Short-Term Electricity Demand Forecasting and Resource Allocation Using Machine Learning and Linear Programming

Submitted in partial fulfillment of the requirements for the degree of

Master of Science
In
Business Statistics

by

Rosemol Jolly
24MBS0069

Under the guidance of

Dr. Amrit Das
School of Advanced Sciences
VIT - Vellore

Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

November, 2025

DECLARATION

I hereby declare that the thesis entitled "**Short-Term Electricity Demand Forecasting and Resource Allocation Using Machine Learning and Linear Programming**" submitted by me, for the award of the degree of *Master of Science in Business Statistics* to VIT is a record of bonafide work carried out by me under the supervision of **Amrit Das.**

I further declare that the work reported in this thesis has not been submitted and will not be submitted, either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university.

Place : Vellore
Date : 25/11/2025

Signature of the Candidate
Rosemol Jolly

CERTIFICATE

This is to certify that the thesis entitled "**Short-Term Electricity Demand Forecasting and Resource Allocation Using Machine Learning and Linear Programming**" submitted by **Rosemol Jolly (Reg. No.: 24MBS0069), School of Advanced Sciences**, VIT, for the award of the degree of *Master of Science in Business Statistics,* is a record of bonafide work carried out by her under my supervision during the period, **09.07.2025** to **14.11.2025**, as per the VIT code of academic and research ethics.

The contents of this report have not been submitted and will not be submitted either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university. The thesis fulfills the requirements and regulations of the University and in my opinion meets the necessary standards for submission.

Place: Vellore

Date: 25/11/2025

25/11/25

Signature of the Guide
Amrit Das
Department of Mathematics
SAS, VIT- Vellore

Dr. Mahamood Usman

Name & Signature of the Examiner

Dr. KHADAR BABU S K
Head, Department of Mathematics
SAS, VIT-Vellore

ACKNOWLEDGEMENT

With immense pleasure and deep sense of gratitude, I wish to express my sincere thanks to my guide **Amrit Das**, School of Advanced Sciences, VIT, Vellore without his motivation and continuous encouragement, this research would not have been successfully completed.

I am grateful to the Chancellor of VIT, Vellore, Dr. G. Viswanathan, the Vice Presidents and the Vice Chancellor for motivating me to carry out research in the Vellore Institute of Technology, Vellore and also for providing me with infrastructural facilities and many other resources needed for my research.

I express my sincere thanks to Dr. Karthikeyan K, Dean, School of Advanced Sciences, VIT, Vellore, for her kind words of support and encouragement. I like to acknowledge the support rendered by my classmates in several ways throughout my research work.

I wish to thank Dr. KHADAR BABU S K, Head of the Department of Mathematics, School of Advanced Sciences, VIT, Vellore for his encouragement and support.

I wish to extend my profound sense of gratitude to my parents and friends for all the support they made during my research and also providing me with encouragement whenever required.

Rose

Signature of student

Rosemol Jolly

ABSTRACT

The demand of electricity fluctuates; it is significant to forecast the electricity demand so that power distribution can be planned in a smooth manner particularly in times where power supply is limited. The provided project is devoted to predicting the short-term electricity demand in Germany, France, and Spain with the help of the Machine Learning tools and subsequently deciding how scarce electricity may be distributed among them. It had developed and compared three models- Random Forest, LightGBM and XGBoost with the latter giving the most credible predictions. Through these projected values, a shortage situation was developed by presuming that 80 percent of the total amount of electricity required is only available. These two allocations were carried out with the help of the linear programming method: proportional allocation, which allocates the electricity equally based on the demand, and priority-based allocation, which lays more weight to the chosen countries. The outcomes are clear on the effects of allocation decision on shortages and priority choices on the distribution. The combination of these forecasting and optimization techniques provides a viable solution into assisting decision-making in the times of electricity shortage.

Keywords: Electricity forecasting, Machine learning, XGBoost, Linear programming, Resource allocation, Optimization.

CONTENTS

Page No.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATION

AEMO	Australian Energy Market Operator
ANN	Artificial Neural Network
API	Application Programming Interface
CNN	Convolutional Neural Network
CSV	Comma Separated Values
GRU	Gated Recurrent Unit
HVDC	High Voltage Direct Current
KNN	K-Nearest Neighbour
LightGBM	Light Gradient Boosting Machine
LP	Linear Programming
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
MW	Megawatt
OSPD	Open Power System Data
RF	Random Forest
RMSE	Root Mean Squared Error
SVM	Support Vector Machine
SVR	Support Vector Regression
UTC	Coordinated Universal Time
XGBoost	Extreme Gradient Boosting

INTRODUCTION

1.1. OBJECTIVE

This project aims at researching, examining, and interpreting short term electricity demand trends of three countries, namely; Germany, France and Spain through real time consumption data. The project also strives to forecast the future electricity demand with various machine learning models and find the most successful one. According to these forecasted values, the research is also keen on exploring the allocation of limited electricity during shortages condition by use of a Linear Programming method.

1.2. MOTIVATION

The reason why this project was created is that I have always had curiosity about how real-world systems operate and what information can be used to make us make a better decision. Electricity is something we all use it almost on a daily basis and hardly do we ever consider the amount of electricity required and how it is handled in various locations. When I began reading about power outages and demand variation that occurred on the global level, I became interested in understanding the variability of electricity demand and its predictability. As data science offers powerful tools to be used in forecasting and optimization, I would prefer to take this project as a chance to learn these techniques practically. Another concept that I encountered as I dug further was the distributing the scarce electricity equitably when supply falls short and this topic became even more significant. This made me research on the demand trend of Germany, France, and Spain, create models to forecast their demands in future and use optimization to understand how electricity can be distributed in case of shortages.

1.3. LITERATURE SURVEY

There are many research works dedicated to electricity demand prediction and the efficient use of electrical energy. Some of the most relevant studies that supported my analysis are discussed as follows. The paper *"Allocation of Energy Resources in Electrical Networks Using Linear Programming: Optimization of Energy Allocation and Distribution in Electrical Networks Using Operations Research Techniques"* by Abdalsatar and Salman [1] presents a linear-programming-based framework for allocating energy resources in power grids while considering operational constraints, cost factors and varying demand levels. Their work demonstrates how mathematical optimization can enhance economic efficiency, reduce system losses and support sustainable grid operation.

The study by Clack, Xie and MacDonald [2], *"Linear Programming Techniques for Developing an Optimal Electrical System Including High-Voltage Direct-Current Transmission and Storage,"* focuses on designing an optimized power system that integrates HVDC transmission and energy storage using linear programming. Their findings show that such optimization models can reduce overall system cost while effectively managing large-scale power flows and renewable energy integration.

Fan and Hyndman [3], in their work *"Forecasting Electricity Demand in the Australian National Electricity Market,"* discuss a comprehensive load forecasting framework built using semi-parametric additive models. Their methodology incorporates weather variables, seasonal patterns and economic indicators and is currently used by the Australian Energy Market Operator. This highlights the importance of combining statistical modelling with practical system requirements to produce reliable forecasts.

