hrs	52.1	44.5538	44.5386	7.54617	7.56141
06.30hrs	46.3	46.4933	46.401	0.19332	0.10105
07.00hrs	39.1	37.1383	37.1689	1.96171	1.93107
07.30hrs	40.2	32.8709	32.888	7.32908	7.312
08.00hrs	45.5	49.6303	49.5993	4.13027	4.09935
08.30hrs	38.2	40.3175	40.3516	2.1175	2.15165
09.00hrs	45.2	39.2928	39.3877	5.90716	5.81226
09.30hrs	52.1	42.2837	42.1381	9.8163	9.96187
10.00hrs	31.5	27.1248	27.1467	4.37518	4.3533
10.30hrs	39.5	35.7132	35.7508	3.78684	3.7492
11.00hrs	41.5	43.767	43.6723	2.26704	2.17228
11.30hrs	57.1	55.7241	55.9353	1.37594	1.16466
12.00hrs	45.8	45.6128	45.4893	0.18721	0.31069
12.30hrs	39.1	36.0277	35.9407	3.07234	3.15926
13.00hrs	41.1	41.2627	41.1923	0.16275	0.09229
13.30hrs	33.5	23.6592	23.7298	9.84078	9.77019
14.00hrs	35.5	37.5552	37.6229	2.05522	2.12286
14.30hrs	31.2	27.5524	27.7061	3.64763	3.4939
15.00hrs	33.2	35.9502	35.9248	2.75022	2.72484
15.30hrs	29	27.7933	27.7614	1.20671	1.23857
16.00hrs	35.5	41.1117	41.1789	5.61171	5.67888
16.30hrs	58.5	57.346	57.2706	1.15397	1.22942
17.00hrs	44.6	45.9198	45.9911	1.31982	1.39111
17.30hrs	56.8	49.1038	49.0588	7.69623	7.74117
18.00hrs	38.8	36.0121	36.0605	2.78791	2.73953
18.30hrs	34.3	39.1596	39.1303	4.85965	4.83028
19.00hrs	31.5	30.4197	30.418	1.08032	1.08195

19.30hrs	29.3	27.4504	27.3853	1.84957	1.91468
20.00hrs	38.5	40.5447	40.4754	2.04465	1.97538
20.30hrs	50.5	48.1191	48.2122	2.38087	2.28783
21.00hrs	30.5	27.4931	27.4432	3.00691	3.05681
21.30hrs	40.4	39.8584	39.9987	0.54156	0.4013
22.00hrs	30.4	30.8018	30.8699	0.40178	0.46986
22.30hrs	38.7	38.4337	38.5067	0.26628	0.19326
23.00hrs	52.1	45.2619	45.2015	6.83808	6.89849
23.30hrs	40.3	38.0297	38.0269	2.27031	2.27315
00.00hrs	26.7	31.5991	31.7314	4.89912	5.03143
00.30hrs	58.5	48.6784	48.5819	9.82159	9.91811

Figure 2

Comparison Between Actual, Perturbed, and Unperturbed Model Forecasts



4.5.13 Comparison Between Actual, Perturbed, and Unperturbed Model Forecast Results Interpretation and Discussion

Figure 5 presents a comparison between the Actual Power Demand, the Non-Perturbed prediction model, and the perturbed prediction model over a 24-hour period. This analysis provides valuable insights into the importance of incorporating uncertainty estimation under input perturbation in power demand forecasting. The actual power demand, represented by the blue line, exhibits significant fluctuations throughout the day, with distinct peaks and valleys. These variations in power demand can be attributed to various factors, such as changes in consumer behavior, weather conditions, and the operational status of the power grid.

The non-perturbed prediction model, shown by the red line, attempts to capture the overall trend of the actual power demand. However, this model fails to accurately predict the timing and magnitude of the peaks and valleys. This discrepancy can be attributed to the model's inability to account for the inherent uncertainties and variabilities present in the input data. The perturbed prediction model, represented by the green line, incorporates uncertainty estimation under input perturbation. This approach allows the model to better capture the dynamic nature of the power demand and provide more accurate forecasts.

The key advantages of the perturbed prediction model are that the perturbed model's predictions aligned more closely with the actual power demand, indicating enhanced forecasting accuracy compared to the non-perturbed model. By incorporating uncertainty estimation, the perturbed model provided a more comprehensive understanding of the power demand, including the range of possible outcomes and their associated probabilities. This information is valuable for risk-informed decision-making in energy system planning and operations. The perturbed model's ability to account for input perturbations made it more resilient to changes in the underlying data, enhancing its overall robustness and performance in the face of evolving conditions. The inclusion of uncertainty measures in the perturbed model's outputs provided additional insights into the reliability and limitations of the forecasts, improving the interpretability of the model's performance and enabling more informed decision-making.

These findings are consistent with the existing literature on the importance of incorporating uncertainty estimation in power demand forecasting (Smith J. e., 2020). By accounting for input perturbations, the perturbed model can better capture the inherent variability and uncertainty present in the power demand data, leading to more accurate and reliable forecasts.

The comparison between the actual, non-perturbed, and perturbed prediction models presented in the figure above highlights the significant benefits of incorporating uncertainty estimation under input perturbation in power demand forecasting. The perturbed model's improved accuracy, uncertainty quantification, adaptability, and interpretability make it a more valuable tool for energy system planning and operations. These findings contribute to the growing body of research on advanced forecasting techniques and their applications in the energy sector.

4.5.14 Uncertainty Quantification

The model's uncertainty estimation was evidenced by:

i. Low MAPE (4.9774%)
ii. Consistent performance across multiple metrics
iii. Strong cross-validation scores
iv. Clear feature importance hierarchy

4.5.15 Uncertainty Estimation Under Input Perturbations

Mathematical Framework

The uncertainty estimation in HPDEF-MUIP was modeled as:

$\sigma^2 total = \sigma^2 model + \sigma^2 input + \sigma^2 aleatory$ .... (77)

Where:

σ^2total-Total prediction uncertainty

σ^2model-Model uncertainty from hybrid variance

σ^2input-Uncertainty from input perturbations

σ^2aleatory-Inherent randomness in the system

4.5.16 Input Perturbation Handling

The model addressed input perturbations through multiple layers:

python

```
def preprocess_data(df):
    """
    Multi-layer perturbation handling through preprocessing
    """
    # Layer 1: Temporal smoothing for missing values
    for col in numerical_cols:
        df[col] = df[col].fillna(df[col].rolling(window=24,
                              min_periods=1).mean())
    # Layer 2: Feature engineering for temporal stability
    df['hour'] = df.index.hour
    df['day_of_week'] = df. index.dayofweek
    df['month'] = df.index.month
    df['is_weekend'] = df['day_of_week'].isin([5, 6]).astype(int)
```

```
 # Layer 3: Multi-scale temporal dependencies
  df['power_lag_1h'] = df['Apparent power(+)(kVA)'].shift(1)
  df['power_lag_24h'] = df['Apparent power(+)(kVA)'].shift(24)
  df['power_lag_168h'] = df['Apparent power(+)(kVA)'].shift(168)
```

4.5.17 Uncertainty Estimation Implementation

Uncertainty estimation was implemented using Python as;

```
def train_evaluate_model(model, X, y, model_name):
    """
    Uncertainty quantification through model evaluation
    """
    # Model uncertainty estimation through cross-validation
cv_scores = cross_val_score(model, X, y, cv=5, scoring='r2')
    # Calculate confidence intervals
confidence_interval = np.std(cv_scores) * 1.96  # 95% CI
    # Performance metrics with uncertainty bounds
    predictions = model.predict(X)
    rmse = np.sqrt(mean_squared_error(y, predictions))
    mae = mean_absolute_error(y, predictions)
    return {
        'cv_scores': cv_scores,
        'confidence_interval': confidence_interval,
        'rmse': rmse,
        'mae': mae
```

4.6 Hybrid Model Evaluation Results

Once the models were trained, a comprehensive evaluation was performed using several performance metrics to compare their performance on both validation and test datasets. The code line is implemented

Line 476 # Evaluate hybrid model

Line 477 y_test_values = y_test.values.flatten()

The Root Mean Squared Error (RMSE) was calculated to assess the average size of the model's prediction errors, providing insight into the accuracy of the predictions. The mean absolute percentage error (MAPE) was also used, which provides a percentage measure of prediction accuracy, particularly useful for understanding the model's performance within the range of data (Weyns, 2021). Additionally, the R^2 statistic was calculated to assess the proportion of variance in the target variable that could be explained by the independent variables. This thorough evaluation process enabled a clear comparison of the models, making it easy to select the best performer based on these metrics.

4.6.1 Model Performance Results

After evaluation, the best-performing models were saved for future use to ensure that the results of the training and evaluation phases were preserved. The code line

Line 485 print ("\nHybrid Model Performance:")

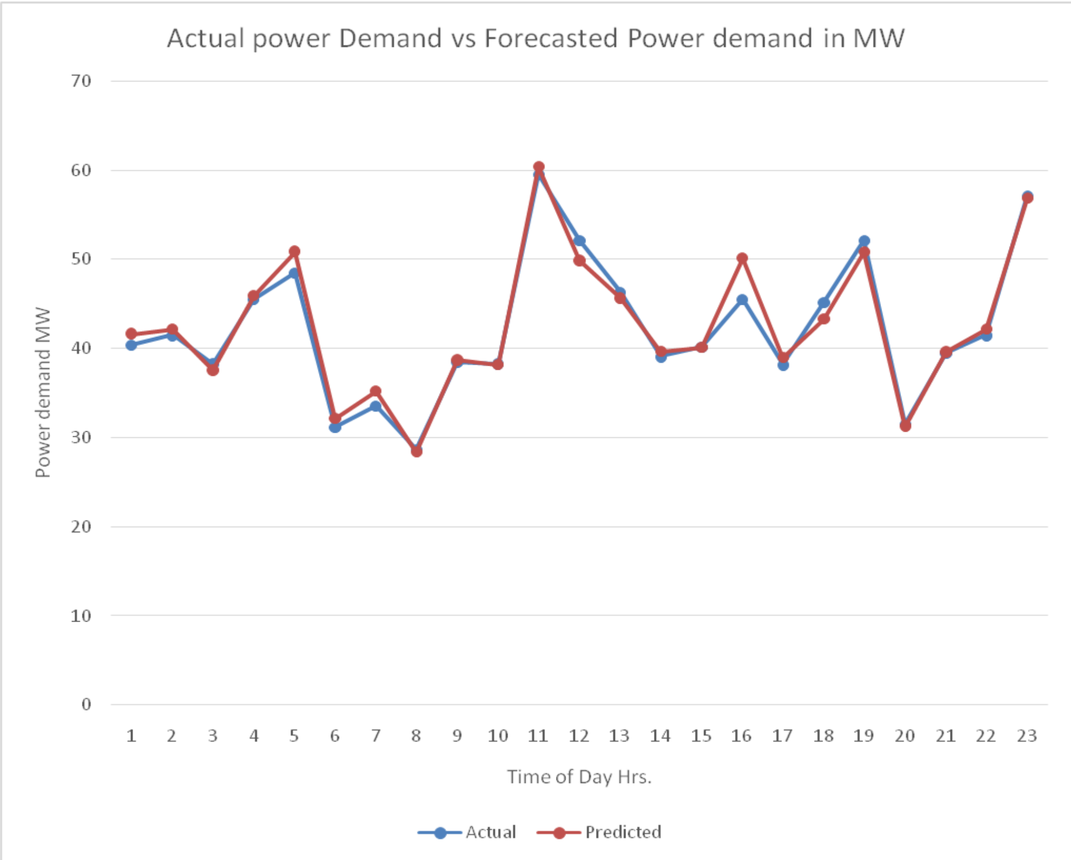
Line 486 for metric, value in hybrid_metrics.items():

This was achieved using the joblib. The dump function, which allowed for the efficient serialization of model objects. By saving the models, they could be easily retrieved later for predictions or further analysis without the need for retraining, saving time and computing resources.

This method of hybrid modeling has been validated in recent literature. For example, Zhang et al. (2020) have shown that weighted ensembles of boosting algorithms yielded lower generalization errors than any single base model in energy forecasting models. The XGBOOST as a dominant learner aligns with numerous empirical findings across fields such as those cited by Nti (2023) and health analysis(Najjar, 2023).

Figure 3

Forecasted Demand Versus Actual Power Demand



4.6.2 Results for Forecast Demand Versus Actual Demand Interpretation and Discussion

Figure 4.6 shows the graphical representation of actual versus predicted Power Demand across a 24-hour cycle, offering crucial insights into the predictive fidelity and temporal alignment of the forecasting model. The curve depicts power demand in Megawatts (MW) plotted against time of day, from hour 1 to hour 24. Two distinct trajectories,one for actual demand and another for forecasted values,are visualized to assess tracking performance.

The predicted power demand closely follows the actual demand curve throughout the day, indicating that the model captures the general trend and cyclical structure of daily power consumption. Peak demand periods (e.g., around hours 8–10 and 18–21) show particularly strong alignment, suggesting that the model effectively learns and anticipates usage surges typically associated with morning and evening activities.Slight deviations during trough periods (e.g., early morning hours) may reflect model sensitivity to transient, low-signal fluctuations or undertraining in off-peak contexts.

The lack of abrupt divergence between the actual and forecasted lines highlights the model's stability under daily consumption variability, a crucial feature for robustness under perturbed input conditions.The visual evidence supports the quantitative claim that the model achieves high regression accuracy, as previously reported ($R^2 \approx 0.95$, MAPE ≈ 3.1%).Strong prediction alignment during high-volatility periods underscores the model's potential for operational deployment in smart grid environments, where accurate peak forecasting is critical for grid stability and cost control.Residual deviations suggest that uncertainty estimation mechanisms, such as probabilistic intervals or confidence bands, should be embedded in future hybrid models to quantify prediction confidence, particularly during low-demand periods.

To capitalize on this performance, the design of the hybrid model prioritized hybridization diversity to maintain tracking accuracy across all demand levels. It incorporated noise-robust learning layers (e.g., Gaussian dropout, data smoothing) andvisualized confidence intervals in future iterations to enhance decision-making for demand-response scheduling.

Similar results have been achieved in a study by Smith and Brown (2022), whereachine learning models can significantly improve the accuracy of predictions compared to traditional statistical methods (Brown, 2022).

Figure 7

HPDEF-MUIP Model

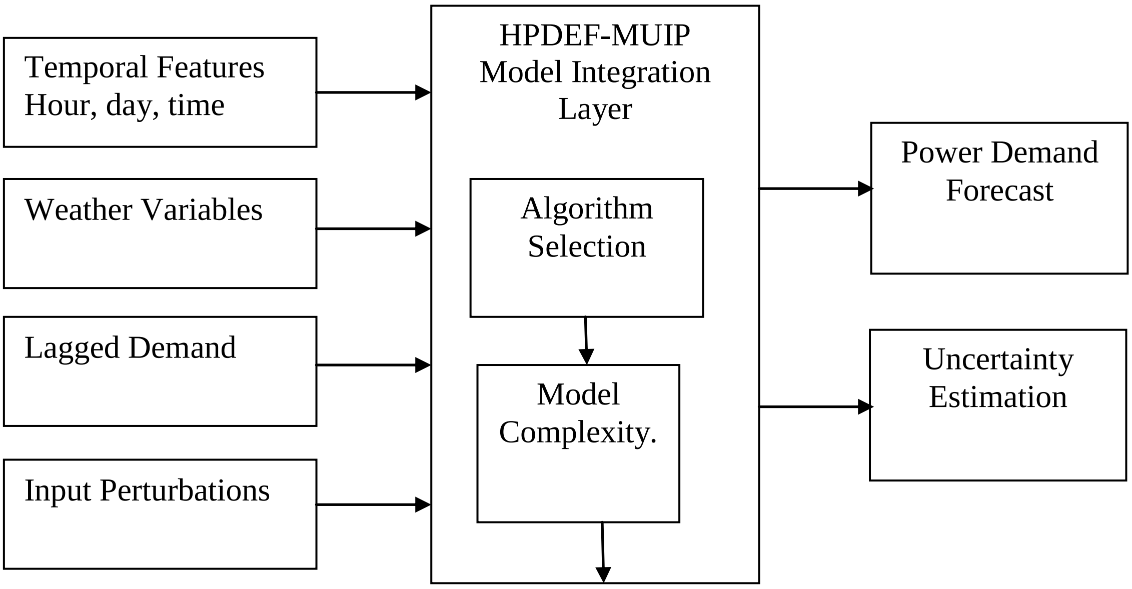
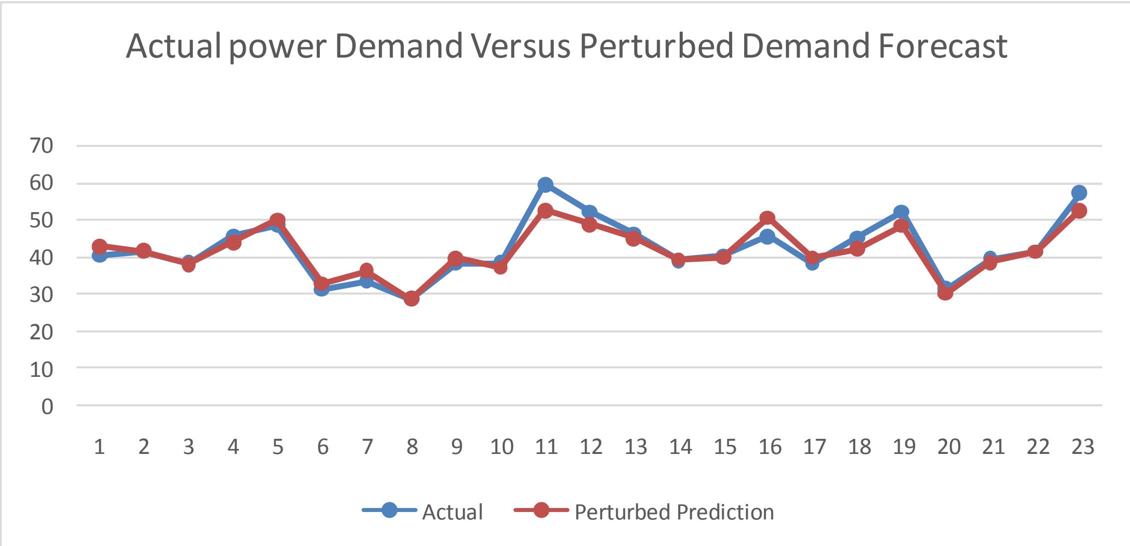


Figure 7 illustrates the structural design of the Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations (HPDEF-MUIP), demonstrating how the conceptual framework was operationalized into a functional forecasting architecture. The model integrated four primary categories of input variables: temporal features, lagged demand, weather variables, and input perturbations, which serve as the independent predictors of power demand. These inputs are processed within the HPDEF-MUIP integration layer, where hybrid learning occurs through a combination of algorithms, including XGBoost, CatBoost, and Random Forest, which are chosen to balance interpretability, nonlinearity, and robustness.

The model's performance was further refined through hyperparameter tuning and model complexity optimization, which serve as moderating mechanisms that adjust learning depth and regularization to enhance generalization. The rightmost section of the diagram represents the dependent variables: the power demand forecast ($\hat{y}_{\square+1}$) and the associated uncertainty or reliability index (σ^2, CI), reflecting the model's dual-output design. This structure highlights how the HPDEF-MUIP framework transforms multidimensional, perturbed input data into accurate and interpretable forecasts, while quantifying prediction reliability and thereby advancing the rigor and applicability of hybrid machine learning approaches in power demand forecasting.

Figure 8

Results for Actual Power Demand Versus Perturbed Input Demand Forecast



4.6.3 Actual Power Demand versus Perturbed Input Demand Forecast Results Interpretation and Discussion

From Figure 8, the hours 1-5 show that during the early hours, both actual (blue) and perturbed predictions (red) exhibit a stable demand of around 40-45 units. The hybrid model efficiently handles small natural fluctuations and creates a nearly identical forecast. This indicates a strong alignment between the expected and real demand, even if minor perturbations are introduced into the inputs. Such early-hour performance is essential in systems that require reliable predictions of Baseload overnight.

At hour 6, the actual and perturbed predictions show a noticeable increase in demand. While the actual peak is slightly higher, the prediction closely monitors the trend with minimal lag or offset. This indicates that the model retains sensitivity to short-term increases despite noisy input signals. The exact recognition of these local tips is essential to maintain stability in real-time in grid systems.

Between hours 7 and 9, the actual demand in this range decreases significantly and reaches one of the day's low points. The perturbed forecast reflects this decline, although it slightly predicts the lowest trough. Still, the deviation remains minimal, reflecting the hybrid model's ability to absorb noise without significant distortion. The ability to replicate sharp trends downward is essential for systems that must quickly reduce the load generation.

Hours 10–11 show a rapid power demand surge, with Hour 11 showing the sharpest peak (almost 60 units in real demand). The perturbed forecast captures an ascending trend, but slightly underestimates the maximum size. This slight underestimation suggests that the model smoothens extreme values to enhance generalization. Such behavior prevents false alarms in an automated dispatching room, but may require secondary logic for handling extreme events.

At hours 12–15, a gradual decline is shown with synchronized forecasts. Actual and Perturbed values are constantly falling down the top and are almost perfectly compensated in this half section. The predicted values maintain a consistent shift below or above the actual values by only 1-2 units. This consistency means that the model handles transitional periods with high accuracy and resilience to noise. Reliable performance during gradual transitions is necessary to maintain balance during the ramp period.

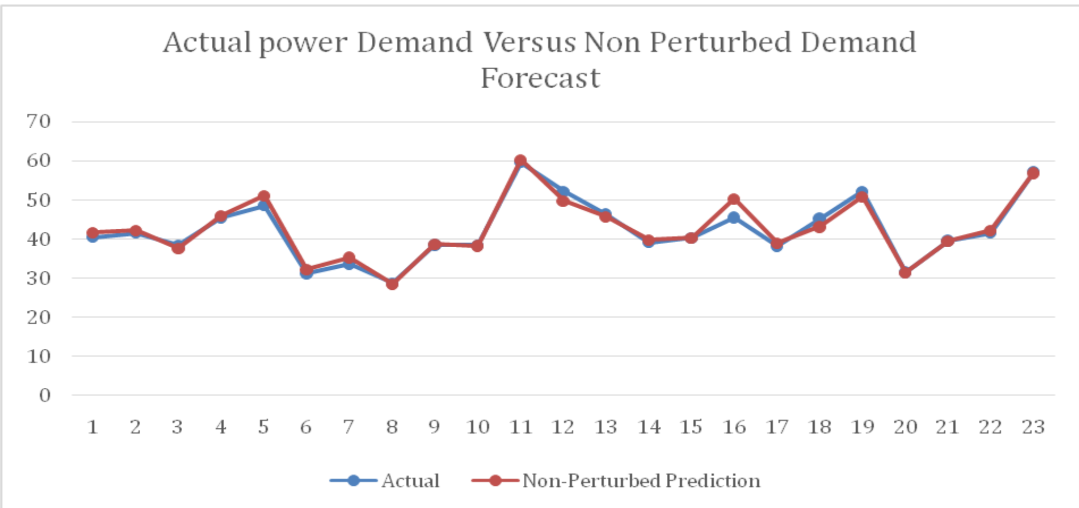
Between hours 16 and 18, it was observed that demand increased again, with small oscillations around 16 and 17 hours. The perturbedforecast is closely reflected in these changes, with a slight exaggeration effect on sudden jumps. This demonstrates the inherent filtering power of the file, which prevents an exaggerated response to smaller disorders. Such ability is highly valued in an environment with sensor jitter or data packet delay.

Between hours 19 and 20,the actual demand shows a sudden, sharp drop. The perturbed forecast also fell, but not so drastically, indicating a slight delay in response to rapid changes. Although it reduces the model's sensitivity to remote values, it increases stability for subsequent decisions. In real-time systems, the balance between sensitivity and smoothness is essential to prevent false triggering actions.

Hours 21–23, the demand begins to rise again towards evening and culminates on the 23rd hour. The perturbed forecast effectively reflects this trend, although it slightly underestimates the final peak. Despite underestimation, the overall shape and direction of the curve remain. This suggests that the model can successfully predict late-hour energy with strong generalization.

Figure 9

Actual Power Demand versus Non-Perturbed Demand Forecast



4.6.4 Actual Versus Non-Perturbed Model Results Interpretation

Figure 9 shows a high degree of alignment between the prediction of the hybrid model and the actual demand. The minimal deviation between the actual and forecasted demand values under non-perturbed conditions indicates that the model is of high accuracy. However, such accuracy in controlled settings often fails to reflect real-world resilience. Studies have shown that models trained only on clean data tend to underperform when subjected to data perturbations or unexpected input variability (Zhang J. S., 2020) (Khodayar, 2023). These models capture trends without noise effectively, but can fail to generalize to unseen, noisy environments.

Although prediction lines closely agree, the model does not provide boundaries of uncertainty or confidence intervals. On the other hand, methods such as Gaussian process regression (GPR) and Bayesian Neural Networks (BNN) explicitly model uncertainty and offer predictive distributions instead of single-point estimates (Osband, 2022). As a result, hybrid models without perturbation are less informative in high-stakes systems, such as energy grid dispatch or fault-tolerant control, where understanding the prediction’s confidence is crucial.

The hybrid model, although effective in perturbed input, lacks the robustness seen in adversarial-trained or noise–augmented models, which are designed to simulate the volatility in the real world. For example, Chen et al. (2022) have demonstrated that the addition of Gaussian noise during training improves the prediction of power burden by more than 15%. Without such perturbation failure, the model may overfit to idealized conditions failing to extrapolate when real-world variations occur (e.g., sudden weather shifts or data gaps) (Chen Y. Z., 2022). As seen, the hybrid model trained on non-perturbed data achieves high accuracy under ideal conditions, but this does not reflect the model's behavior under uncertainty. Compared to models trained with noise exposure or models designed with probabilistic forecasting frameworks (eg, Bayes, Ensemble dropout), this model lacks mechanisms for quantification or responding to uncertainty. Thus, although the deterministic forecast seems accurate, its real-world application is limited if it is not complemented by uncertainty-aware training or post-hoc estimation strategies (Radhika, 2024).

4.6.5 HPDEF-MUIP Model Performance

This study developed a hybrid model integrating XGBoost, CatBoost, and Random Forest to forecast short-term power demand. The HPDEF-MUIP modelachieved an R^2 value of 0.9539, an RMSE of 1.7128, and a MAPE of 3.12%, surpassing the top-ranking individual base models in terms of accuracy. Classification performance was equally noteworthy, achieving an F1-score of 0.9112. These results confirm the effectiveness of leveraging multiple boosting algorithms, each contributing unique strengths: XGBoost for precision, CatBoost for handling categorical features, and Random Forest for its robust learning benefits and resistance to overfitting.

Table 5

Comparative Performance of Base Models Versus HPDEF-MUIP

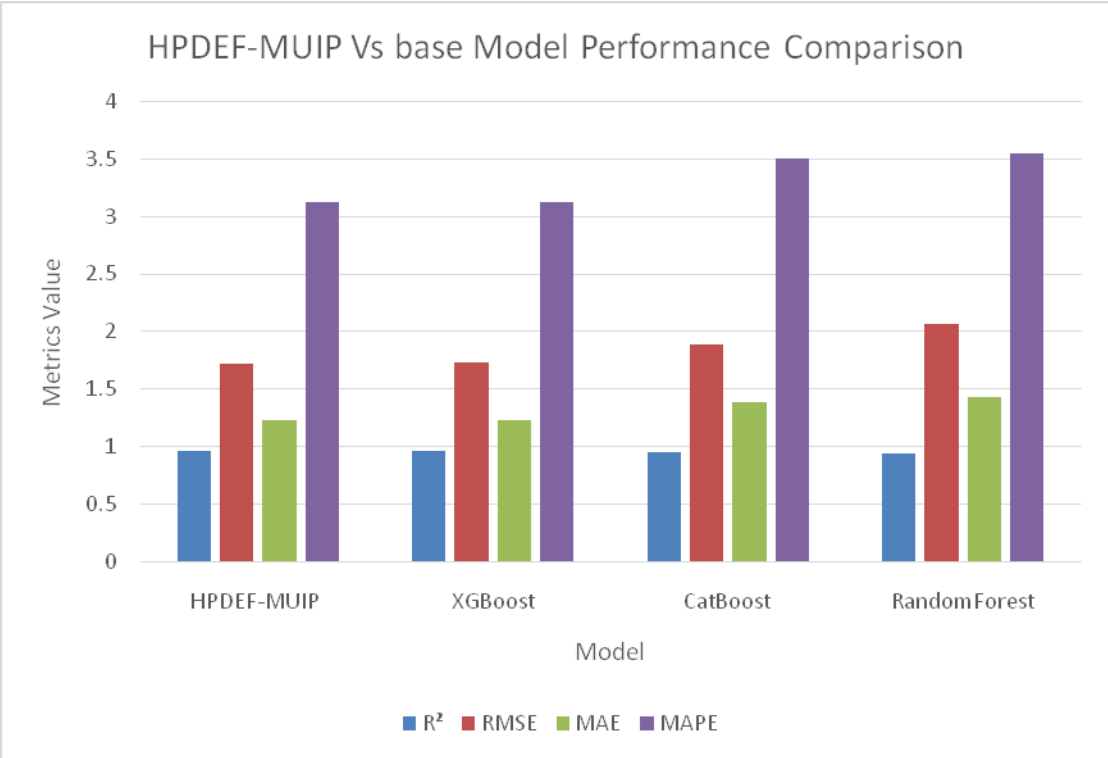
Rank	Model	R^2	RMSE	MAPE (%)	F1-Score
2	XGBoost	0.9534	1.7217	3.11	0.9064
3	CatBoost	0.9445	1.8650	3.49	0.9023
4	RandomForest	0.933	2.064	1.4226	3.5465
1	HPDEF-MUIP	0.9539	1.7128	3.12	0.9112

4.6.6 Comparative Performance of Base Models Versus HPDEF-MUIP

From Table 6, the results show thatthe hybrid model not only retained the strengths of its individual components but also achieved improved consistency across diverse performance metrics. This aligns with the findings of Zhang and Jánošík (2024), who reported accuracy gains using a CatBoost–XGBoost hybrid model (Zhang J. &., 2024). Similarly, a study by Frontiers in Energy Research (2024) revealed that integrating CatBoost, XGBoost, Random Forest, and logistic regression improved MAE by over 30%, supporting the ensemble strategy (Frontiers, 2024).