Javidsharifi et al. [4] compare a baseline LSTM, a transfer-learning-based LSTM and XGBoost for hourly demand forecasting in Nordic cities. Their results show that a model trained on data-rich Copenhagen can be successfully fine-tuned for data-scarce Reykjavik, demonstrating the value of transfer learning when historical data is limited.

The work of Miyao and Nakamura [7], *"Time and Day-Based Peak Electricity Demand Forecasting,"* proposes a method that relies solely on time-based features rather than weather data. Among the models evaluated, LightGBM provides the highest accuracy and computational efficiency, making it suitable for peak-cut and peak-shift strategies aimed at decarbonization.

Pandiyan et al. [8], in their study *"Electrical Load Demand Forecasting Using Machine Learning Algorithm,"* compare LSTM, RNN, CNN-LSTM, SVM and ARIMA models. Their results indicate that LSTM performs best for capturing long-term temporal dependencies, while ARIMA performs poorly on nonlinear demand patterns, reinforcing the strength of deep learning models for modern load forecasting.

Saji, Prakash and Krishnan [9] develop forecasting models using ten years of KSEB data and evaluate several algorithms including Linear Regression, Decision Tree, Random Forest, SVR, KNN, XGBoost and ANN. Their findings show that Random Forest delivers the highest accuracy with the lowest error values, making it highly suitable for regional demand forecasting in Kerala.

Thakur et al. [10], in *"Predictive Analysis of Energy Consumption and Electricity Demand Using Machine Learning Techniques,"* use boosting models such as XGBoost, CatBoost and LightGBM for demand prediction. They also use autocorrelation plots and wavelet transforms to capture deeper time-series patterns, demonstrating that boosting-based models can effectively handle complex energy consumption behaviours.

Overall, these studies show that machine learning, deep learning, transfer learning and linear programming all play a significant role in improving electricity demand forecasting and energy allocation. However, there is still a need for models that are highly accurate, computationally efficient and adaptable to different regional and operational conditions.

PROJECT DESCRIPTION

The short-term electricity demand of the three countries in Europe namely Germany, France, and Spain are the subjects of this project and the countries are characterized by organized and documented patterns of power consumption. The central concept in the paper is to appreciate the dynamics of electricity demand and how the values can be forecasted with the assistance of the modern machine learning algorithms. The analysis is implemented in two sections, namely, Machine Learning Analysis and Optimization Analysis.

In the Machine Learning Analysis, we take a glance at the historical electricity consumption data of all three countries and examination of their demand patterns. Various machine learning algorithms like the Random Forest, LightGBM, and XGBoost are adopted to forecast the electricity demand in the future. These models are compared with each other to come up with the model that has the best performance and XGBoost is signed due to its better accuracy. It is based on this most appropriate model to predict the future hourly demand of electricity in each country.

The Optimization Analysis assumes a shortage situation using the predicted values, that only 80 percent of the necessary electricity is generated. In order to manage this state of affairs, a Linear Programming (LP) model is designed to distribute the scarce electricity in a coherent and significant manner. Two allocation schemes, which are proportional allocation and priority-based allocation, are used to see the effect of various rules on the distribution and shortages. The outcomes are very clear in revealing the effects of each technique on the ultimate allocation in Germany, France and Spain.

All in all, the project is the integration of forecasting and optimization to gain the full picture of electricity demand and resource allocation under the circumstances of shortage that illustrates the ways data-driven methods could be applied to the management of energy in a real-world situation.

PROJECT GOALS

i. Examining the electricity demand trends of France and Spain and Germany.
ii. The comparison and identification of performance of various machine learning models to predict short-term demand.
iii. Choosing the model that works the best (XGBoost) and forecasting hourly demands of electricity in the future.
iv. Sum total of the predicted demand in each country with regards to the model outputs.
v. It created an electricity shortage situation by reducing the supply to 80 percent of the total demand.
vi. Formulating a Linear Programming (LP) model in order to distribute scarce electricity between the three nations.
vii. Using the proportional and priority methods of allocation when examining the effect of allocation rules during shortages.
viii. Comparing the results of both allocation methods to study their effectiveness in real-world shortage situations.

TECHNICAL SPECIFICATIONS

The primary environment was Google Colab (Python 3) with the support of NumPy(numerical calculators) and pandas (Series, DataFrame) to process and figure out the electricity consumption dataset. The demand patterns, comparisons of the models and the allocation results were visualized using Matplotlib.pyplot and seaborn packages. Random Forest Regressor, Gradient Boosting, train-test-split, and evaluation metrics (RMSE, MAE, MAPE) were installed in the scikit-learn libraries to build and compare the various machine learning models. XGBoost and LightGBM packages were applied to implement the advanced graduate boosting algorithm to enhance the performance of the forecasting.to solve the optimization issue, pulP package was installed to model and solve the Linear Programming(LP) model of electricity allocation. All the workflow was run on Google Colab cloud-based hardware that provides support on the fast computation and big data processing.

SCHEDULE, TASKS AND MILESTONES

Table 1

S.NO	MONTH-WEEK	PLAN
1.	JULY- WEEK 1	Identification of the problem.
2.	JULY- WEEK 2, 3	Literature review on the decided problem.
3.	AUGUST- WEEK 1	Discussion on the aims, objectives and outcomes of the problem.
4.	AUGUST-WEEK 2-4	Formation of abstract.
5.	SEPTEMBER-WEEK 1,2	Collection of data.
6.	SEPTEMBER-WEEK 3,4	Methodology: Adaptation of the appropriate methods for the gathered data.
7.	OCTOBER- WEEK 1,2	Appropriate analysis, relevant discussion and valid conclusions.
8.	OCTOBER- WEEK 3	Feedback from guide.
9.	OCTOBER- WEEK 4	Final documentation and report writing.
10.	NOVEMBER - WEEK 1,2	Report review
11.	NOVEMBER – WEEK 3 (NOV 24)	Final review

METHODOLGY

6.1 OVERVIEW OF THE METHODOLOGY

The research approach followed in this project was to predict the electricity demand of Germany, France and Spain systematically with the help of machine learning models and optimally allocate the demand forecasted with the help of linear programming models. The whole workflow can be divided into five key steps, i.e., data collection, data preprocessing. feature engineering, model development and evaluation, and, last but not least, the design of an optimization model of energy allocation. The steps are well formulated to be accurate, consistent and meaningful in making a decision.

6.2 DATA COLLECTION

This study considered the Open Power System (OPSD) platform as the source of data since it is a publicly available and standardized source of electricity-related data. To this end, hourly electricity demand (in megawatts) in Germany (DE), France (FR) and Spain (ES) have been chosen to use in this project. The data extends to several years and has a consistent timestamp structure and thus can be used in a time series analysis and prognosis.

Once the countries of interest were chosen the raw CSV was extracted and loaded into python environment to continue with the processing. Preliminary examination was done to learn the column format, missing rows, time format and unit of measurement.