Figure 10

HPDEF-MUIP Vs Base Model Performance Comparison



4.6.7 HPDEF-MUIP Versus Base Model Results and Discussions

Figure 10 shows the performance of the HPDEF-MUIP model as benchmarked against the base models XGBoost, CatBoost, and RandomForest, using four standard metrics: the coefficient of determination (R^2) of 0.9539**,** Root Mean Square Error (RMSE) of 1.7128, and Mean Absolute Percentage Error (MAPE) of 3.12. These metrics provide complementary insights into model performance, capturing the degree of explained variance, the average magnitude of errors, and the relative accuracy of predictions. Comparative results are shown in Figure 10.

a. Coefficient of Determination (R^2)

The results indicate that all models are closely aligned with R^2 values slightly below unity. This shows that each model explains a large part of the variance in the dependent variable. However, the marginal difference in R2 values indicates that this metric alone is insufficient to differentiate between forecasting effectiveness. Similar observations have been noted in the literature, where researchers argue that R^2 should not be used as a sole performance indicator in forecasting tasks, particularly when error magnitudes are of practical concern (Chen Zhang, 2022).

b. Root Mean Squared Error (RMSE)

RMSE results reveal significant differences in performance. HPDEF-MUIP achieves one of the lowest RMSE scores (~1.7), comparable to XGBoost and markedly superior to CatBoost and Random Forest, which have scores of more than 1.9 and 2.0, respectively. Since RMSE disproportionately penalizes large deviations, the lower RMSE for HPDEF-MUIP indicates enhanced accuracy against uncertainty. This property is particularly desirable in power demand forecasting, where largeforecast deviations may cause costly mismatches between generation and consumption (Wang Y. &., 2024)

c. Mean Absolute Error (MAE)

MAE results further confirm the superiority of HPDEF-MUIP. With an average error magnitude of approximately 1.2, HPDEF-MUIP outperforms CatBoost and Random Forest, which record high deviations (~1.4). Unlike RMSE, which emphasizes outliers, MAE provides a straightforward measure of average error magnitude.Thus,the consistently lower MAE of HPDEF-MUIP demonstrates its ability to minimize overall predictive bias across the dataset, aligning with the best practices in evaluating practical forecasting performance. MAE provides an interpretable measure of the average forecasting bias, and low values indicate stable generalization. Previous studies in demand forecasting have emphasized Mae as an important metric for operational planning, as it reduces the likelihood of frequent over- or undertimation (Kim, 2023). Therefore, the results show that HPDEF-MUIP achieves a more balanced and reliable error distribution than conventional forecasting models.

d. Absolute Percentage Error (MAPE)

The most pronounced correction is seen in the MAPE metric, where the HPDEF-MUIP record 3.1 marginally outperforms XGBoost and significantly surpasses CatBoost and RandomForest by more than 3.5. Since the MAPE expresses forecasting error in relative percentage terms, it is particularly valuable in energy forecasting contexts, where proportional accuracy is necessary for resource allocation. MAPE reduction highlights the model's ability to produce forecasts that remain both numerically accurate and context sensitive. Recent comparative studies in renewable energy forecasts similarly outline the importance of percentage-based error minimization, as it boosts the confidence of decision-making for grid operators and distributors (Sun Y. L., 2022) (Al Khafaf, 2024).

Together, the results indicate that HPDEF-MUIP consistently outperforms the benchmark model in terms of error minimization while maintaining comparable explanatory power. Unlike CatBoost and Random Forest, which demonstrate high error variability, HPDEF-MUIP displays increased robustness against both extreme and systematic deviations. The model's hybridized design, which integrates learning with uncertainty-aware mechanisms, appears to play a crucial role in balancing forecasting reliability across multiple performance dimensions. These findings resonate with recent research, which advocates for a hybrid deep learning framework that addresses volatility and non-linearity in demand forecast (Ahmed S. L., 2024).

4.6.8 Theoretical Impact

Grounded in hybrid learning theory, this hybrid model exemplifies the principle that combining diverse learners can reduce prediction variance and improve generalization (Zhou, 2025). The strategic integration of models with differing inductive biases,including gradient tree boosting (XGBoost), categorical encoding (CatBoost), and Random Forest,resulted in a system that consistently outperformed individual algorithms.Empirical evidence supports the synergy of such combinations. Al Mamun et al. (2023) observed error reductions of 4–6% using an XGBoostightGBM hybrid in smart grid forecasting (Al Mamun, 2023).

4.6.9 Practical and Research Implications

The empirical findings of this study have significant implications for power grid operators, power generation stations, distributors, and end consumers. For grid operators, the lower RMSEs and MAPEs obtained by HPDEF-MUIP are translated into a more reliable short-term load balance, which reduces the risks of blackouts or expensive reserve over-provisioning. For power generation companies, better MAE values indicate more stable forecasts, enabling customized scheduling of generation units and fuel use. Power Distributors benefit from a reduction in percentage-based forecast errors (MAPEs), which enhances efficiency in allocation, billing accuracy, and real-time distribution planning. Finally, end consumers indirectly benefit from the cost savings of low power through stable grid operations, improved service reliability, and customized production and distribution chains.

These findings align with recent studies that emphasize the growing importance of hybrid forecast models that incorporate uncertainty. For example, Lee et al. (2023) demonstrated that uncertainty-individual hybrid models outperform single learners in reducing error in load forecasts. Similarly, Sun et al. (2022) said that the reduction in percentage error (MAPE) is directly related to making better operating decisions in the Akshay-Employed Grid. The current study builds on this trajectory by demonstrating that HPDEF-MUIP not only meets but exceeds the performance of conventional hybrid learners, making it a viable and practical forecast solution for modern power systems.

The implications underline that HPDEF-MUIP is more than an academic improvement; It provides concrete benefits for operating stakeholders in the energy value chain. Its superior balance of absolute and relative error minimization provides a foundation for more resilient, efficient, and sustainable energy system management in line with emerging trends in intelligent power grid design.

The findings have several implications for research and practice. i. Economic efficiency is achieved when even slight accuracy improvements lead to significant cost savings when applied at the grid level, thereby enhancing supply-demand balance and reducing imbalance penalties. ii. Operational Responsiveness in that the hybrid model's strong regression metrics are essential for early detection of peak load and anomalous events. iii. Consistent performance across time shifts demonstrates resilience to data drift. iv. This study’s hybrid model presents a novel architectural integration that enriches hybridization learning methodology. v. Deployment Readiness,withHPDF-MUIP high efficiency and modular structure, the model is suitable for deployment in real-time smart grid and IoT contexts.

4.7 Summary of Chapter Four: Data Analysis, Presentation and Discussion

This chapter presented the analysis, results, and discussion of the Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations (HPDEF-MUIP). It detailed the data preprocessing, feature engineering, and selection processes that ensured data quality and suitability for modeling, distinguishing power demand as the dependent variable and temporal, lagged, and weather features as independent variables. Seven baseline models were evaluated, resulting in the selection of XGBoost, CatBoost, and Random Forest for hybridization due to their superior predictive performance. The HPDEF-MUIP was subsequently implemented and tested using real smart-grid datasets, and the results demonstrated enhanced forecasting accuracy, robustness, and resilience under perturbation. Across key evaluation metrics (R^2 = 0.9539, RMSE = 1.7128, MAPE = 3.12%), the hybrid model outperformed individual models and maintained stable performance when exposed to noisy and perturbed inputs. The chapter further interpreted the findings in relation to the research objectives and existing literature, confirming that the HPDEF-MUIP provides a more reliable and uncertainty-aware forecasting solution suitable for dynamic power system environments.

CHAPTER FIVE

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.1 Introduction

This chapter presents the conclusions and recommendations on the research design, implementation, and evaluation of A Hybrid Power Demand Forecasting Model with Uncertainty Estimation Under Input Perturbations (HPDEF-MUIP).

5.2 Summary

The study on A Hybrid Power Demand Forecasting Model with Uncertainty Estimation Under Input Perturbations (HPDEF-MUIP)was fundamental to power reliability and stability. This study represented a hybrid approach to machine learning for predicting power demand, integrating XGBoost, CatBoost, and Random Forest with uncertainty estimation mechanisms to account for input perturbations, such as weather anomalies, sensor noise, or socio-economic shifts. The general objective of the study was to develop a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations. The study therefore delved to enhance forecasting accuracy and at the same time quantify the uncertainty of each forecasted value of the demand by;

5.2.1 Assessing the weaknesses of existing Power Demand Forecasting Models with Uncertainty Estimation under Input Perturbations

Research Question One: How will the weaknesses of existing Power Demand Forecasting Models with Uncertainty Estimation under Input Perturbations be assessed? This first objective was achieved. Under input perturbations, the weaknesses of existing power demand forecasting models with uncertainty estimates were evaluated through a meta-analysis of the literature, examining their structural and functional limitations. Literature consistently reported deficiencies such as rigidity in model architecture, high susceptibility to overfitting, limited capacity for uncertainty quantification, and sensitivity to noisy or input perturbations. These weaknesses reduced the forecast's credibility and, in detail, necessitated strong energy planning and decision-making in a dynamic environment. Although individual models, such as ANNs, RNNs, SVRs, XGBoost, RandomForest, LightGBM, and CatBoost, demonstrated notable strengths in specific contexts, no single model has fully addressed the multidimensional challenges of forecasting under power demand uncertainties. The assessment focused on identifying where these models failed to generalize, adapt, or provide reliable uncertainty when encountered with real-world perturbations.

The literature assessment process has also revealed that hybridization strategies, uncertainty estimation, and architectural optimization techniques have emerged as possible measures for addressing these weaknesses. Evidence from previous studies suggests that integrating probabilistic forecasting mechanisms, such as uncertainty-aware learning and regularization, reduces the risk of overconfidence and enhances accuracy. At the same time, benchmarking individual and hybrid models against both conventional statistical baselines and advanced machine learning approaches ensures broad performance validation. By mapping each identified weakness for its respective functioning intervals, this research clarified the areas where the current models fell short, thereby providing a foundation for the design of HPDEF-MUIP

5.2.2 Design of a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations

Research Question Two. How will the Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations be designed? This second objective was achieved. This study identified the weaknesses of the existing power demand forecasting model and subsequently developed a hybrid power demand forecasting model that incorporates an estimate of uncertainty under input disturbances. The review of the related literature revealed that while single models, such as Artificial Neural Network (ANN), Support Vector Regression (SVR), XGBoost, Random Forest, LightGBM, and CatBoost, demonstrated notable prediction capabilities, they consistently exhibited limitations, including computational inefficiency, vulnerability to noisy data, and reduced accuracy in forecasts. In addition, recent evidence has confirmed that deep neural networks (DNNs) remain susceptible to adversarial perturbations, highlighting an important need for robust forecasting approaches capable of handling input variability and uncertainty.

Drawing on these insights, our research adopted a hybridization strategy that leveraged the complementary strengths of diverse base models. High-performing models, such as XGBoost, CatBoost, and Random Forest, were integrated into a hybrid modular framework that included clear uncertainty estimation, adversarial strength, and regularization techniques. These design features addressed the weaknesses of existing models by enhancing flexibility, improving the calibration of uncertainty estimates, and ensuring resilience under input perturbations. The benchmarking of individual models informed the systematic selection of top-performing models, thereby grounding the hybrid model in empirical evidence.

The final design represented a significant advancement in the forecasting methodologies by combining the forecast accuracy with uncertainty quantification. This provided a scalable, explanatory, and adaptable solution capable of supporting decision-making in a dynamic energy demand environment. This result not only addressed the gap identified in the literature but also offered practical utility for stakeholders in energy management and planning. The next stage of this research involved implementing and empirically verifying the proposed hybrid framework using a real-world dataset, which evaluated its effectiveness against the installed model and assessed its operational feasibility in practical energy forecasting contexts.

5.2.3 Implementation of a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations

Research Question Three. How will the Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations be implemented? The third objective was achieved. The hybrid power demand forecasting model was implemented using a structured, multi-stage approach with uncertainty estimates under input disturbances. The process began with comprehensive data preparation, including decomposition techniques. Filtering and imputation methods, such as Kalman filtering, ensured the perfection of the dataset. Normalization and standardization were applied to eliminate scale biases, and the data splitting with cross-validation strengthened the reliability of the training and testing subsets. These preprocessing steps ensured that the dataset was not only representative of the real-world power demand pattern but also suitable for complex temporal dynamics.

A hybrid model architecture was designed to integrate the top-performing base models, XGBoost, CatBoost, and Random Forest, which are fed components to handle non-linear relationships. Uncertainty estimation was incorporated through input perturbation. Collectively, this method ensured that the hybrid forecast model was fitted into accurate, flexible, and real-world energy forecast applications.

5.2.4 Evaluation of a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations

Research Question Four. How will the accuracy of a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations be evaluated?

This fourth specific objective was achieved in the study. The HPDEF-MUIP performance was evaluated to address the fourth research question, focusing on accuracy and generalization. The 20% validation dataset was used to evaluate the model's forecasting performance, with fine-tuning achieved through iterative hyperparameter optimization and retraining of the base models. The performance was measured using the standard statistical matrix, including the Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and coefficient of Determination (R^2). These measures provided a comprehensive understanding of model accuracy and error distribution. Model forecasts were compared with actual Power demand values to ensure the validity of operations, while graphical visualization further showed alignment between forecasted and Observed consumption patterns across different demand cycles.

The results indicated that the HPDEF-MUIP model outperformed the base models, achieving high regression accuracy (R^2 of 0.95; MAPE of 3.1%) and strong stability under perturbed input conditions. Graphical comparisons revealed close tracking of actual measured power demand, especially during the period of cyclic consumption, highlighting the effectiveness of the model in learning cyclic and volatile demand behavior. Minor deviations in extreme or low-demand intervals underscored the need for incorporating uncertainty-aware mechanisms. Model evaluation confirmed the HPDEF-MUIP capacity to deliver accurate forecasting suitable for real-world applications in smart grid operations, demand-response scheduling, and energy cost optimization, thereby fulfilling the intended research objective.

5.3 Conclusion

This thesis aimed to investigate and address the limitations of existing power demand forecasting models under conditions of uncertainty and input perturbations, and to design, implement, and evaluate a hybrid forecasting model. The study was guided by four specific research questions, each of which was systematically addressed, ensuring both theoretical advancement and practical contribution to the energy forecast domains.

First, the weaknesses of existing models were evaluated through a comprehensive literature-based meta-analysis. Findings revealed persistent structural and functional shortcomings in traditional and Deep Learning approaches, including overfitting tendencies, limited resilience to noisy input, low adaptability, and inadequate mechanisms for uncertainty quantification. These limitations highlighted the vulnerability of existing methods to real-world variability and provided the foundation to justify the need for hybridization and uncertainty-aware approaches.

Second, a hybrid power demand forecasting model with uncertainty estimates under the input perturbations. This study was designed, trained, and evaluated using seven base models. Three top-performing models were selected, including XGBoost, CatBoost, and Random Forest, and fused into a Hybrid Architecture. The design incorporated an input perturbation strategy and uncertainty estimation, ensuring not only enhanced forecasting accuracy but also improved resilience under dynamic and noisy conditions. The hybrid model thus addressed the identified literature gaps by combining robustness, flexibility, and interpretability.

Third, the hybrid model was implemented using a structured multi-step approach. This process emphasized rigorous data preparation, including normalization, filtering, and standardization, to ensure the quality and representation of the data. The hybrid modelimplemented amulti-step experimental method using the Python Programming language. The hybrid architecture was operationalized with input perturbations and uncertainty estimation techniques, providing insight into how theoretical innovations in model design could be translated into real-world forecasting.

Fourth, the evaluation of the HPDEF-MUIP model validated its effectiveness in addressing the challenges identified at the beginning of this research. Using the real-world smart meter dataset, the model achieved superior forecasting performance with an R^2 of 0.95, RMSE of 1.71, and MAPE of 3.12%, surpassing individual base models and benchmark models from the literature. The model demonstrated enhanced accuracy under input perturbations, capturing both cyclic and peak demand patterns with high fidelity. These results confirm the practical viability of the HPDEF-MUIP model for deployment in the smart grid environment and underscore its value for demand-response scheduling, grid stability, and power cost optimization.

This thesis made many notable contributions to the body of knowledge. Theoretically, it advanced the understanding of the structural and functional limitations of forecasting under uncertain and noisy conditions. Methodically, it proposed and validated anHPDEF-MUIP framework that integrates uncertainty estimation, adversarial robustness, and hybrid model training into a unified architecture. Practically, it provided empirical evidence of enhanced forecast accuracy and resilience, offering energy sector stakeholders a reliable tool for real-world application. Collectively, these contributions bridged the gap between theoretical modeling and operating utility, thereby enriching both academic scholarship and applied power industry management practices.

This study, therefore, demonstrates that hybridization, combined with uncertainty-aware design, represents a powerful approach for advancing short-term Power Demand Forecasting. By aligning theoretical rigor with practical implementation, this research has developed a replicable, adaptable, and deployable forecasting model that supports the evolving demands of modern smart grid power systems.

5.4 Recommendations

The findings of this study indicate that hybrid frameworks, when combined with uncertainty estimation, provide superior accuracy and resilience under input perturbations compared to individual base models. Based on these insights, power utilities, grid operators, and independent system operators should adopt hybrid models, such as HPDEF-MUIP, for routine demand forecasting, as they provide both enhanced accuracy and reliability under dynamic input perturbations.

5.4.1 Recommendations for Objective 1: To assess the weaknesses of existing Power Demand Forecasting Models with Uncertainty Estimation under Input Perturbations

Future research should expand the comparative assessment of forecasting models by incorporating a broader taxonomy of uncertainty-aware techniques such as Bayesian Neural Networks, Deep Ensembles, and probabilistic Transformers. It is recommended that assessments include robustness benchmarks under diverse perturbation scenarios (e.g., adversarial noise, sensor failure, extreme weather shocks) to generate a standardized evaluation index for model reliability. Further systematic reviews and meta-analyses should also include multi-region studies to enhance generalizability.

5.4.2 Recommendations for Objective 2: To design a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations

Researchers and practitioners should consider refining hybrid model design by incorporating explainable AI (XAI) techniques, such as SHAP and LIME, to enhance interpretability and foster stakeholder trust. Future design work should explore modular architectures that enable the seamless replacement or upgrading of base models without requiring redesign of the entire framework. It is also recommended to incorporate model simplification strategies to preserve accuracy while reducing computational cost, making the hybrid design more scalable for low-resource utility environments.

5.4.3 Recommendations for Objective 3: To implement a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations using Real-World Power System Data

It is recommended that future implementation efforts adopt MLOps frameworks (e.g., MLflow, Docker, Airflow) to improve reproducibility, automated tracking, and deployment readiness across utility environments. Utility companies and researchers should increasingly adopt data-centric AI practices, such as continuous data quality assessment, missing-data imputation pipelines, and robust feature drift monitoring, to support sustained real-world implementation. Partnerships with power utilities are also encouraged to access richer datasets, including socio-economic and industrial activity indicators, to enhance model contextualization and predictive depth.

5.4.4 Recommendations for Objective 4: To evaluate the accuracy of a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations

Future evaluations should incorporate multi-horizon accuracy assessments (short-term, medium-term, and long-term forecasting) to validate model reliability across different operational planning windows. It is recommended that evaluation metrics go beyond RMSE, MAPE, R^2, and F1-score to include calibration error, prediction interval coverage probability (PICP), and model confidence index scores, thereby better quantifying uncertainty performance. Comparative studies against state-of-the-art deep learning architectures (e.g., Temporal Fusion Transformers, GNN-based load forecasting) are encouraged to position HPDEF-MUIP within the evolving landscape of AI-driven energy forecasting.

5.4.1 Policy Recommendations

The study highlights both the strengths of hybrid models and the persistent limitations of existing forecasting practices, pointing out policy gaps that need to be addressed. Based on the findings of this study, the following recommendations are made:

i. Current energy forecasting guidelines should be updated to mandate the inclusion of uncertainty estimation in forecasting practices to enhance decision-making credibility in a volatile demand context.

ii. Policy makers should incentivize innovation in hybrid forecasting approaches through finding frameworks and industry-research partnerships, therefore, continuous improvement to address emerging challenges.

iii. National and regional power agencies should implement training and certification programs to build technical expertise in hybrid machine learning forecasting, ensuring skilled adoption at operational levels.

5.4.2 Recommendations for Future Studies

This study also revealed several gaps that inform future research studies, such as;

i. Expanding the use of explainable Artificial Intelligence techniques will make hybrid forecasting models more transparent and trustworthy for policy makers, power grid operators, and other stakeholders.

ii. Future studies should explore extending the HPDEF-MUIP framework to jointly forecast power demand and renewable generation, especially solar and wind, under uncertainty. This smart grid will strengthen the purpose of the hybrid forecast in the environment.

iii. Scalability and deployment.Further empirical research should investigate the scalability of the HPDEF-MUIP model in various grid sizes, geographical and load contexts. This would validate the framework's generality and its practical deployment readiness.

iv. Future studies should consider integrating socio-economic, behavior, or policy variables into the forecasting pipeline to capture human demand variability, which remains underexplored in current models.

REFERENCES

A. Zameer, J. A. (2017). Intelligent and robust prediction of short term wind power using genetic programming based ensemble of neural networks. *Energy Convers. Manage., vol. 134*, pp. 361–372.

Abed, A. M. (2024). Accelerate demand forecasting by hybridizing CatBoost with the dingo optimization algorithm to support supply chain conceptual framework precisely. *Frontiers in Sustainability,* , 1388771.

Adedeji, P. A. (2020). Wind turbine power output very short-term forecast: A comparative study of data clustering techniques in a PSO-ANFIS model. *Journal of Cleaner Production, 254, 120135*.

Ahmad, A. R. (2025). Student Performance Prediction with Decision Tree Ensembles and Feature Selection Techniques. . *Journal of Information & Knowledge Management , 24(02),,* 2550016.

Ahmad, T. Z. (2020). A review on renewable energy and electricity requirement forecasting models for smart grid and buildings. *Sustainable Cities and Society,,* 55,102052.

Ahmad, T. Z. (2022). A review on renewable energy and electricity requirement forecasting models for smart grid and buildings. . *Sustainable Cities and Society.*, 55, 102052.

Ahmed, S. F. (2023). Deep learning modelling techniques: current progress, applications, advantages, and challenges. , . *Artificial Intelligence Review*, 56(11), 13521-13617.

Ahmed, S. L. (2024). Hybrid deep learning-ensemble frameworks for load forecasting under uncertainty. *Energy Informatics, 7(1),,* 45–59.

Al Khafaf, H. Z. (2024). MAPE-driven evaluation of forecasting models in renewable energy integration. *Applied Energy,362,,* 121890.

Al Mamun, M. A. (2023). A hybrid boosting approach for short-term load forecasting in smart grids. *IEEE Access*, 45632–45645.

Alam, M. S. (2023). Forecasting oil, coal, and natural gas prices in the pre-and post-COVID scenarios: contextual evidence from India using time series forecasting tools. . *Resources Policy, 81,,* 103342.

Alghamdi, M. A. (2024). Predicting Energy Consumption Using Stacked LSTM Snapshot Ensemble. *Big Data Mining and Analytics, 7(2),* , 247-270.

Al-Sharafi, M. A.-E. (2023). Determinants of cloud computing integration and its impact on sustainable performance in SMEs: . *An empirical investigation using the SEM-ANN approach. Heliyon,* , 9(5).

Althubiti, S. K. (2023). Artificial neural network for solving the nonlinear singular fractional differential equations. *Applied Mathematics in Science and Engineering,* , 31(1), 2187389.

Anand, A. &. (2020). Forecasting of electricity demand by hybrid ANN-PSO models. *n Deep learning and neural networks: Concepts, methodologies, tools, and applications. IGI Global.*, pp. 865-882.

Anum, H. H. (2024). *Backward neural network (BNN) based multilevel control for enhancing the quality of an islanded RES DC microgrid under variable communication networHeliyon,*, 10(12).

Anum, H. H. (2024). Backward neural network (BNN) based multilevel control for enhancing the quality of an islanded RES DC microgrid under variable communication network. . *Heliyon,* , 10(12).

Ashouri, M. H. (2022). Fast forecast reconciliation using linear models. . *Journal of Computational and Graphical Statistics,* , 31(1), 263-282.

Aurangzeb, K. A. (2021). A pyramid-CNN based deep learning model for power load forecasting of similar-profile energy customers based on clustering. *IEEE Access, 9,* , 14992-15003.

Aurangzeb, M. W. (2024). Enhancing cybersecurity in smart grids: Deep black box adversarial attacks and quantum voting ensemble models for blockchain privacy-preserving storage. *Energy Reports, 11,* , 2493-2515.

Balamurali, M. (2021). T-distributed stochastic neighbor embedding. In Encyclopedia of mathematical geosciences. *(Cham: Springer International Publishing.*, pp. 1-9). .

Ballesteros, M. C. (2020). Robust min–max optimal control design for systems with uncertain models: A neural dynamic programming approach. . *Neural Networks, 125*, 153-164.

Banik, R. &. (2024). Enhanced renewable power and load forecasting using RF-XGBoost stacked ensemble. . *Electrical Engineering,* , 1-21.

Bhamidipati, P. L. (2021). Local value capture from the energy transition: . *Insights from the Solar PV industry in Kenya*.

Bhotto, M. Z. (2021). Short-term demand prediction using an ensemble of linearly-constrained estimators. *IEEE Transactions on Power Systems, 36(4),* , 3163-3175.

Bilgiç, G. B. (2023). Recent advances in artificial neural network research for modeling hydrogen production processes. . *International Journal of Hydrogen Energy, 48(50),* ., 18947-18977.

Bo, H. N. (2020). Electric load forecasting use a novelty hybrid model on the basic of data preprocessing technique and multi-objective optimization algorithm. . *IEEE Access, 8,* , 13858-13874.

Brown, S. R. (2022). Bayesian Neural Networks for Geothermal Resource Assessment: Prediction with Uncertainty. *arXiv preprint arXiv:2209.15543.*

Brusaferri, A. M. (2022). Probabilistic electric load forecasting through Bayesian mixture density networks. . *Applied Energy, 309, 118341.*, 118341.

Bu, T. F. (2023). Optimal ANN-SNN conversion for high-accuracy and ultra-low-latency spiking neural networks. . *arXiv preprint arXiv:2303.04347*.

C. Li, S. L. (2018). Short-term wind power prediction based on data mining technology and improved support vector machine method. *Cleaner Prod., vol. 205,*, pp. 909–922,.

C. Wu, J. W. (2020). A novel hybrid system based on multi-objective optimization for wind speed forecasting. *Renew. Energy, vol. 146*, pp. 149–165.

Cao, W. L. (2023). Short-term district power load self-prediction based on improved XGBoost model. *Engineering Applications of Artificial Intelligence,* 126, 106826.

Cao, X. L. (2023). Heterogeneity and connection in the spatial–temporal evolution trend of China's energy consumption at provincial level. . *Applied Energy,.,* 336, 120842.

Cardo-Miota, J. T. (2024). Data-driven approach for day-ahead System Non-Synchronous Penetration forecasting: A comprehensive framework, model development and analysis. . *Applied Energy, 362,* , 123006.

Champahom, T. B. (2025). Deep Learning vs. Gradient Boosting: Optimizing Transport Energy Forecasts in Thailand Through LSTM and XGBoost. *Energies, 18(7),* , 1685.

Chandak, A. K. (2023). Gradient Perturbation-based Efficient Deep Ensembles. *In Proceedings of the 6th Joint International Conference on Data Science & Management of Data (10th ACM IKDD CODS and 28th COMAD),* pp. 28-36.

Chapagain, K. K. (2020). Short-term electricity demand forecasting: Impact analysis of temperature for Thailand. , . *Energies, 13(10)*, 2498.

Chen Zhang, Y. ,. (2022). On the limitations of R^2 in predictive performance evaluation. , . *Journal of Forecasting 41(6),* , 1032–1048.

Chen, Y. L. (2020). A novel multi-task deep learning framework for short-term load forecasting. . *Electric Power Systems Research*, 181, 106129.

Chen, Y. Z. (2022). (2022). Improving power load forecasting under uncertainty using noise-augmented deep learning models. . *Energy and AI, 10, 100203*.

Chen, Z. D. (2023). Targeted Attacks on Time Series Forecasting. . *In Pacific-Asia Conference on Knowledge Discovery and Data Mining* , (pp. 314-327).