6.3 DATA PREPROCESSING

Electricity demand is a continuous time-series variable, hence data preprocessing was required to guarantee quality, consistency and modelling preparation. The step carried were as follows:

a. Handling Missing Values

   The nature of some timestamps had missing demand values because of missing recording, these were solved by the techniques of interpolation which approximate the lost entries by the observations around them. This step guarantees that the model gets a full-course to learn.

b. Timestamp Standardization

   The data had hourly timestamps that were in the UTC format. They were checked and normalized to allow the time index to have no anomalies like duplication or gaps. Any discrepancies were eliminated to ensure the order of events.

c. Selectivity of Countries to be filtered.

   Outliers Time-series subsets were designed separately between Germany, France and Spain. This enabled the data of each country to be modelled separately but with a similar pipeline.

d. Train-Test Split

   To measure objectively the model performance the data was split into training and testing sets. The latest part of the dataset was tested because time-series forecasting demands the maintenance of time-sequence.

6.4 FEATURE ENGINEERING

In order to enhance model learning ability a number of new features were formed based on original time series:

a. Lag Features

   Lag values like demand of past hour, previous day and previous week were built. The lag features aid the model to detect any temporal dependencies and recurring patterns of demand.

b. Rolling Statistics

   The features of rolling mean and rolling standard deviation were calculated to be able to meet the short term and medium-term trends in the consumption.

c. Time-Based Features

   The timestamp was broken down into hour, day of the week, month and the day being weekend or weekday. These features are often of good significance since electricity demand is usually affected by season and daily patterns of usage.

d. Normalization

   In models that are sensitive to scale difference, as in XGBoost and LightGBM, normalization was done to achieve steady convergence of the model.

6.5 MODEL DEVELOPMENT

Identity forecasting electricity demand was done with three popular machine learning models namely Random Forest (RF), XGBoost and LightBGM. These models have been selected relying on their good results in the time-series problems, their capacity to operate with nonlinearity and resistance to outliers.

a. Random Forest Regression

   Random Forest is an ensemble model, which builds a number of decision trees and combines their predictions. It has the benefit of nonlinear relationships and the absence of overfitting. The number of trees, maximum depth and minimum samples per split examples are hyperparameters that were tuned to enhance performance.

b. XGBoost Regression

   XGBoost is a powerful and efficient boosting-based algorithm. It constructs trees in sequence with a new tree correcting the error of previous trees. To optimize the accuracy, learning rate, max depth, subsample ratio and estimators number were tuned.

c. LightGBM Regression

   LightGBM is a gradient-boosting implementation that is optimized to be fast and scalable to large datsets.it was computationally efficient and content to execute. Optimization was done on parameters like number of leaves, learning rate and boosting rounds.

The models of Germany. France and Spain were trained separately where national demand characteristics were taken into account in the country.

6.6 MODEL EVALUATION

Forecasting metrics that are commonly used to measure model performance were used:

Mean Absolute Error (MAE)

Mean Absolute Percentage Error (MAPE)

Root Mean Squared Error (RMSE)

These measures are used to determine the similarity of the predicted values to the actual values in terms of both the magnitude and percentage of the values. It was observed that XGBoost offered the most accurate results in the three countries with the lowest error value in all cases. The XGBoost predictions were utilized in the subsequent analysis and optimization because of its accuracy.

6.7 DEMAND FORECASTING OF ELECTRICITY.

Once XGBoost was chosen as the most performing model, it was re-trained with the complete training dataset and used to predict the electricity demand in Germany, France and Spain. The forecast covered the short-term fluctuations and long-term consumption trends. These predictions were used as the input into the next optimization step.

The following is a linear programming model of the optimal allocation. After the demand values were modeled, Linear programming (LP) model was developed in order to efficiently distribute electricity using two strategies:

a. Proportional Allocation Model

   This model allocates electricity proportionately according to the expected demand of each country. The proportionate distribution is higher in countries with greater demand.

b. The Priority-Based Allocation Model.

   In this model, countries were given priority scores (high, medium, low) depending on the sensitivity of demand, pattern of use or preference of policy. These priorities are then used to apportion the electricity.

LP Objective Function: - The overall goal of the LP model was to reduce the discrepancy between the provided and demanded energy within a range of restrictions that included:

Available electricity in totality.

The lowest and highest supply limits.

Priority weights (priority-based model)

The results of the LP allocation were compared by country giving an insight on how the proportional and priority systems vary in the distribution pattern.

6.8 ENDING INTEGRATION AND ENDING.

Combination of the forecasting and optimization was the last step. The forecasted demand curves were plotted and the optimized values of allocation of each country were plotted. The findings underscore the role of using ML-driven prediction and optimization tools to assist in efficient entry planning and make decisions.

MATHEMATICAL FORMULATION OF THE LINEAR PROGRAMMING MODEL.

Linear programming in the project is adopted in the case of electricity shortage in order to distribute the scarce electricity supply among the three counties, Germany (DE), France (FR) and Spain (ES). The distribution is determined by the values of forecasted demand which are obtained with the best model XGBoost. There are two allocation techniques that are deemed:

1. Proportional Allocation where the electric power is distributed equitably on the basis of demand.
2. Priorities based on allocation where one country is allocated more priority than the other.

The objective of the LP formulation will be to reduce the shortage of electricity, meeting the limitations of supply and demand.

7.1 NOTATIONS

i. Let i = index of country, where
   i =1 Germany (DE)
   i=2 France (FR)
   i=3 Spain (ES)
ii. D_i = annual electric demand of country i (XGBoost) predicted.
iii. X_i = the amount of electricity that LP assigns to country i.
iv. S_i = electricity deficit in country i.

$$S_i = D_i - X_i$$

v. S= amount of electricity supply that is available when there is a shortage (in this case S is considered to be 80 students of total demand)

$$S = 0.8\, X \sum_{i=1}^{3} Di$$

vi. For priority-based method:
   Wi = the weight given to country i.
   The variables Xi and Si are all non-negative.

7.2 RELATIONSHIP BETWEEN DEMAND, ALLOCATION AND SHORTAGE

The allocation, demand and shortage of each country are related as follows.

$$X_i + S_i = D_i \quad for\ i = 1,2,3$$

i.e., Demand = Allocated Electricity + Shortage

when X_i increases the shortage S_i decreases, and vice versa.

7.3 PROPORTIONAL ALLOCATION MODEL

The purpose of the proportional allocation is to allocate scarce supply in a manner that the cumulative shortage to all the countries is reduced, with no country being prioritized.

Objective Function: -

$$Minimize\ Z_{prop} = \sum_{i=1}^{3} Si$$

This implies that the LP model minimizes shortages of countries with a high priority weight. The purpose attempts to decrease the total number of unsatisfied demand.

The objective tries to reduce the total amount of unmet demand.

Constraints:-

1. Demand Balance for each country

$$X_i + S_i = D_i \quad for\ i = 1,2,3$$

2. Total supply constraint

$$\sum_{i=1}^{3} Xi \leq S$$

The total electricity allocated to all country can't exceed the available supply.