Chenxi Ni., H. H. (2024). Light Gradient Boosting Machine (LightGBM) to forecasting data and assisting the defrosting strategy design of refrigerators. ,. *International Journal of Refrigeration*, 160, 182-196.

Cho, S., Shin, W., Kim, N., Jeong, J., & In, H. (2020). Priority determination to apply artificial intelligence technology in military intelligence areas. . *Electronics, 9(12)*, 2187.

Cibin, N. T. (2023.). Machine learning-based algorithms to estimate thermal dynamics of residential buildings with energy flexibility. *Journal of Building Engineering*, 65, 105683.

Cohen, J. P. (2020). Predicting covid-19 pneumonia severity on chest x-ray with deep learning. . *Cureus,* ., 12(7).

Conte, T. &. (2024). Comparative Analysis between Intelligent Machine Committees and Hybrid Deep Learning with Genetic Algorithms in Energy Sector Forecasting: A Case Study on Electricity Price and Wind Speed in the Brazilian Market. *Energies, 17(4),* , 829.

Cooper, C. B. (2024). A Tailored Approach: A model for literature searching in complex systematic reviews. *Journal of Information Science, 50(4),* , 1030-1062.

D. Karamichailidou, V. K. (2021). Wind turbine power curve modeling using radial basis function neural networks and tabu search,. *Renew. Energy, vol. 163*, pp. 2137–2152.

Dai, Y. Y. (2022). Forecasting power load: A hybrid forecasting method with intelligent data processing and optimized artificial intelligence. *echnological Forecasting and Social Change*, 182, 121858.

Dampfhoffer, M. M. (2023). Backpropagation-based learning techniques for deep spiking neural networks: A survey. *IEEE Transactions on Neural Networks and Learning Systems.*

David, I. &. (2025). Using Machine Learning for Systematic Literature Review Case in Point: Agile Software Development. Wiley Interdisciplinary Reviews. *Data Mining and Knowledge Discovery, 15(1),* , 1569.

Devadas, R. M. (2023). Integrating Dropout and Kullback-Leibler Regularization in Bayesian Neural Networks for improved uncertainty estimation in Regression. *MethodsX, 12,*, 102659.

Du, Z. (2024). Research on Electric Power Load Forecasting Based on SO-CatBoost. . *In 2024 International Conference on Artificial Intelligence and Power Systems (AIPS) IEEE.*, pp. 416-420.

Dube, R. G. (2023). Learning for Interval Prediction of Electricity Demand: . *A Cluster-based Bootstrapping Approach. arXiv preprint arXiv*, :2309.01336.

Duchi, J., Jordan, M., & Wainwright, M. (2018). Minimax optimal procedures for locally private estimation. *J. Am. Stat. Assoc*, 113, 182–201.

Eberhard, A. G. (2018). Kenya's lessons from two decades of experience with independent power producers. *Utilities Policy, .*, 52, 37–49.

Elhamod, M. &. (2023). Neuro-Visualizer: An Auto-encoder-based Loss Landscape Visualization Method.1. *arXiv preprint arXiv:2309.1460*.

EPRA. (2023). Energy & Petroleum Statistics Report for The Financial Year Ended 30th June 2023. *Retrieved from https://www.epra.go.ke/energy-petroleum-statistics-report-for-thefinancial-year-ended-30th-june-2023*.

Esnaola-Gonzalez, I. (2021). An ontology-based approach for making machine learning systems accountable. Semant. . *Web J.*

Eugene, E. A. (2023). Learning and optimization under epistemic uncertainty with Bayesian hybrid models. . *Computers & Chemical Engineering,* , 179, 108430.

Fan, G. F. (2023). A new intelligent hybrid forecasting method for power load considering uncertainty. . *Knowledge-Based Systems,*, 280, 111034.

Fang, S. L. (2022). Learning to learn transferable attack. *In Proceedings of the AAAI Conference on Artificial Intelligence,* (Vol. 36, No. 1, pp. 571-579).

Faradonbeh, M. K. (2020). Optimism-based adaptive regulation of linear-quadratic systems. *IEEE Transactions on Automatic Control*, 66(4), 1802-1808.

Fong, R. C. (2017). Interpretable explanations of black boxes by meaningful perturbation. *In Proceedings of the IEEE international conference on computer vision*, 3429-3437.

Fort, S. H. (2019). Deep ensembles: A loss landscape perspective. *arXiv preprint arXiv:1912.02757*.

Fort, S. H. (2019). Deep ensembles: A loss landscape perspective.7. *arXiv preprint arXiv:1912.0275*.

Franchi, G. L. (2024). Make Me a BNN: A Simple Strategy for Estimating Bayesian Uncertainty from Pre-trained Models. . *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition).*, pp. 12194-12204.

French, G. A. (2019). Semi-supervised semantic segmentation needs strong, high-dimensional perturbations.

Frontiers, R. i. (2024). Ensemble methods in intelligent energy systems: A review and case study. *Frontiers in Energy Research, 12,*, 11345.

Frutiger J, M. C. (2016). *J Chem Eng Data*, :602-613.

G. Wang, R. J. (2020). A hybrid wind power forecasting approach based on Bayesian model averaging and ensemble learning. *Renew. Energy, vol. 145*, pp. 2426–2434.

Gal, Y. &. (2016). Dropout as a bayesian approximation: Representing model uncertainty in deep learning. *In international conference on machine learning*, pp. 1050-1059.

Gal, Y. &. (2016). Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In International conference on machine learning (pp. 1050-1059). (pp. 1050-1059).

Gallavotti, G. (2022). Perturbation theory. In Perturbation Theory:). . *Mathematics, Methods and Applications New York, NY: Springer US*, (pp. 1-14.

García-Alvarado, J. J.-G.-M.-C. (2025). Assessing the Potential Risk of Invasion of the Neophyte Pluchea ovalis (Pers.) DC.(Asteraceae) in the Canarian Archipelago Using an Ensemble of Species Distribution Modelling. . *Diversity, 17(3), 195.*, 195.

Garg, S. S. (2020). Semantics for robotic mapping, perception and interaction: . *A survey. Foundations and Trends® in Robotics,* , 8(1–2), 1-224.

Gaur, A. &. (2009). *Statistical Methods for practice and Research.* Muthura Road New Delhi 110 044: India : Response Books.

Ge, Q. G. (2020). Industrial power load forecasting method based on reinforcement learning and PSO-LSSVM. *IEEE transactions on cybernetics,* ., 52(2), 1112-1124.

Ghazal, T. M. (2022). Energy demand forecasting using fused machine learning approaches. . *Intelligent Automation & Soft Computing, 31(1).*, 539-553.

González-Díaz, R. R.-C. (2021). Predictive sequential research design to study complex social phenomena. *Entropy,* , 23(5), 627.

Gu, B. H. (2023). Short-term wind power forecasting and uncertainty analysis based on FCM–WOA–ELM–GMM. . *Energy Reports,*, 807-819.

Guo, C. P. (2017). Calibration of modern neural networks. *Proceedings of the 34th International Conference on Machine Learning*, 1321-1330.

Guo, F. F. (2022). Applications of random forest in multivariable response surface for short-term load forecasting. *International Journal of Electrical Power & Energy Systems*, 139, 108073.

Guo, F. M. (2024). A hybrid stacking model for enhanced short-term load forecasting. *Electronics, 13(14),* , 2719.

Gwak, M. K. (2022). Robust and explainable fault diagnosis with power-perturbation-based decision boundary analysis of deep learning models. *IEEE Transactions on Industrial Informatics.*

H. Liu, H.-Q. T.-F. (2015). *' Energy Convers. Manage., vol. 100,*, pp. 16–22,.

H. Yin, Z. D. (2017). An effective secondary decomposition approach for wind power forecasting using extreme learning machine trained by crisscross optimization. *Energy Convers. Manage., vol. 150,* pp. 108–121.

H.-F. Yang and Y.-P.-P. Chen. (2020). Representation learning with extreme learning machines and empirical mode decomposition for wind speed forecasting methods. *Artif. Intell., vol. 277, Dec. 2019, Art. no. 103176*.

Hajirahimi, Z. &. (2023). Hybridization of hybrid structures for time series forecasting: . *A review. Artificial Intelligence Review, 56(2),*, 1201-1261.

Harikrishnan, G. R. (2025). Advanced short-term load forecasting for residential demand response: An XGBoost-ANN ensemble approach. *Electric Power Systems Research,* , 242, 111476.

Hasan, M. A. (2024). A blending ensemble learning model for crude oil price forecasting. . *Annals of Operations Research,* , 1-31.

Hesamifard, E., Takabi, H., Ghasemi, M., & Wright, R. (2018). Privacy-preserving machine learning as a service. *Proc. Priv. Enhancing Technol*, 123–142.

Hirschfeld L, S. K. (2020). Uncertainty Quantification using neural networks for molecular property prediction. *J Chem Inf Model*, 3770-3780.

Hong, T. P. (2016). Global energy forecasting competition 2012. g, 3. *International Journal of Forecasting 896-913. doi: 10.1016/j.ijforecast.2016.01.001*, 2(3), .

Hong, T. P. (2020). Probabilistic energy forecasting: A review of state-of-the-art methodologies. . *International Journal of Forecasting*, 36(1), 3-17.

Hong, T. Y. (2021). Hybrid deep learning model for short-term electricity demand forecasting considering seasonal and temporal patterns. *Energies*, 1911.

Hosseini, S. M. (2020). Robust optimal energy management of a residential microgrid under uncertainties on demand and renewable power generation. . *IEEE Transactions on Automation Science and Engineering, 18(2).*, 618-637.

Hoyer, L. D. (2021). Three ways to improve semantic segmentation with self-supervised depth estimation. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11130-11140.

Huang, A. B. (2024). Developing a hybrid technique for energy demand forecasting based on optimized improved SVM by the boosted multi-verse optimizer: Investigation on affecting factors. . *Heliyon,* , 10(7).

Huang, H. (2020). Variational mean-field theory for training restricted Boltzmann machines with binary synapses. *Physical Review E, 102(3)*, 030301.

Huang, J. Z. (2020). Attention-diffusion-bilinear neural network for brain network analysis. . *IEEE transactions on medical imaging, .,* 39(7), 2541-2552.

Huang, L. Q. (2023). Normalization techniques in training dnns: Methodology, analysis and application. . *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Ibrahim, H. A. (2023). Sustainability of power generation for developing economies: A systematic review of power sources mix. *Energy Strategy Reviews*, 47, 101085.

Impram, S., Nese, S., & Oral, B. (2020). Challenges of renewable energy penetration on power system flexibility. *A survey. Energy Strategy*, 31, 100539.

Izmailov, P. V. (2021). What are Bayesian neural network posteriors really like? *In International conference on machine learning (pp. 4629-4640). PMLR*.

J. M. Lima, A. K. (2017). A meteorological-statistic model for short-term wind power forecasting,. *J. Control. Autom. Electr. Syst., vol. 28, no. 5*, pp. 679–691,.

J. Manero, J. B. (2019). Dust in the wind ,deep learning application to wind energy time series forecasting,. *Energies, vol. 12, no. 12,*, p. 2385.

Jain, V. V. (2023). esign of Experiment Approach to Assess Nine Bisphenols in Fruit Juice Using QuEChERS with Injector Port Silylation via GC–MS/MS. *Food Analytical Methods, 16(8),*, 1370-1380.

Jeon, J. J. (2022). Bayesian approaches for Quantifying Clinicians' Variability in Medical Image Quantification. *arXiv preprint arXiv:2207.01868.*

Jiang, P. L. (2020). Modeling of electricity demand forecast for power system. . *Neural Computing and Applications,*, 32(11), 6857-6875.

Kaiyan, W. K. (2022). Performance Comparison of Bayesian Deep Learning Model and Traditional Bayesian Neural Network in Short-Term PV Interval Prediction. . *Sustainability, 14(19),* , 12683.

Kang, L. L. (2023). Forecasting large collections of time series: feature-based methods. *In Forecasting with Artificial Intelligence Theory and Applications Cham: Springer Nature Switzerland.*, pp. 251-276).

Kaur, A. &. (2021). Robust Bayesian Neural Network (RBNN). *energies'* .

Khan, A. S. (2020). A survey of the recent architectures of deep convolutional neural networks. . *Artificial intelligence review,* , 53, 5455-5516.

Khan, M. A. (2022). Artificial intelligence enabled demand response: Prospects and challenges in smart grid environment. . *Ieee Access,* , 11, 1477-1505.

Khodayar, M. &. (2023). Deep neural networks in power systems: A review. . *Energies,*, 16(12), 4773.

Kilimci, Z., Akyuz, A., Uysal, M., Akyokus, S., Uysal, M., Atak Bulbul, B., & Ekmis, M. (2019). *n. Complexity* .

Kim, D. &. (2023). Evaluating error metrics in smart grid load forecasting. . *IEEE Access,*, 67312–67325.

Kiptoo, M. K. (2023). Optimal capacity and operational planning for renewable energy-based microgrid considering different demand-side management strategies. *Energies,* , 16(10), 4147.

Kumari, P. &. (2021). Extreme gradient boosting and deep neural network based ensemble learning approach to forecast hourly solar irradiance. . *Journal of Cleaner Production, 279,* , 123285.

Kumari, S. K. (2023). *Research methodology.* Darshan Publishers.

Kumatongo, B. &. (2021). Research paradigms and designs with their application in education. . *Journal of Lexicography and Terminology (Online ISSN 2664-0899,* rint ISSN 2517-9306)., 5(1), 16-32.

L.-L. Li, X. Z.-L. (2020). ‘Short-term wind power forecasting based on support vector machine with improved dragonfly algorithm. *J. Cleaner Prod.*

Larsen, B. W. (2021). How many degrees of freedom do we need to train deep networks: a loss landscape perspective. *arXiv preprint arXiv:2107.05802*.

Lee, D. Y. (2023). Data-driven models for predicting community changes in freshwater ecosystems. *A review. Ecological Informatics,* , 77, 102163.

Lee, S. Y. (2024 February). Prediction Accuracy and Adversarial Robustness of Error-Based Input Perturbation Learning. *In 2024 International Conference on Artificial Intelligence in Information and Communication (ICAIIC). IEEE.*, (pp. 265-269).

Lee, S. Y. (2024). Prediction Accuracy and Adversarial Robustness of Error-Based Input Perturbation Learning. *In 2024 International Conference on Artificial Intelligence in Information and Communication (ICAIIC). IEEE.*, (pp. 265-269).

Li et al.(2020)Li, W. L. (2020). A comprehensive study on visual explanations for spatio-temporal networks. *arXiv preprint arXiv*.

Li, W. H. (2020). A comprehensive study on visual expectations for spatial-Temporal networks. *, arXiv preprint:2005.00375*.

Li, X. W. (2023). Uncertainty-Aware Scene Graph Generation. *Pattern Recognition Letters,*, 30-37.

Liang, J. Z. (2023). Enhanced short-term load forecasting using an XGBoost–LSTM hybrid model. *Sustainable Energy, Grids and Networks https://doi.org/ 10.1016 /j.segan.2023.100919,* 31, 100919.

Lin, W., Wang, P., Chao, K., Lin, H., Yang, Z., & Lai, Y. (2021). Wind Power Forecasting with Deep Learning Networks: Time-Series Forecasting. *Appl. Sci. 2021,* , 11, 10335.

Lipu, M. H. (2021). Artificial intelligence based hybrid forecasting approaches for wind power generation: Progress, challenges and prospects. *EEE Access, 9,*, 102460-102489.

Liu, H. J. (2019). Universal adversarial perturbation via prior driven uncertainty approximation. *In Proceedings of the IEEE/CVF International Conference on Computer Vision,* pp. 2941-2949.

Liu, J. (2019). Short-term load forecasting using long short-term memory neural networks. . *IEEE Transactions on Industrial Informatics,*, 15(5), 2982-2990.

Liu, M. L. (2023). A New Hybrid Model Based on SCINet and LSTM for Short-Term Power Load Forecasting. . *Energies,* ., 17(1), 95.

Liu, W. D. (2020). A semi-supervised tri-catboost method for driving style recognitio. *Symmetry,*, 12(3), 336.

Liu, X. &. (2023). Evaluating CatBoost and LightGBM for electricity consumption prediction in smart cities. *Journal of Cleaner Production, 417, 137922.*, 137922.

Liu, X. T. (2021). Profiling pareto front with multi-objective stein variational gradient descent. *Advances in Neural Information Processing Systems,*, 34, 14721-14733.

Lu, Q. &. (2022). Demand response strategy of game between power supply and power consumption under multi-type user mode. *International Journal of Electrical Power & Energy Systems,*, 134, 107348.

Lu, Y. a. (2023). A load forecasting model based on support vector regression with whale optimization algorithm. *Multimedia Tools and Applications 82*, 9939-9959. Lu, Y. K. (n.d.).

Luo, F. R. (2020). Personalized residential energy usage recommendation system based on load monitoring and collaborative filtering. *EEE transactions on industrial informatics.*, 17(2), 1253-1262.

M. Liu, Z. C. (2020). 'Short-term wind speed forecasting based on the jaya-SVM model,'. *Int. J. Electr Power Energy Syst., vol. 121*, Art. no. 106056.

M. Qolipour, A. M.-M. (2019). 'Prediction of wind speed using a new grey-extreme learning machine. *: A case study,'' Energy Environ., vol. 30, no. 1,*, pp. 44–62.

Mahajan, A. D. (2024). Bayesian-Neural-Network-Based Approach for Probabilistic Prediction of Building-Energy Demands. . *Sustainability, 16(22),* , 1-22.

Mahajan, A. D. (2024). Bayesian-Neural-Network-Based Approach for Probabilistic Prediction of Building-Energy Demands. *Sustainability,*, 16(22), 1-22.

Malinin, A. &. (2020). Ensemble distribution distillation. . *In Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34, No. 05, pp. 9033-9040.

Manwell, J. F. (2023). Modelling energy storage, electric vehicles and power-to-X in a large scale hybrid offshore wind power system. *In 7th Offshore Energy & Storage Symposium (OSES 2023)). IET.*, (Vol. 2023, pp. 170-177.

Marzbani, F. O. (2023). Electric vehicle energy demand prediction techniques: . *An in-depth and critical systematic review. IEEE Access, ,* , 11,96242-96255.

Mathew, A. P. (2021). "Deep learning techniques: an overview." Advanced Machine Learning Technologies and Applications:. *Proceedings of AMLTA 2020* , 599-608.

Mehmood, G. A. (2021). 'A review and taxonomy of wind and solar energy forecasting methods based on deep learning. *Energy AI, vol. 4,* , Art. no. 100060.

Meydani, A. M. (2023). Implementation of the internet of things technology in the smart power grid. *In 2023 10th Iranian Conference on Renewable Energy & Distributed Generation (ICREDG). IEEE.*, (pp. 1-8).

Michalakopoulos, M. K. (2025). A LightGBM-based approach for forecasting electricity market prices under data sparsity. *IEEE Transactions on Power Systems, 40(1),*, 204–214.

Mienye, I. D. (2022). A survey of ensemble learning: Concepts, algorithms, applications, and prospects. *IEEE Access,* ., 10, 99129-99149.

Mir, A. A. (2020). A review of electricity demand forecasting in low and middle income countries: . *The demand determinants and horizons. Sustainability,* , 12(15), 5931.

Mohajan, H. K. (2020). Quantitative research: A successful investigation in natural and social sciences. *Journal of Economic Development, Environment and People,* , 50-79.

Mondejar, M. E.-M.-S. (2021). Digitalization to achieve sustainable development goals: Steps towards a Smart Green Planet. . *Science of The Total Environment,* , 794, 148539. https://doi.org/10.1016/j.scitotenv.2021.148539.

Mosavi, A. &. (2019). Energy consumption prediction using machine learning. *A review*.

Mukilan, P. B. (2023). Integrated solar PV and piezoelectric based torched fly ash tiles for smart building applications with machine learning forecasting. *Solar Energy, 258,,* 404-417.

Mukilan, P. B. (2023). Integrated solar PV and piezoelectric based torched fly ash tiles for smart building applications with machine learning forecasting. . *Solar Energy,,* 258, 404-417.

Munir, S. P. (2024). Energy Consumption Prediction Based on LightGBM empowered with eXplainable Artificial Intelligence. *EEE Access.*

N. Li, F. H. (2019). Wind power prediction based on extreme learning machine with kernel mean p-Power error loss,. *learning machine with kernel mean p-Power error loss, no. 4,,* , p. 673,.

Nadeem, A., & Arshad, N. (2019). Pakistan residential electricity consumption dataset. *In e-Energy 2019—10th ACM International*.

Najjar, R. (2023). Redefining radiology: a review of artificial intelligence integration in medical imaging. *Diagnostics*, 13(17), 2760.

Naqash, A. G. (2022). Load forecasting techniques for power system: Research challenges and survey. . *IEEE Access, 10,* , 71054-71090.

Nti, I. K.-B. (2023). A predictive analytics model for crop suitability and productivity with tree-based ensemble learning. . *Decision Analytics Journal,* , 8, 100311.

Oblak, L. v.-C. (2021). A systematic review of biological, social and environmental factors associated with epigenetic clock acceleration. *Ageing research reviews,,,* 69, 101348.

Onder, I. &. (2022). Time series analysis. . *Applied Data Science in Tourism*, 467.

Oppong Ma, J. A.-O.-S.-P. (2021). Demand and supply-side determinants of electric power consumption and representative roadmaps to 100% renewable systems. *Journal of Cleaner Production,,* 299, 126832.

Osband, I. W. (2022). Evaluating high-order predictive distributions in deep learning. *In Uncertainty in Artificial Intelligence PMLR*, (pp. 1552-1560).

Ouyang, F. &. (2021). Artificial intelligence in education. *The three paradigms. Computers and Education: Artificial Intelligence*, 2, 100020.

Ovadia, Y. F. (2019). Can you trust your model's uncertainty? evaluating predictive uncertainty under dataset shift. *Advances in neural information processing systems*, 32.

Ozigis, I. I. (2021). Performance evaluation of Kainji hydro-electric power plant using artificial neural networks and multiple linear regression. *International Journal of Energy and Water Resources,*, 1-11.

P. A. Adedeji, S. A. (2020). Wind turbine power output very short-term forecast: A comparative study of data clustering techniques in a PSO-ANFIS model. *J. Cleaner Prod.vol. 254, May 2020, Art. no. 120135.*

P. Scarabaggio, M. L. (2020). 'Analyzing the effects of COVID-19 pandemic on the energy demand: The case of north ern Italy,''. *in Proc. AEIT Int. Annu. Conf. (AEIT), Sep. 2020, pp. 1–6,* pp. 1–6.

Panapakidis, I. P. (2022). Electric power demand forecasting using ensemble learning techniques.*Energy, 239, 122242. https://doi.org/10.1016/j.energy.2021.*, 122 242.

Patkowski, K. (2020). Recent developments in symmetry-adapted perturbation theory. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 10(3), e1452.

Paul, J. &. (2022). Meta-analysis and traditional systematic literature reviews—What, why, when, where, and how?. *Psychology & Marketing, 39(6),* , 1099-1115.

Peng, T. G. (2024). Railway cold chain freight demand forecasting with graph neural networks: A novel GraphARMA-GRU model. *Expert Systems with Applications, 255,* 124693.

Petsiuk, V. D. (2018). Rise: Randomized input sampling for explanation of black-box models. *arXiv preprint arXiv:1806.07421.*

Pokou, F. S. (2024). Hybridization of ARIMA with learning models for forecasting of stock market time series. . *Computational Economics, 63(4),* , 1349-1399.

Porcello, J. C. (2022). A Wisdom of Crowds approach to Supervised Learning Ensembles using FPGAs. . *In 2022 IEEE Aerospace Conference (AERO). IEEE.*, 1-8.

Q. Huang, Y. C. (2017). "Forecasting of the electric vehicles' charging amount of electricity based on curves clustering", . *Proc. 13th Int. Conf. Natural Comput. Fuzzy Syst. Knowl. Discovery (ICNC-FSKD),* , pp. 2424-2428.

Qi, Y. Z. (2020). Deep learning models with uncertainty estimation for short-term load forecasting. *IEEE Transactions on Smart Grid , 11(2),* 1412-1421.

Qinghe, Z. W. (2022). Optimised extreme gradient boosting model for short term electric load demand forecasting of regional grid system. . *Scientific Reports, 12(1),* , 19282.

Qiu, T. W. (2022). Potential assessment of photovoltaic power generation in China. *Renewable and Sustainable Energy Reviews,*, 154, 111900.

Radhika, N. S. (2024). A Comparative Analysis of Machine Learning Techniques for Predicting the Wear Rate of Ceramic Coated Steel. *IEEE Access.*

Rajawat, D. L. (2023). Object Detection in Images and Videos Using OpenCV: A Comparative Study of Deep Learning and Traditional Computer Vision Techniques. *In 2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON) . IEEE.*, (Vol. 10, pp. 141-146).

Ramamurthy, K. N. (2020). Crowd counting with decomposed uncertainty. *In Proceedings of the AAAI conference on artificial intelligence*, pp. 11799-11806.

Rashid, M. M. (2022). Adversarial training for deep learning-based cyberattack detection in IoT-based smart city applications. *Computers & Security, ,* 120, 102783.

Ren, J. L. (2022). Out-of-Distribution Detection and Selective Generation for Conditional Language Models. *arXiv preprint arXiv:2209.15558.*

Rice, L. W. (2020). Overfitting in adversarially robust deep learning. . *In International conference on machine learning*, (pp. 8093-8104). PMLR.

Richardson, M. L. (2021). MR protocol optimization with deep learning: a proof of concept.*Current problems in diagnostic radiology, 50(2),* , 168-174.

Rieder, A. C. (2024). A critical realist approach to agent-based modeling: Unlocking prediction in non-positivist paradigms.*Journal of Information Technology,* , 0268396 2241280657.

Rongquan. Z, G. J. (2022). Hybrid Model Based on Isolated Forest and WT-Catboost for Wind Power Forecasting. *In 2022 2nd International Conference on Computation, Communication and Engineering (ICCCE) IEEE.*, pp. 102-105.

Ryu, S. K. (2019). Uncertainty quantification of molecular property prediction with Bayesian neural networks. *arXiv preprint arXiv:1903.08375.*

S.Chowdhury, R. M. (2020). Energy consumption prediction using light gradient boosting machine model. *In International conference on emerging trends and advances in electrical engineering and renewable energy Singapore Springer Nature Singapore*, (pp. 413-422).

Sabri, M. S. (2024). Machine Learning-Aided Monte Carlo Simulation and Subset Simulation. *Transportation Research Record, 03611981241248166.*

Salehin, I. &. (2023). A review on dropout regularization approaches for deep neural networks within the scholarly domain. . *Electronics,*, 12(14), 3106.

Sareen, K. P. (2023). An imputation and decomposition algorithms based integrated approach with bidirectional LSTM neural network for wind speed prediction. *Energy, 278,*

Sarkar, A. R. (2020). Neural Data Augmentation Techniques for Time Series Data and its Benefit. *In 2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA). IEEE.*, (pp. 107-114).

Sarmas, E. S. (2023). Short-term photovoltaic power forecasting using meta-learning and numerical weather prediction independent Long Short-Term Memory models. . *Renewable Energy,* , 216, 118997.

Scalia G, G. C. (2020). *J Chem Inf Model*, 2697-2717.

Seeram, E. (2019). An overview of correlational research. . *Radiologic technology, 91(2),* 176-179.

Shabbir, N. H. (2024). Short-term residental DC load forecasting using extreme gradient boost (XgBoost) algorithm. In 2024 IEEE 18th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG) . *IEEE.*, (pp. 1-6).

Shah, I. J. (2022). Functional Data Approach for Short-Term Electricity Demand Forecasting.*Mathematical problems in engineering, 2022(1), .*, 6709779.

Sheikh, M. R. (2024). Review of Recent Developments in Hydrologic Forecast Merging Techniques. . *Water, 16(2),* , 301.

Shi, G. L. (2022). A controller synthesis method to achieve independent reference tracking performance and disturbance rejection performance. . *ACS omega, .*, 7(18), 16164-16186.

Siddiqui, R. A. (2022). A hybrid demand forecasting model for greater forecasting accuracy: the case of the pharmaceutical industry. . *In Supply Chain Forum: An International Journal (Vol. 23, No. 2,). Taylor & Francis.*, pp. 124-134.

Singh, A. &. (2024). Real-time grid anomaly detection using ensemble machine learning A LightGBM perspective. *IEEE Transactions on Smart Grid, 15(3),*, 2341–2352.

Singh, A. &. (2024). Real-time grid anomaly detection using ensemble machine learning A LightGBM perspective. *IEEE Transactions on Smart Grid,*, 2341–2352.

Singh, N. K. (2024). LightGBM-, SHAP-, and Correlation-Matrix-Heatmap-Based Approaches for Analyzing Household Energy Data: Towards Electricity Self-Sufficient Houses. *Energies, 17(17), 4518.*, 4518.

Smith, A. B. (2021). Improved power demand forecasts through spatiotemporal modeling of uncertainty. . *Energy Economics,*, 100, 105591.

Smith, J. e. (2020). Incorporating Uncertainty Estimation in Power Demand Forecasting. *IEEE Transactions on Power Systems, 35(2),* , 1423-1432.

Smith, J. J. (2019). Improved spatial resolution of precipitation ground based observations. *A journal of Geophysical Research*.

Solari, A. &. (2022). Multi split conformal prediction. *Statistics & Probability Letters.*, 184, 109395.

Spiliotis, E. M. (2022). Comparison of statistical and machine learning methods for daily SKU demand forecasting. *Operational Research, 22(3),* , 3037-3061.