3. Non-negativity constraints

$$X_i \geq 0, \quad S_i \geq 0 \quad for\ i = 1,2,3$$

4. No over-allocation

   Since $X_i + S_i = D_i$ and , $S_i \geq 0$,it ensures that:

$$X_i \leq D_i$$

   This implies that no nation can get more electricity than demand.

7.4 PRIORITY BASED ALLOCATION MODEL

The priority-based allocation model allocates resources based on the priority assigned to the process. In real life situation, there may exist countries that can be considered more critical than others based on other such factors like industrial value, population or vital services such as hospitals. In that place this is allocated. In this case, nations are given a weight Wi of priority. The greater the weight value, the heavier is the penalty on shortage by that country in the objective function.

Objective function:

$$Minimize\ Z_{prio} = \sum_{i=1}^{3} Wi\ Si$$

This implies that LP model makes more effort in minimising shortages to countries whose weights of priority are higher.

Constraints: -

1. Demand Balance

$$X_i + S_i = D_i \quad for\ i = 1,2,3$$

2. Total supply constraint

$$\sum_{i=1}^{3} Xi \leq S$$

3. Non-negativity constraints

$$X_i \geq 0, \quad S_i \geq 0 \quad for\ i = 1,2,3$$

Such formulation makes sure that the total amount of electricity allocated does not exceed the amount of electricity available, but the distribution will vary according to the priority weights.

7.5 INTERPRETATION OF THE LP MODELS

In the proportional method, the allocation is equitable in demand share as all countries have the same percentage of shortage. With the priority model, priority countries will receive more allocation as compared to low-priority countries. The total supply that has been allotted is the same in both and it is only distribution pattern that varies.

PROJECT OUTPUTS

8.1 MODEL COMPARISON TABLE

MODEL	MAE	RMSE	MAPE
Random Forest	165.308	256.75	0.33889
XGBoost	149.148	244.842	0.31287
LightGBM	169.703	269.283	0.35443

Table 2: Performance Comparison of Machine Learning Models for Germany

MODEL	MAE	RMSE	MAPE
Random Forest	571.704	1573.473	1.1652
XGBoost	609.013	1507.726	1.1253
LightGBM	624.842	1509.439	1.2965

Table 3: Performance Comparison of Machine Learning Models for France

MODEL	MAE	RMSE	MAPE
Random Forest	281.498	404.376	1.0846
XGBoost	285.124	419.626	1.1322
LightGBM	297.714	442.156	1.1835

Table 4: Performance Comparison of Machine Learning Models for Spain

8.2 FORECASTING OUTPUT GRAPHS

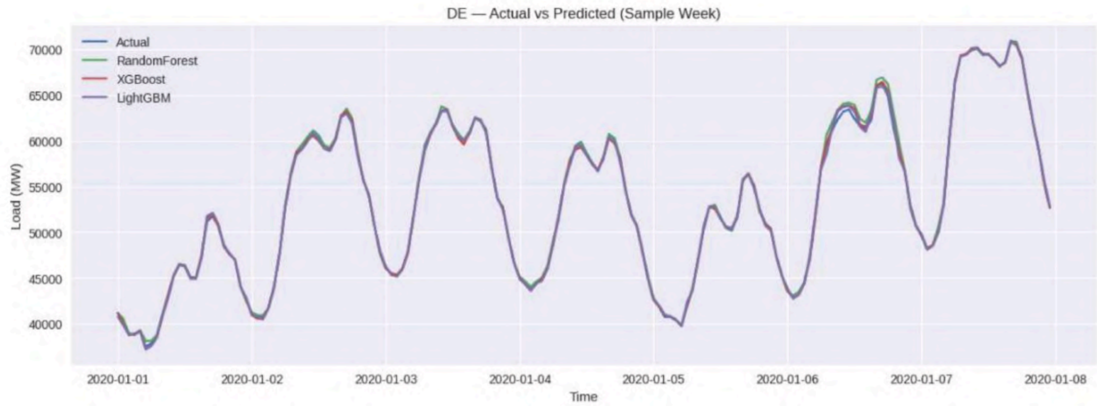

Figure 1: Actual vs Predicted Load for Germany (Sample Week)

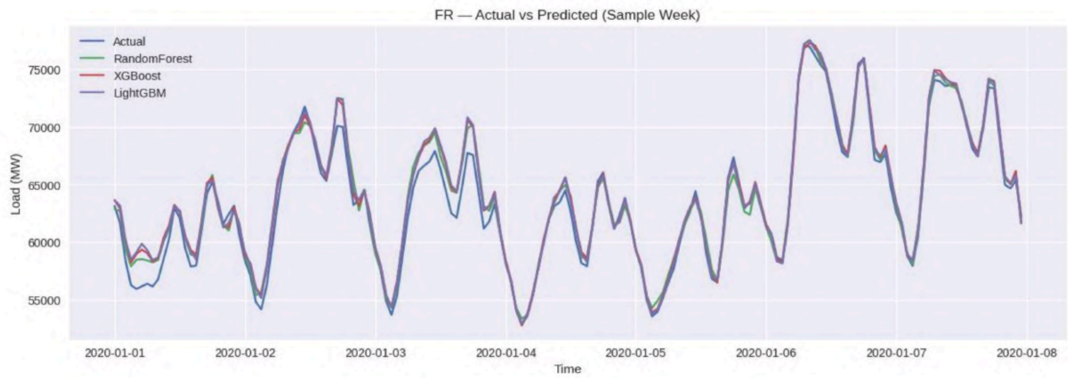

Figure 2: Actual vs Predicted Load for France (Sample Week)

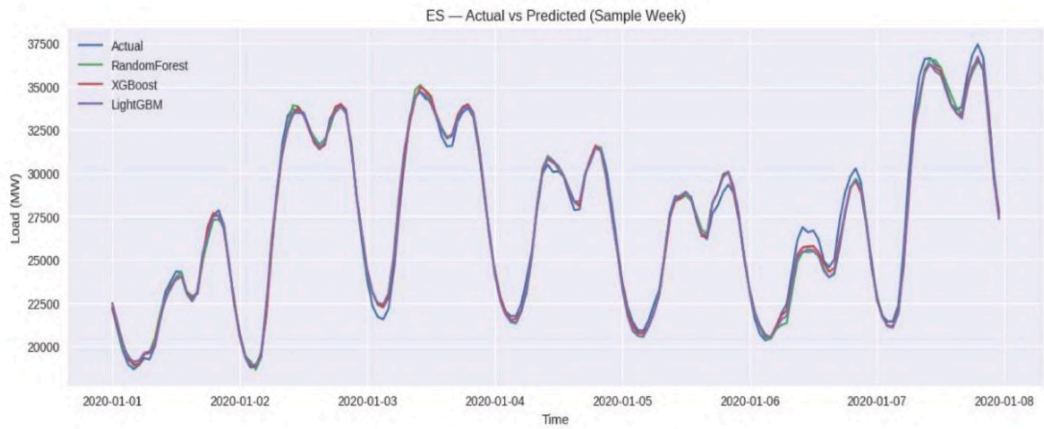

Figure 3: Actual vs Predicted Load for Spain (Sample Week)

8.3 ERROR METRIC COMPARISON GRAPHS

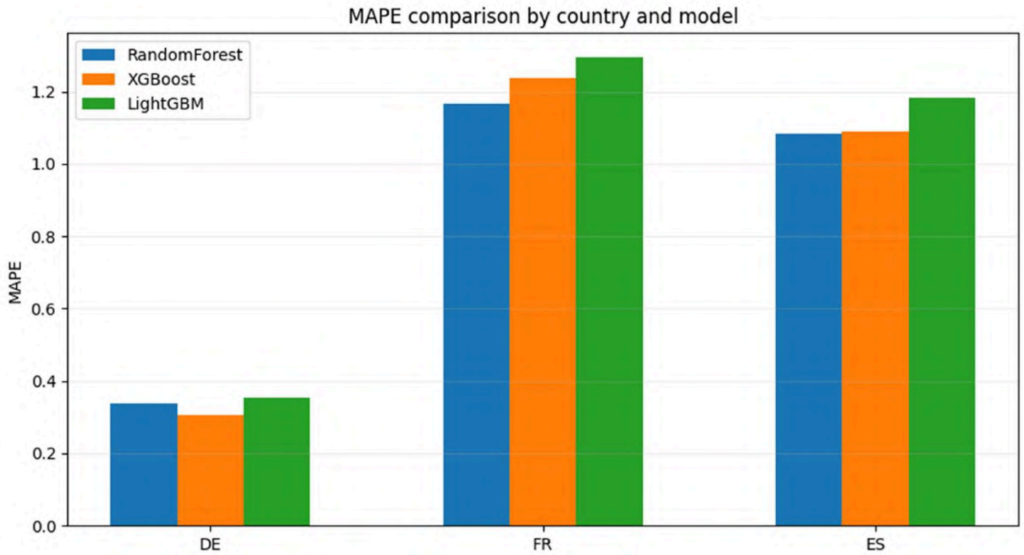

Figure 4: MAPE Comparison of Machine Learning Models (DE, FR, ES)

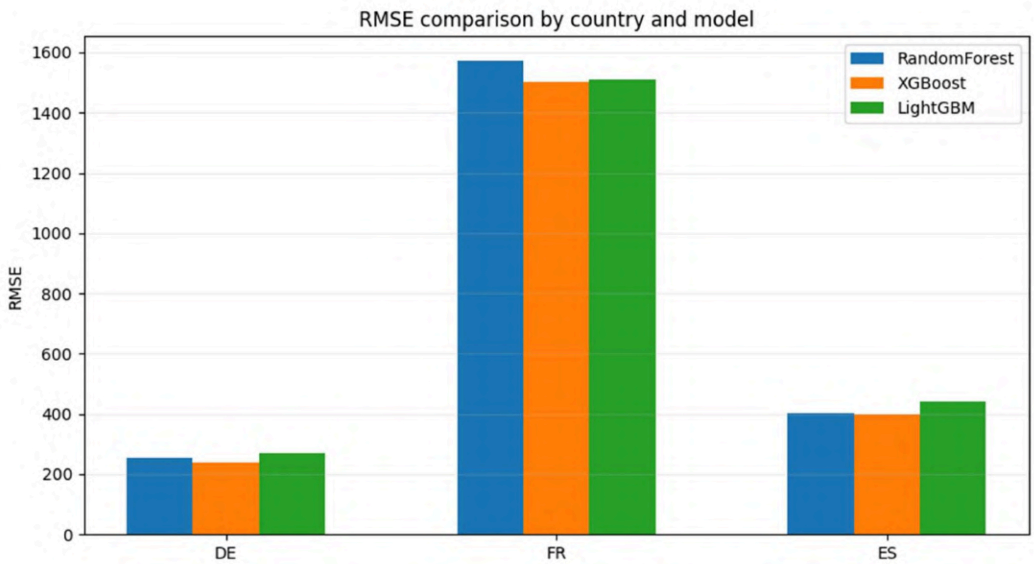

Figure 5: RMSE Comparison of Machine Learning Models (DE, FR, ES)

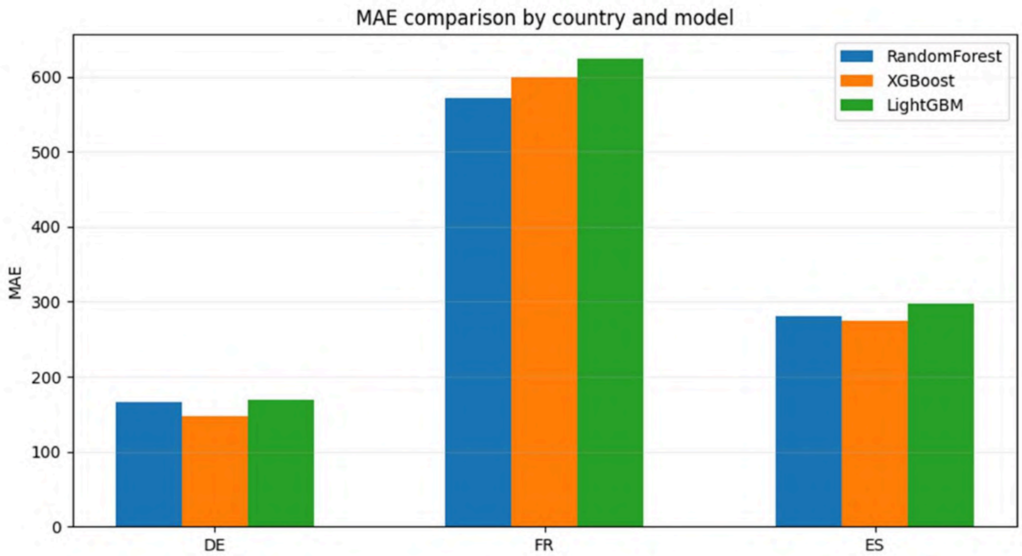

Figure 6: MAE Comparison of Machine Learning Models (DE, FR, ES)