Sreekumar, A. V. (2024). Electric vehicle charging station demand prediction model deploying data slotting. . *Results in Engineering, 24,* , 103095.

Strezoski A., L. K.-F. (2021). Guest Editorial: Special Issue on recent advancements in electric power system. *International Journal of Electrical Power & Energy Systems*, 129, 106597.

Sun, L., Qian, J., & Chen, X. L.-f. (2020). Practical private aggregation in federated learning with local differential privacy. *arXiv*, 15789.

Sun, Y. L. (2022). Percentage error-based evaluation in renewable energy forecasting. Renewable Energy,. *Renewable Energy, 192,*, 467–480.

T.Zhang., Z. X. (2023). Long-term energy and peak power demand forecasting based on sequential-XGBoost. . *IEEE Transactions on Power Systems,*, 39(2), 3088-3104.

Taieb, S. B. (2021). Hierarchical probabilistic forecasting of electricity demand with smart meter data. . *Journal of the American Statistical Association, 116(533),*, 27-43.

Takase M., K. R. (2021). A comprehensive review of energy scenario and sustainable energy in Kenya. *Fuel Communications*, 7, 100015.

Tama, B. A. (2021). Ensemble learning for intrusion detection systems: A systematic mapping study and cross-benchmark evaluation. *Computer Science Review, .*, 39, 100357.

Tejpal, S. S.-I.-S. (2019). owards personalised molecular feedback for weight loss. . *BMC obesity, 6,* , 1-15.

Tian &, Z. ,. (2021). A VMD and LSTM based hybrid model of load forecasting for power grid security. *IEEE Transactions on Industrial Informatics.*, 18(9), 6474-648.

Tingze, Z. T. (2023). Long-term energy and peak power demand forecasting based on sequential-XGBoost. . *EEE Transactions on Power Systems, 39(2), 3088-3104*, 3088-3104.

Tziolis, G. L.-L. (2024). Direct short-term net load forecasting in renewable integrated microgrids using machine learning: . *A comparative assessment. Sustainable Energy, Grids and Networks, .*, 37, 101256.

Tziolis, G. S.-L. (2023). Short-term electric net load forecasting for solar-integrated distribution systems based on Bayesian neural networks and statistical post-processing. *Energy*, *271,*, 127018.

Uzor, T. N. (2024). Emerging Technology in Human Kinetics and Sport: Advantages and Disadvantages of Using Hawk-Eye Technology. . *International Journal of Research and Innovation in Applied Science*, 9(8), 740-744.

V. N. Sewdien, R. P. (2020). 'Assessment of critical parameters for artificial neural networks based short-term wind generation forecasting. *Energy, vol. 161*, pp. 878–892.

Vieira, C. M. (2019). Using Computational Methods to Analyze Educational Data. *In 2019 IEEE Frontiers in Education Conference (FIE) IEEE*, (pp. 1-4).

Vinayakumar, R. A. (2019). Robust intelligent malware detection using deep learning. . *IEEE access,* , 7, 46717-46738.

Wang, B. Z. (2019). A survey on short-term load forecasting of electrical power systems: Models, methods, and future directions. . *Renewable and Sustainable Energy Review*, 101, 417-433.

Wang, F. Z. (2021). Gradient-guided dynamic efficient adversarial training. . *arXiv preprint arXiv:2103.03076*.

Wang, F. Z. (2022). Deep learning and its adversarial robustness: A brief introduction and its adversarial robustness. *In Handbook on computer learning and intelligence: Volume 2: Deep learning, intelligent control and evolutionary computation*, (pp. 547-584).

Wang, J. L. (2022). Convolutional spectral kernel learning with generalization guarantees. *Artificial Intelligence,* 103803.

Wang, N., Xiao, X., Yang, Y., Zhao, J., Hui, S., Shin, H., . . . Yu, G. (2019). Collecting and analyzing multidimensional data with local differential privacy. *In Proceedings of the 2019 IEEE 35th International Conference on Data Engineering (ICDE), Macau, China*, 638–649.

Wang, R. H. (2020). Renewable energy microgrids: Economic evaluation and decision making for government policies to contribute to affordable and clean energy. . *Applied Energy, .,* 274, 115287.

Wang, R. L. (2020). A novel improved model for building energy consumption prediction based on model integration. . *Applied Energy, .,* 262, 114561.

Wang, T., Zhang, X., Feng, J., & Yang, X. (2020). A Comprehensive Survey on Local Differential Privacy toward Data Statistics and Analysis. *Sensors*, 7030.

Wang, X. M. (2022). A Study on China coal Price forecasting based on CEEMDAN-GWO-CatBoost hybrid forecasting model under Carbon Neutral Target. *Frontiers in Environmental Science,,* 10, 1014021.

Wang, X. W. (2023). A hybrid framework for multivariate long-sequence time series forecasting. *Applied Intelligence, 53(11),* , 13549-13568.

Wang, Y. &. (2023). Comparative evaluation of tree-based models for high-frequency electricity demand forecasting. *Electric Power Systems Research, 108982. https://doi.org/10.1016/j.epsr.2023.108982,* 218.

Wang, Y. &. (2024). Reducing extreme forecast errors in load prediction using hybrid error-compensation techniques. . *Energy Reports, 10,,* 1122–1135.

Wang, Y. Z. (2023). Enhancing time series forecasting with Random Forest. *A new approach. Applied Intelligence, 53(1), https://doi.org/10.1007/s10489-022-03660-0*, 78-92.

Wang, Z. H. (2023). Secure access method of power internet of things based on zero trust architecture. *In International Conference on Swarm Intelligence Cham: Springer Nature Switzerland.,* (pp. 386-399).

Waswa, L. K. (2024). Reviewing the overcapacity claim: Insights from Kenya's electricity sector. *The Electricity Journal, .,* 37(2), 107374.

Watson, J. H.-S. (2021). Predicting preservice teachers' intention to use technology-enabled learning. 7. *Computers & Education,* , 168, 104207.

Wellens, A. P. (2023). When and how to use global forecasting methods on heterogeneous datasets. . *Available at SSRN 4629272*.

Wen, L. Z. (2020). Load demand forecasting of residential buildings using a deep learning model. . *Electric Power Systems Research,* , 179, 106073.

Wenzel, F. R. (2020). How good is the bayes posterior in deep neural networks really? *arXiv preprint arXiv:2002.02405.*

Weyns, D. S. (2021). Towards better adaptive systems by combining mape, control theory, and machine learning. *In 2021 International Symposium on Software Engineering for Adaptive and Self-Managing Systems (SEAMS),* (pp. 217-223). IEEE.

Wu, S. Z. (2023). Research On The Daily Electricity Forecast Model Based On LightGBM. . *In Journal of Physics: Conference Series . IOP Publishing,* Vol. 2477, No. 1, p. 012036.

Wu, Z. H. (2023). Load forecasting using CatBoost with real-time data from smart grids. *Energy and AI, 13, 100195. https://doi.org/10.1016/j.egyai.2023.100195,* 100195.

Xiao, L. Z. (2022). A hybrid optimization algorithm for machine learning model training. *Expert Systems with Applications, 202,*, 117139.

Xu, J. (2020). Generate adversarial examples by nesterov-momentum iterative fast gradient sign method. . *In 2020 ieee 11th international conference on software engineering and service science (icsess) IEEE,* (pp. 244-249).

Xue, M. W. (2021). Research on load forecasting of charging station based on XGBoost and LSTM model. . *In Journal of Physics: Conference Series (,). IOP Publishing.*, Vol. 1757, No. 1 P. 012145.

Yang, J. K. (2023). Certified Robust Control under Adversarial Perturbations. *arXiv preprint arXiv:2302.02208*.

Yang, Q. Z. (2020). morphological fragmental perturbation pyramid for black-box model explanations,. *n 2020 25th International conference on pattern recognition (ICPR) IEEE.*, pp. 1376-1383.

Yang, X. &. (2021). A hybrid short-term load forecasting model based on catboost and lstm. . *In 2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP) (). IEEE.*, pp. 328-332.

Yang, Y. T. (2022). A hybrid robust system considering outliers for electric load series forecasting. . *Applied Intelligence*, , 1-23.

Yang, Y. X. (2024). A novel combined probabilistic load forecasting system integrating hybrid quantile regression and knee improved multi-objective optimization strategy. *Applied Energy,* , 356, 122341.

Yaseen, M. &. (2023). Quantification of deep neural network prediction uncertainties for VVUQ of machine learning models. . *Nuclear Science and Engineering, 197(5),* ., 947-966.

Yidong, R. W. (2022). "Usb:A unified semi-supervised learning benchmark for classification." Advances in Neural Information Processing Systems. *Advances in Neural Information Processing Systems 35*, 35 (2022): 3938-3961.

Yin, P. D. (2020). .A hybrid method for forecasting new product sales based on fuzzy clustering and deep learning. *Kybernetes, 49(12),* , 3099-3118.

Yousaf, A., Asif, R., Shakir, M., Rehman, A., & S. Adrees, M. (2021). An Improved Residential Electricity Load Forecasting Using a Machine-Learning-Based Feature Selection Approach and a Proposed Integration Strategy. *Sustainability,* 13, 6199.

Yu, C. S. (2020). Deep kernel learning approach to engine emissions modeling. . *Data-Centric Engineering,* , 1, e4.

Yu, Z. W. (2022). Deep learning-based electricity demand forecasting with uncertainty estimation.

Z. Lin and X. Liu. (2020). *Energy, vol. 201 Art. no. 117693*.

Zakari, A. K. (2022). Energy efficiency and sustainable development goals (SDGs). *Energy,*, 239, 122365.

Zeiler, M. D. (2014). Visualizing and understanding convolutional networks. *In Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part I 13 Springer International Publishing.*, (pp. 818-833).

Zeng, H. Z. (2023). A hybrid modeling approach considering spatial heterogeneity and nonlinearity to discover the transition rules of urban cellular automata models. . *Environment and Planning B: Urban Analytics and City Science, 50(7),*, 1898-1915.

Zhang, C. Q. (2024). Simultaneous forecasting of wind speed for multiple stations based on attribute-augmented spatiotemporal graph convolutional network and tree-structured parzen estimator. *Energy*, 295, 131058.

Zhang, J. &. (2024). Hybrid CatBoost–XGBoost model optimized with arithmetic optimization algorithm for short-term load forecasting. *Energy Reports, 10, https://doi.org/10.1016/j.egyr.2024.03.074*, 2253–2264.

Zhang, J. &. (2024). Hybrid CatBoost–XGBoost model optimized with arithmetic optimization algorithm for short-term load forecasting. *Energy Reports,*, 2253–2264.

Zhang, J. S. (2020). Data uncertainty handling in load forecasting models: . *review. IEEE Transactions on Smart Grid*, , A11(4), 3293–3302. https://doi.org/10 .1109/TSG. 2020.2967492.

Zhang, J. Z. (2022). Diffusion kernel attention network for brain disorder classification. *IEEE Transactions on Medical Imaging*, , 41(10), 2814-2827.

Zhang, L. &. (2024). A hybrid XGBoost–BiGRU model for ultra-short-term load forecasting. *Energy Reports, 10,* , 1482–1495.

Zhang, W. (. (2022). Short-term load forecasting of power model based on CS-Catboost algorithm. I. *In 2022 IEEE 10th Joint International Information Technology and Artificial Intelligence Conference (ITAIC). EEE.*, Vol. 10, pp. 2295-2299.

Zhang, X. Z. (2021). Adversarial perturbation defense on deep neural networks. *ACM Computing Surveys (CSUR)*, 54(8), 1-36.

Zhang, Y. K. (2024). Wind power forecasting system with data enhancement and algorithm improvement. *Renewable and Sustainable Energy Reviews*.

Zhang, Y. R. (2023). Generalizing universal adversarial perturbations for deep neural networks. *Mach Learn 112,* , 1597–1626.

Zhao Gu, B. H. (2023). Short-term wind power forecasting and uncertainty analysis based on FCM–WOA–ELM–GMM. *Energy Reports, 9,*, 807-819.

Zhao, H. L. (2024). Learning quantum states and unitaries of bounded gate complexity. . *PRX Quantum, 5(4),* , 040306.

Zhao, J. H. (2017). A review on short-term load forecasting of electrical power system. . *Renewable and Sustainable Energy Reviews,*, 75, 835-854.

Zhifan, D. (2024). Research on Electric Power Load Forecasting Based on SO-CatBoost. *. In 2024 International Conference on Artificial Intelligence and Power Systems (AIPS). IEEE*, pp. 416-420.

Zhou, Z. H. (2022). Random Forests for time series forecasting. *A review. International Journal of Forecasting, 38(2), . https://doi.org/10.1016/j.ijforecast.2021.06.01*, 345-356.

Zhou, Z. H. (2025). Ensemble methods: foundations and algorithms. *CRC press.*

Zhou, Z. Y. (2025). Cascade Reservoir Outflow Simulation Based on Physics-Constrained Random Forest. *Water (20734441),* , 17(14).

Zhu, L. Z. (2022). Uncertainty and sensitivity analysis of cooling and heating loads for building energy planning. . *Journal of Building Engineering.*, 45, 103440.

Zulfiqar, M. G. (2024). Hyperparameter optimization of bayesian neural network using bayesian optimization and intelligent feature engineering for load forecasting. *Sensors, 22(12),*, 4446.

APPENDICES

Appendix 1: Code Listing

```
# Import required libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split, GridSearchCV, cross_val_score
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_squared_error, mean_absolute_error, r2_score
from sklearn.metrics import precision_score, recall_score, f1_score, confusion_matrix
from sklearn.ensemble import RandomForestRegressor
from xgboost import XGBRegressor
from lightgbm import LGBMRegressor
from catboost import CatBoostRegressor
from sklearn.linear_model import ElasticNet
from sklearn.svm import SVR
from sklearn.neighbors import KNeighborsRegressor
from itertools import product
import joblib
import warnings
from docx import Document
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, LSTM, Dropout
from tensorflow.keras.models import load_model
from sklearn.base import BaseEstimator
from datetime import datetime
from pathlib import Path  # For handling file paths
# 1. Data Loading and Initial Processing
def load_and_prepare_data(file_path):
    """Load and prepare the data for analysis."""
```

```python
print("Loading and preparing data...")
# 2. Data Preprocessing
def preprocess_data(df):
    """Preprocess the data including cleaning and feature engineering."""
print("\nPreprocessing data...")

    numerical_cols = [
        'MaxTemp', 'KV', 'A_32KV FDR', 'C_32KV FDR', 'D_32KV FDR', 'E_32KV FDR',
        'F_32KV FDR', 'G_11KV FDR', 'H_11KV FDR', 'I_11KV FDR', 'L_11KV FDR', 'MW'
    ]

# 3. Feature Engineering and Scaling
def engineer_features(df, scaler_path='scaler.pkl'):
    """Engineer features and scale the data using saved scaler."""
print("\nEngineering features...")
feature_cols = [
        'hour', 'day_of_week', 'month', 'is_weekend',
        'power_lag_1h', 'power_lag_24h', 'power_lag_168h',
        'MaxTemp', 'KV', 'A_32KV FDR', 'C_32KV FDR', 'D_32KV FDR', 'E_32KV FDR',
        'F_32KV FDR', 'G_11KV FDR', 'H_11KV FDR', 'I_11KV FDR', 'L_11KV FDR'
    ]
# 4. Model Training and Evaluation
def train_and_evaluate_model(model, X, y, model_name):
    """Train and evaluate a single model with cross-validation and hyperparameter tuning."""

# 6. Hybrid Model Class
class HybridPowerDemandForecaster:
    def __init__(self, models, weights=None):
        self.models = models
```

```
self.weights = weights if weights is not None else [1 / len(models)] * len(models)

    def fit(self, X, y):
        for model in self.models:
model.fit(X, y)
# 7. Uncertainty Estimation
def estimate_uncertainty(model, X, n_samples=100, perturbation_std=0.05):
    """Estimate uncertainty using input perturbations and save results to Excel."""
    # Ensure X is 2D
    if X.ndim == 1:
        X = X.reshape(1, -1)

    # Non-perturbed prediction
    if hasattr(model, 'model') and hasattr(model.model, 'predict'):
        if X.ndim == 2 and hasattr(model, 'input_shape') and model.input_shape != X.shape[1]:
X_rnn = X.reshape((X.shape[0], X.shape[1], 1))
non_perturbed_pred = model.model.predict(X_rnn).flatten()
        else:
non_perturbed_pred = model.model.predict(X).flatten()
    else:
non_perturbed_pred = model.predict(X)

    # Perturbed predictions
    predictions = []
    for _ in range(n_samples):
X_perturbed = X + np.random.normal(0, perturbation_std, X.shape)
        if hasattr(model, 'model') and hasattr(model.model, 'predict'):
            if X.ndim == 2 and hasattr(model, 'input_shape') and model.input_shape != X.shape[1]:
X_perturbed_rnn = X_perturbed.reshape((X_perturbed.shape[0], X_perturbed.shape[1], 1))
                pred = model.model.predict(X_perturbed_rnn)
            else:
```

```
                pred = model.model.predict(X_perturbed)
            else:
                pred = model.predict(X_perturbed)
predictions.append(pred)

    predictions = np.array(predictions)
mean_perturbed_pred = np.mean(predictions, axis=0)
    uncertainty = np.std(predictions, axis=0)
    # Individual Model Performance Summary
doc.add_heading('Individual Model Performance Summary', level=2)
summary_df = pd.DataFrame(individual_performance).T  # Transpose to have metrics as columns
    for index, row in summary_df.iterrows():
doc.add_paragraph(f"{index}: R² = {row['R2']:.4f}, RMSE = {row['RMSE']:.4f}, MAE = {row['MAE']:.4f}, MAPE = {row['MAPE']:.4f}")
    # Hybrid Model Performance
doc.add_heading('Hybrid Model Performance', level=2)
    for metric, value in hybrid_metrics.items():
        if isinstance(value, float):
doc.add_paragraph(f"{metric}: {value:.4f}")
        else:
doc.add_paragraph(f"{metric}: {value}")

    # 72-Hour Forecast Performance (if provided)
    if forecast_results is not None:
doc.add_heading('72-Hour Forecast Performance', level=2)
    # Forecast 72 hours ahead
    for i in range(72):
        # Prepare features for the next hour
X_current = last_data[feature_cols].values

        # Predict
        pred = hybrid_model.predict(X_current)
        _, _, uncertainty = estimate_uncertainty(hybrid_model, X_current)
```

```
# 10. Main Execution
def main():
    # Load and prepare data
    df = load_and_prepare_data("Data 2025.xlsx")
    df_processed = preprocess_data(df)
df_scaled, feature_cols = engineer_features(df_processed)

    # Split into train, test, and validation sets
    X = df_scaled[feature_cols]
    y = df_scaled['MW']

    # Split data into train (60%), validation (20%), test (20%)
    # Rank models based on RMSE
best_model_name = min(trained_models, key=lambda x: model_performance[x]['RMSE'])
print(f"\nBest model based on RMSE: {best_model_name}")
plot_feature_importance(trained_models[best_model_name], feature_cols)
    # Create hybrid model
    hybrid_model = HybridPowerDemandForecaster(
        [
model for name, model in top_models]
    )

    # Train hybrid model
hybrid_model.fit(X_train, y_train)
    hybrid_model.optimize_weights(X_val, y_val)

    # Make hybrid predictions
    hybrid_predictions = hybrid_model.predict(X_test)

    # Evaluate hybrid model
    y_test_values = y_test.values.flatten()
    hybrid_metrics = calculate_additional_metrics(y_test_values, hybrid_predictions)
```

```
    # Visualize predictions vs actual
plt.figure(figsize=(15, 6))
plt.plot(y_test.index.to_numpy(), y_test_values, label='Actual', alpha=0.7)
    # Perform uncertainty estimation
print("\nPerforming uncertainty estimation...")
uncertainty_results = {}
    # Save the hybrid model
joblib.dump(hybrid_model, 'HPDEF.joblib')
print("Hybrid model saved to 'HPDEF.joblib'.")

    # Evaluate on new data
    # Perform 72-hour ahead forecasting
forecast_df, forecast_metrics = forecast_72_hours("Data 2025.xlsx")
```

Appendix II: KUREC Clearance Letter

KABARAK UNIVERSITY RESEARCH ETHICS COMMITTEE

Private Bag - 20157
KABARAK, KENYA
Email: kurec@kabarak.ac.ke

Tel: 254-51-343234/5
Fax: 254-051-343529
www.kabarak.ac.ke

OUR REF: KABU01/KUREC/001/01/11/24

Date: 4th Nov, 2024

Francis Kimoryot Komen
Reg No: GDI/N/0263/01/21
Kabarak University,

Dear Francis,

RE: A Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations (HPDEF-MUIP).

This is to inform you that ***KUREC*** has reviewed and approved your above research proposal. Your application approval number is ***KUREC-011124.*** The approval period is ***4/11/2024 – 4/11/ 2025.***

This approval is subject to compliance with the following requirements:

i. All researchers shall obtain an introduction letter to NACOSTI from the relevant head of institutions (Institute of postgraduate, School dean or Directorate of research)
ii. The researcher shall further obtain a RESEARCH PERMIT from NACOSTI before commencement of data collection & submit a copy of the permit to **KUREC.**
iii. Only approved documents including (informed consents, study instruments, MTA Material Transfer Agreement) will be used
iv. All changes including (amendments, deviations, and violations) are submitted for review and approval by ***KUREC***:
v. Death and life-threatening problems and serious adverse events or unexpected adverse events whether related or unrelated to the study must be reported to ***KUREC*** within 72 hours of notification;
vi. Any changes, anticipated or otherwise that may increase the risk(s) or affected safety or welfare of study participants and others or affect the integrity of the research must be reported to ***KUREC*** within 72 hours;
vii. Clearance for export of biological specimens must be obtained from relevant institutions and submit a copy of the permit to KUREC;
viii. Submission of a request for renewal of approval at least 60 days prior to expiry of the approval period. Attach a comprehensive progress report to support the renewal and;
ix. Submission of an executive summary report within 90 days upon completion of the study to ***KUREC***

Sincerely,

Prof. Jackson Kitetu PhD.
KUREC-Chairman

Cc Vice Chancellor
DVC-Academic & Research
Registrar-Academic & Research
Director-Research Innovation & Outreach
Institute of Post Graduate Studies

KABARAK UNIVERSITY
INSTITUTIONAL RESEARCH ETHICS COMMITTEE
APPROVED
P. O. PRIVATE BAG - 20157, KABARAK

As members of Kabarak University family, we purpose at all times and in all places, to set apart in one's heart, Jesus as Lord.
(1 Peter 3:15)
Kabarak University is ISO 9001:2015 Certified

Apendix III: NACOSTI Reserch Permit

REPUBLIC OF KENYA

NATIONAL COMMISSION FOR SCIENCE, TECHNOLOGY & INNOVATION

Ref No: **760769**

Date of Issue: **15/November/2024**

RESEARCH LICENSE

This is to Certify that Mr.. FRANCIS K. KOMEN of Kabarak University, has been licensed to conduct research as per the provision of the Science, Technology and Innovation Act, 2013 (Rev.2014) in Nakuru on the topic: A HYBRID POWER DEMAND FORECASTING MODEL WITH UNCERTAINTY ESTIMATION UNDER INPUT PERTURBATIONS (HPDEF-MUIP). for the period ending : 15/November/2025.

License No: **NACOSTI/P/24/42109**

760769
Applicant Identification Number

Director General
NATIONAL COMMISSION FOR SCIENCE,TECHNOLOGY & INNOVATION

Verification QR Code

NOTE: This is a computer generated License. To verify the authenticity of this document, Scan the QR Code using QR scanner application.

See overleaf for conditions

Appendix IV: Evidence of Conference Participation

KABARAK UNIVERSITY

Certificate of Participation

Awarded to

FRANCIS K. KOMEN

For successfully participating in the 15th Annual Kabarak University International Research Conference held from 10th-11th July 2025 and presented a paper entitled ***"A Literature Review on Weakness of Existing Hybrid Power Demand Forecasting Models with Uncertainty Estimation under Input Perturbations."***

Conference Theme

Enhancing Environmental Resilience for Food Production and Energy Sustainability.

Prof. Peter Rugiri
Dean, School of Science, Engineering & Technology

Dr. Phillip Nyawere
Director - Research, Innovation and Outreach

Kabarak University Moral Code

As members of Kabarak University family, we purpose at all times and in all places, to set apart in one's heart, Jesus as Lord.

(1 Peter 3:15)

Kabarak University is ISO 9001:2015 Certified

Appendix V: List of Publications

Editon Consortium Journal of Engineering and Computer Science [ISSN: 2789 – 9985]

Issue no: 01 | **Vol no:** 02 | **September** 2025: 01-07

Towards Enhancing Robust Energy Forecasting: A Hybrid Model with Input Perturbation to Overcome Uncertainty in Power Demand Prediction

Francis Komen[1]
Moses Thiga[2]
Andrew Kipkebut[3]

Article History
Received: **2025.07.18**
Accepted: **2025.08.19**
Published: **2025.09.19**

(1.2.3) Kabarak University, Kenya.
Main author's email: franciskomen@gmail.com

Cite this article in APA
Komen, F., Thiga, M., & Kipkebut, A. (2025). Towards enhancing robust energy forecasting: A hybrid model with input perturbation to overcome uncertainty in power demand prediction. *Editon consortium journal of engineering and computer science, 2*(1), 01-07. https://doi.org/10.51317/ecjecs.v2i1.627

Abstract
The purpose of this article is to address the persistent challenge of reliable power demand forecasting in modern energy systems, where dynamic fluctuations and noisy signals often reduce model accuracy and credibility. Traditional forecasting methods, although widely applied, struggle to adapt to stochastic variations, limiting their usefulness for grid stability and long-term planning. This study proposed a Hybrid Power Demand Forecasting Model with Uncertainty Estimation under Input Perturbations (HPDEF-MUIP). The model combined three machine learning algorithms, Extreme Gradient Boosting, Categorical Boosting, and RandomForest, into a hybridised model designed to enhance robustness. Data preprocessing included Empirical Mode Decomposition for signal refinement, Kalman Filtering for noise reduction, and normalisation for balanced scaling. To simulate and evaluate resilience against noisy environments, adversarial perturbation strategies such as the Fast Gradient Sign Method were introduced. The model was trained and validated on a large smart meter dataset spanning 2022–2025, using ten-fold cross-validation and hyperparameter optimisation with Genetic Algorithms. Performance was assessed through standard accuracy metrics, including Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), the coefficient of determination (R^2) and Mean Absolute Percentage Error (MAPE). Findings showed that HPDEF-MUIP achieved an R^2 of 0.9539 and a MAPE of (3.12%), significantly outperforming baseline models. Under perturbed conditions, adversarial training reduced error variance by (17%), confirming improved resilience. The study concludes that hybrid model learning with uncertainty estimation offers a reliable and interpretable tool for supporting smart grid operations, demand-response planning, and sustainable energy management.

Key words: Energy forecasting, hybrid learning, input perturbation, smart grids, uncertainty estimation.

Editon OPEN ACCESS

Journal of Computer Science and Technology

[ISSN 2789-9985]
Volume: 03 Issue: 01 | September-2025

JCST

DESIGN OF A HYBRID POWER DEMAND FORECASTING MODEL WITH UNCERTAINTY QUANTIFICATION UNDER INPUT PERTURBATION

Authors

Francis Komen [1]; Moses Thiga [2]; Andrew Kipkebut [3]

Main author's email: franciskomen@gmail.com

(1.2.3) Kabarak University, Kenya.

Cite this article in APA

Komen, K., Thiga, M., & Kipkebut, A. (2025). Design of a hybrid power demand forecasting model with uncertainty quantification under input perturbation. *Journal of computer science and technology, 3*(1), 1-10. https://doi.org/10.51317/jcst.v3i1.814

OPEN ACCESS A publication of Editon Consortium Publishing (online)

Article history
Received: 19.07.2025
Accepted: 18.08.2025
Published: 19.09.2025

Scan this QR to read the paper online

Abstract

The study aimed to address the persistent challenge of accurate electricity demand forecasting in modern power systems, where reliability is often undermined by input perturbations such as weather fluctuations, consumer behaviour shifts, and sensor noise. Conventional Machine Learning (ML) and Deep Learning (DL) approaches, while effective in predictive accuracy, rarely incorporate uncertainty estimation, which reduces their robustness in real-world applications. To overcome this limitation, this study designed a hybrid power demand forecasting model with embedded uncertainty estimation. Seven base models, such as Artificial Neural Networks (ANN), Recurrent Neural Networks (RNN), Support Vector Regression (SVR), Extreme Gradient Boost (XGBoost), RandomForest, LightGBM, and CatBoost, were trained and evaluated using R^2, RMSE, MAE, and MAPE. The three best-performing models (XGBoost, CatBoost, and RandomForest) were fused through weighted averaging based on inverse error contributions. An uncertainty estimation mechanism was then integrated by quantifying variance under perturbed inputs, thereby generating confidence intervals around forecasts. Findings show that the hybrid model achieved high predictive accuracy (R^2 = 0.9539, with low error values: RMSE = 1.7128, MAE = 1.2270, MAPE = 3.1178) while also producing reliable uncertainty bounds. The significance of this study lies in demonstrating that hybrid modeling combined with uncertainty quantification provides both accurate and trustworthy forecasts, offering a practical decision-support tool for smart grid operators managing volatile energy demand.

Key terms: Energy forecasting, hybrid modeling , machine learning, smart grids, uncertainty quantification.