8.4 FEATURE IMPORTANCE OUTPUT

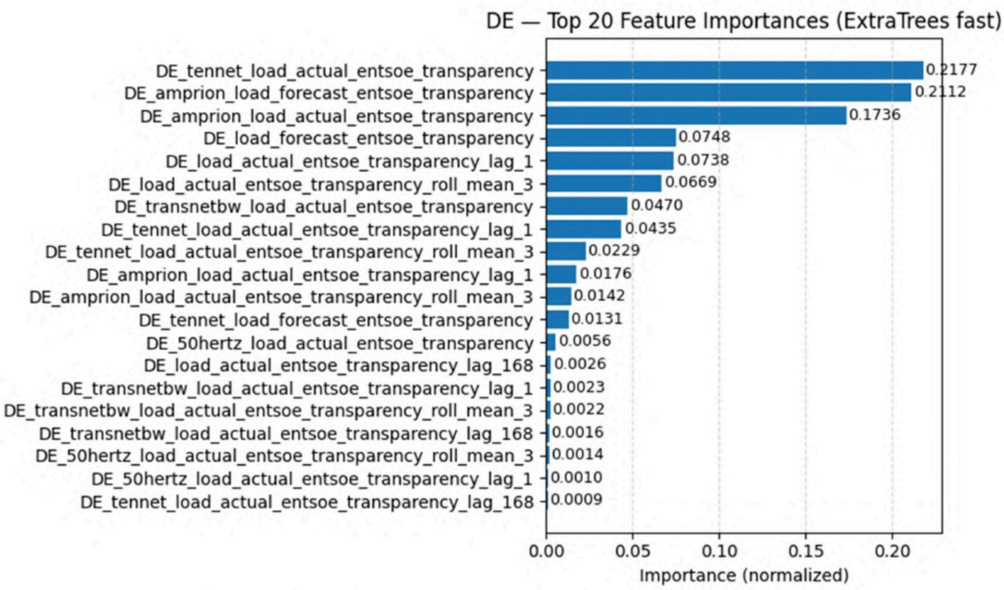

Figure 7: Feature Importance Plot of XGBoost Model

8.5 TOTAL FORECASTED DEMAND OUTPUT

COUNTRY	TOTAL PREDICTED DEMAND (SUM OF HOURLY PREDICTION)
Germany (DE)	348,815,845
France (FR)	320,082,511
Spain (ES)	177,299,405

Table 5.: Total Predicted Annual Electricity Demand Using XGBoost

8.6 LINEAR PROGRAMMING ALLOCATION OUTPUTS

COUNTRY	DEMAND	PROPORTIONAL ALLOCATION	PROPORTIONAL SHORTAGE	PRIORITY ALLOCATION	PRIORITY SHORTAGE
Germany	348,815,845	279,052,676	69,763,169	315,465,067	33,350,778
France	320,082,511	256,066,009	64,016,502	241,232,473	78,850,038
Spain	177,299,405	141,839,524	35,459,881	120,260,668	57,038,737

Table 6: LP-Based Electricity Allocation for Proportional and Priority Models

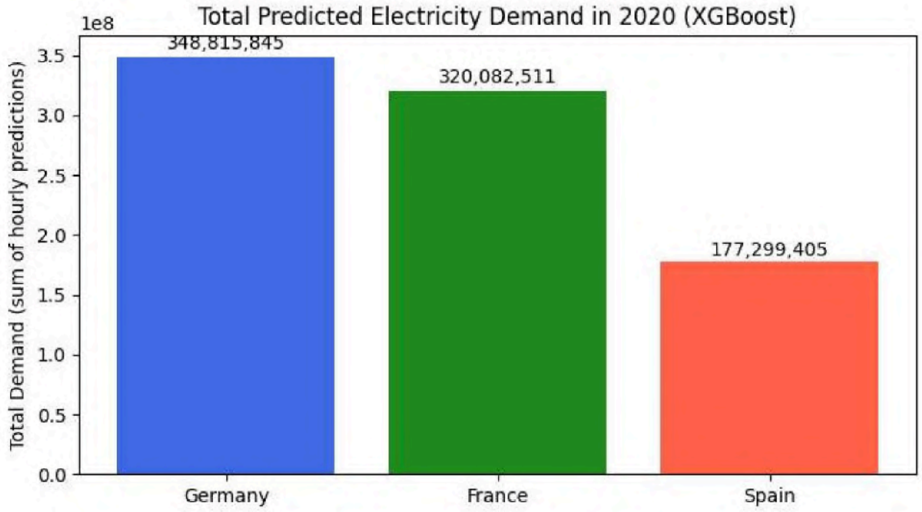

Figure 8: Total Predicted Annual Electricity Demand for Germany, France and Spain Using XGBoost

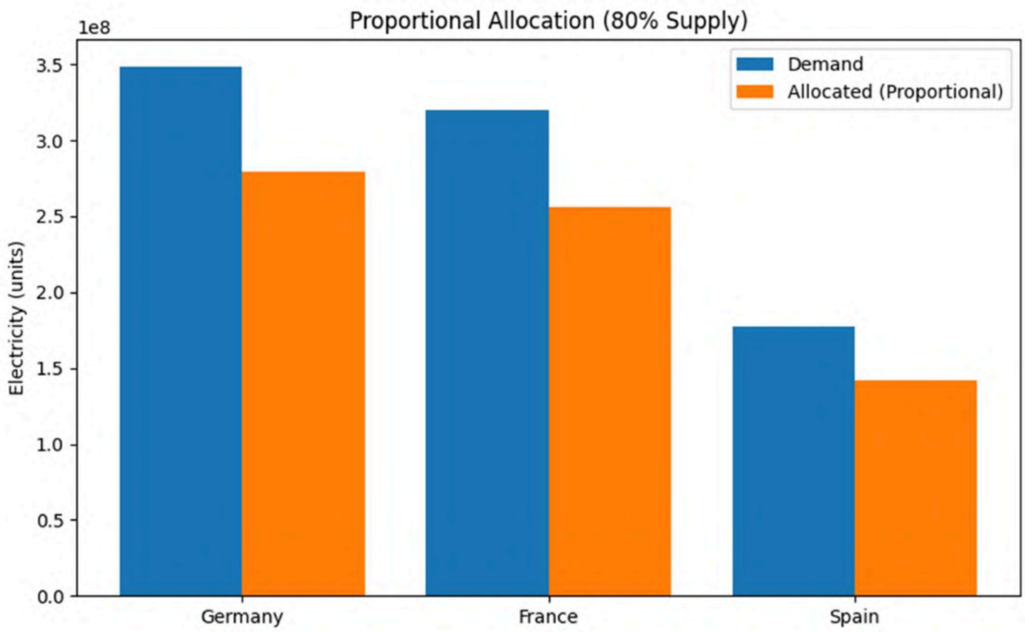

Figure 9: Proportional Allocation Output (Graph)

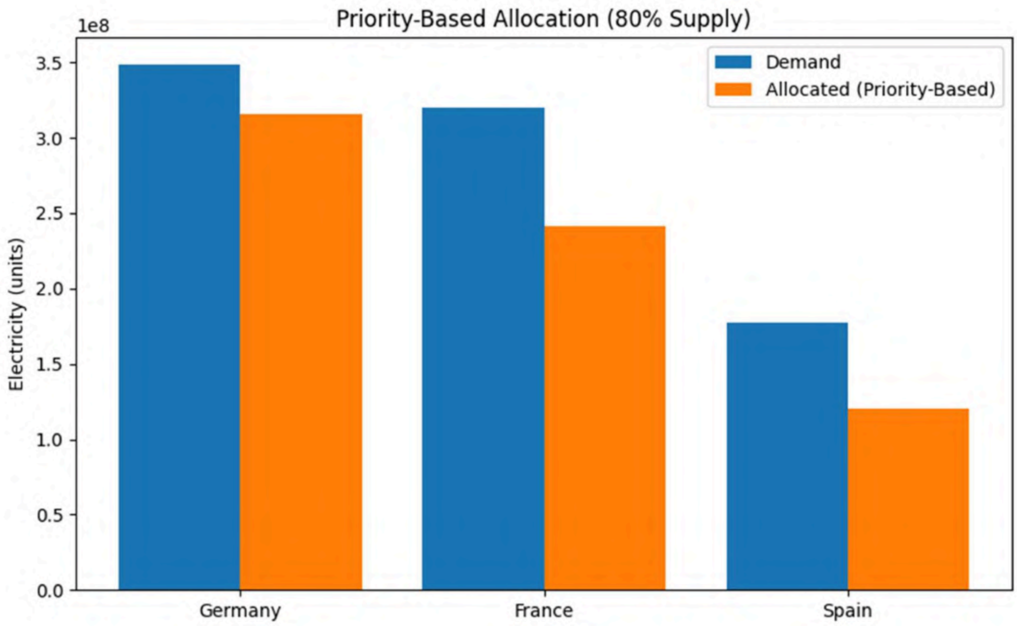

Figure 10: Priority-Based Allocation Output (Graph)

RESULTS AND DISCUSSION

In this section, it has been given in detail how machine learning forecasting results have been made to be accurate in Germany, France and Spain and then the results of the linear programming-based allocation is presented when there are constrained electricity supply. The discussion gives an explanation of the behaviour of the models, the nature of the forecasted demand values and implications of the allocation strategies. All the interpretations are founded on the outputs that are displayed in the Project Output section.

9.1 DISCUSSION OF MODEL PERFORMANCE

Random Forest, XGBoost and LightGBM performance on Germany are shown in Table 1. The findings reveal that XGBoost has the minimal MAE (149.148), RMSE (244.842) and MAPE (0.31287%), which means that it best fits the demand patterns of Germany. The boosting effect of XGBoost could be explained by the fact that it is more effective in the minimization of residual errors. Random Forest has an average level of performance, yet its error values are still high. LightGBM has the largest value of errors compared to the other two models, which indicates the inclination towards error fluctuations at the value peaks. All of this leads to considering XGBoost as the best model to use in Germany.

The France model performance (summarised in Table 2) also shows that again, XGBoost gives the lowest percentage error (MAPE=1.1253%). Even though the MAE of random forest is a bit smaller, the generalisation of XGBoost is shown by the lower values of MAPE and RMSE. The mistakes made by France are quite high in all models than that of Germany because of more sporadic consumption patterns which are dependent on the seasonal residential consumption. LightGBM also has the lowest performance, which proves that it is less adapted to changing load profile in France.

In case of Spain, in Table 3, the Random Forest has the lowest values of MAE and RMSE (281.498 and 404.376 respectively), which is numerically the best. Nevertheless, XGBoost continues to be competitive with a MAOE of 1.1322% and it also exhibits the same behaviour in all countries. The fact that Spain has a lower total demand and comparatively less turbulent patterns of consumption means that Random Forest will work better than XGBoost in this area through latter is more consistent in general. This is what makes this consistency useful to multi

country forecasting systems, and is the reason why XGBoost is applied in the end of year.
Comparison of the real and forecasted load patterns.
The predicted and actual load curves show how well each model is able to capture the weekly behaviour. XGBoost in all three countries adheres to the real load curve, a fact that is more evident than the other models. It generates the morning peak and evening peak patterns with the least amount of deviation. Random Forest gives smoother curves, thus, it is underestimated in case of sharp peaks or abrupt changes in demand. LightGBM shows significant variations during the high load times, particularly in the case of Germany and France. These visual trends are in agreement with these error measures, which supports the fact that XGBoost offers the most alignment with the actual demand data.

The comparison graphs in the error metric also indicate the similarity of the models performance. In the case of Germany and France, XGBoost has the smallest errors in all metrics, which means that is highly effective in nonlinear and time-varying load behaviours modelling. France has generally more MAPE values since it is more volatile in demand. The results of Spain are slightly different, with random forest showing a better number, although the gap between RF and XGBoost is insignificant. The high errors Achieved by LightGBM in all countries are a confirmation that leaf-wise growing of trees cause fluctuations that decrease the accuracy of the forecasts. These findings justify the use of XGBoost as the main forecasting model.

9.2 ANNUAL DEMAND FORECAST DISCUSSION.

Table 4 indicates that the overall predicted annual demand of the electricity is the highest in Germany with the demand being 348,815,845 kWh. France comes next at 320,082,517 kwh with Spain at the lowest annual consumption of 177,299,405,000 kWh. These variations are consistent with established European trends of electricity consumption with Germany having a large industrial sector contributing towards its load. The demand in France is affected by the domestical heating and cooling, Spanish climate is relatively mild, which makes the use less significant. These predictions can be justified based on the trend in the model performance as explained earlier. the values are the foundation of the linear programming allocation model applied in the following step of analysis.

9.3 LINEAR PROGRAMMING CAPACITY DISCONTINUITIES ANALYSIS.

Table 5 summarising the results of the linear programming offers an understanding of the way electricity may be allocated fairly or strategically in case of a shortage of supply. With proportional allocation, every nation gets an allocation in a strictly proportional way with regards to its estimated demand. This leads to the effect of equal proportionate shortages in all the three countries. The deficit is 69,763,169 units in Germany, 64,016,502 units in France and 35,459,881 units in Spain. Such an approach is fairly consistent, but not considerate of different economic or industrial priorities.

The weighted importance between the countries is brought in by the priority-based allocation. Germany, which is given a priority, is allocated more units, namely 315,465,067 units of 120,260,668 units. This causes an increase in shortages in France and Spain as opposed to proportional method. The priority-based model thus represents real-life resource allocation policies where the high-impact or needy regions are given priority in the event of limited availability.

9.4 OVERALL INTERPRETATION

The predictive blend and optimisation give substantial information with respect to the multi country electricity planning. XGBoost has proven to be a great predictor model in Germany and France but also competitive in Spain. The forecasted demand values of the project per year are in line with what would be utilized and justifies the forecasting method. The outputs of the linear programming show that the allocation strategies are very different when fairness or priority are put into focus. All in all, machine learning and the use of optimisation methods create an efficient decision support tool in electricity demand management and resource planning in the short term.

LIMITATIONS

Despite the fact that the results of this study were correct and significant, there are limitations associated with the work. The models that were used in forecasting were based on the past demand and the time aspect mainly; they did not incorporate external conditions that may change the consumption pattern like the temperature, the humidity, the change in the economic activity, the change in the policy, etc. The dataset was also thought to be clean and consistent, but actual load data can have anomalies, or sudden changes of behavior. The linear programming model employed in the allocation stage made simplified assumptions, including fixed supply constraints and constant priority weights and omitted operational constraints, including transmission losses, a regional generation capacity, renewable intermittency or dynamic tariffs. These simplifications allow the model to be analyzed in the academic world but have the potential to restrict its direct implementation in an actual power system.

FUTURE SCOPE

This can be further developed by adding other features like weather variables, indicators of industrial activity and social-economic parameters to enhance better predicting ability particularly in those countries where consumption is not consistent. Incorporating LSTM, GRU or hybrid CNN-LSTM networks which are more advanced deep learning architectures could be considered to learn long-term and nonlinear behavior better. The optimization model can be developed into multi-objective model where the cost, fairness, carbon emissions, availability of renewable and grid constraints are taken into account at the same time. The probabilistic forecasting could also be considered in the future research to determine the uncertainty and the creation of a real-time dynamic allocation system updating the forecast and allocations in the future. These additions would enhance the model to become stronger and fit for real-life energy planning.

CONCLUSION

The project included a statistically based model of predicting electricity demand in Germany, France and Spain and optimizing resource allocation of the countries in the circumstances of scarce supply. There were three machine learning models (Random Forest, XGBoost and LightGBM) prepared and assessed based on ten years of past hourly demand data. The comparative analysis revealed XGBoost to have the most stable performance in general with the lowest error values in Germany and France and is competitive in Spain. The actual- versus-predicted load curves and the comparison of error metrics also confirmed the high quality of trend-tracking of XGBoost. This model was used to predict the annual electricity demand of a respective country with realistic estimates that were in line with what was known about the consumption behavior in Europe.

Such predictions were the foundation to an optimization model that was aimed at analyzing supply allocation in the case of a shortage. In order to create a realistic constraint, the supply was purposely limited to 80 per cent of the overall estimated annual demand. Linear programming was used to implement two strategies, namely, proportional allocation and priority-based allocation. The proportional approach apportioned electricity equitably on a demand share basis, and the priority-based approach was a strategic approach that shifted a greater volume of supply to areas with higher priorities. The discrepancy between the two results demonstrated the importance of optimization decisions in making a significant difference to the distribution of supplies in a situation where resources are not adequate. In general, the forecasting and optimization combination illustrates the existence of a realistic and efficient decision support system in managing the electricity demand of several countries.

REFERENCES

[1] A. H. Abdalsatar and R. Y. Salman, “Allocation of energy resources in electrical networks using linear programming: Optimization of energy allocation and distribution in electrical networks using operations research techniques,” *International Journal of Professional Studies*, vol. 20, no. 2, pp. 88–106, 2025, doi: 10.37648/ijps.v20i01.006.

[2] C. Clack, Y. Xie, and A. MacDonald, “Linear programming techniques for developing an optimal electrical system including high-voltage direct-current transmission and storage,” *International Journal of Electrical Power & Energy Systems*, vol. 68, pp. 103–114, 2015, doi: 10.1016/j.ijepes.2014.12.049.

[3] S. Fan and R. J. Hyndman, “Forecasting electricity demand in Australian National Electricity Market,” in *2012 IEEE Power and Energy Society General Meeting*, San Diego, CA, USA, 2012, pp. 1–4, doi: 10.1109/PESGM.2012.6345304.

[4] M. Javidsharifi, H. P. Arabani, J. C. Vasquez, and J. M. Guerrero, “Transfer learning for electricity demand forecasting in Nordic cities: A case study of Copenhagen and Reykjavik,” in *2025 32nd International Conference on Systems, Signals and Image Processing (IWSSIP)*, Skopje, North Macedonia, 2025, pp. 1–5, doi: 10.1109/IWSSIP66997.2025.11151865.

[5] A.-R. Kim, D. Park, C.-S. Lee, and H. Chang, “A comparative analysis of time series forecasting methods for short-term electricity demand prediction,” in *2023 14th International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju Island, Republic of Korea, 2023, pp. 1235–1238, doi: 10.1109/ICTC58733.2023.10392314.

[6] A. V. Mendoza, J. H. Tovar Hernández, and H. F. Ruiz Paredes, “Electricity demand forecasting for the BCS region of the Mexican SEN using artificial neural networks,” in *2023 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC)*, Ixtapa, Mexico, 2023, pp. 1–6, doi: 10.1109/ROPEC58757.2023.10409474.

[7] D. Miyao and M. Nakamura, “Time and day-based peak electricity demand forecasting: A comparative analysis of machine learning models for peak cut and decarbonization,” in *2024 13th International Conference on Renewable Energy Research and Applications (ICRERA)*, Nagasaki, Japan, 2024, pp. 1429–1433, doi: 10.1109/ICRERA62673.2024.10815184.

[8] P. Pandiyan, A. R. Gurudeep, M. Gowtham, M. Naveen Kumar, and S. Bala Subramanian, “Electrical load demand forecasting using machine learning algorithm,” in *2025 Second International Conference on Cognitive Robotics and Intelligent Systems (ICC-ROBINS)*, Coimbatore, India, 2025, pp. 1–6, doi: 10.1109/ICC-ROBINS64345.2025.11086163.

[9] A. Saji, A. Prakash, and M. S. Krishnan, “Electricity demand forecasting in Kerala using machine learning models,” in *2023 3rd International Conference on Intelligent Technologies (CONIT)*, Hubli, India, 2023, pp. 1–6, doi: 10.1109/CONIT59222.2023.10205925.

[10] A. Thakur, K. A. Shukla, A. Choudhary, and J. Atrey, “Predictive analysis of energy consumption and electricity demand using machine learning techniques,” in *2023 International Conference on Smart Systems for Applications in Electrical Sciences (ICSSES)*,Tumakuru, India, 2023, pp. 1–6, doi: 10.1109/ICSSES58299.2023.10200636.

[11] B. Yan, A. Beltre, P. Saha, and M. Govindaraju, “Exploring enhanced dominant resource fairness using linear programming calculated weights,” in *2024 IEEE/ACM 17th International Conference on Utility and Cloud Computing (UCC)*, Sharjah, United Arab Emirates, 2024, pp. 174–181, doi: 10.1109/UCC63386.2024.00031.

[12] Grandón, T. G., Schwenzer, J., Steens, T., & Breuing, J. (2023). Electricity demand forecasting with hybrid classical statistical and machine learning algorithms: Case study of Ukraine. *Applied Energy*, *355*, 122249. https://doi.org/10.1016/j.apenergy.2023.122249

[13] Wang, C., & Lai, Y. (2022). Integrating Deep Learning with Bi-Objective Linear Programming to Conduct Electricity Demand Forecasting and Supply Optimization. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4040686

[14] Rao, H., Li, J., & Sun, X. (2025b). Demand forecasting and allocation Optimization of green power grid supply chain based on Machine Learning algorithm: A study based on the Whole-Process data of Power Grid materials. *Sustainability*, *17*(3), 1247. https://doi.org/10.3390/su17031247

[15] Gulaydin, O., & Mourshed, M. (2025). Machine learning for subnational residential electricity demand forecasting to 2050 under shared socioeconomic pathways: Comparing tree-based, neural and kernel methods. *Energy*, *336*, 138195. https://doi.org/10.1016/j.energy.2025.138195

[16] Lotfi, H., Vafadoost, P., Rokhsati, H. *et al.* An optimized machine learning approach for short-term electricity demand forecasting considering environmental and historical influencing factors. *Discov Appl Sci* **7**, 1015 (2025). https://doi.org/10.1007/s42452-025-07681-z