VAASAN AMMATTIKORKEAKOULU
UNIVERSITY OF APPLIED SCIENCES

Dinesh Karunarathne

MACHINE LEARNING AND STATISTICAL APPROACHES FOR FORECASTING ELECTRICITY DEMAND IN VAASA

Information Technology

2025

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to the many individuals who have helped and supported me in numerous ways throughout the journey of this Master's thesis. Their guidance, assistance, and encouragement have been invaluable to the successful completion of this research.

First and foremost, I am deeply indebted to my supervisor, Dr. Ghodrat Moghadampour, Principal Lecturer at VAMK, for his insightful guidance, unwavering support, and constant encouragement throughout my study. His expertise and constructive feedback were instrumental in shaping this research.

I am also very grateful for the crucial assistance I received in gathering the necessary data for this project. I would like to extend my thanks to Tiia Haapakoski of Statistics Finland for her helpful guidance in navigating data collection processes. My appreciation also goes to Hanna Nurmilo of Fingrid Oyj for her valuable direction in obtaining relevant data. Furthermore, I wish to sincerely thank Mari Böling and Jaakko Yliaho of Vaasan Sähkö, who were instrumental in providing the historical hourly electricity consumption data for Vaasa, which formed the backbone of this study.

The completion of this thesis would not have been possible without the contributions of these individuals.

Dinesh Karunarathne
Vaasa, Finland
May 26, 2025

VAASAN AMMATTIKORKEAKOULU
UNIVERSITY OF APPLIED SCIENCES
Masters in Cloud-Based Software Engineering

ABSTRACT

Author	Dinesh Karunarathne
Title	Machine Learning and Statistical Approaches for Forecasting Electricity Demand in Vaasa
Year	2025
Language	English
Pages	86
Supervisor	Dr. Ghodrat Moghadampour

Electricity is vital to modern life, but stable supply is increasingly challenged by changing consumption patterns, climate variability, and renewable integration. This research focuses on forecasting electricity demand in Vaasa, Finland, a regional energy hub. Tailored short and long-term prediction models were developed to address grid stability challenges driven by factors such as climate variability.

The study utilized established load forecasting literature, comparing traditional statistical models against machine learning techniques and identifying key demand drivers. A quantitative methodology was employed using historical Vaasa-specific data, including electricity consumption and weather information. This involved comprehensive data acquisition, preprocessing, feature engineering (such as deriving 'Is_workday' and 'Sun_Flag' variables), and the iterative development and evaluation of Multiple Linear Regression (MLR), Random Forest, and XGBoost models for monthly, daily, and hourly forecasts, using metrics like R^2 and MAPE.

The results demonstrated strong predictive capabilities, with an MLR model using temperature variables proving best for monthly forecasts with R^2=96.3% (Table 6). For daily forecasts, MLR outperformed the tested machine learning models with R^2=92.8% (Table 7). In hourly forecasting, XGBoost achieved a marginally better R^2=90.2% (Table 17), closely followed by MLR with R^2=90.1%. Temperature and various temporal features were consistently the most dominant demand drivers. The study concludes that for Vaasa's context, well-specified multiple linear models can be highly effective, offering valuable and interpretable tools for local energy management and strategic planning.

Keywords electricity demand forecasting, machine learning, multiple linear regression, short-term load forecasting, long-term forecasting

CONTENTS

FIGURES

TABLES

ABBREVIATIONS

AI	Artificial Intelligence
ANN	Artificial Neural Network
ANOVA	Analysis of Variance
AR	Autoregressive Model

ARIMA	Autoregressive Integrated Moving Average
c/kWh	Cents per kilowatt-hour
CV-RMSE	Cumulative Variation of Root Mean Square Error
DNN	Deep Neural Network
EVs	Electric Vehicles
GDP	Gross Domestic Product
GWh	Gigawatt-hour
ISO	International Organization for Standardization (used in context of ISO Weekday)
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
MLR	Multiple Linear Regression
MAPE	Mean Absolute Percentage Error
MSE	Mean Squared Error
MWh	Megawatt-hour
NN	Neural Network
OHE	One-Hot Encoded
Q-Q	Quantile-Quantile
R^2	Coefficient of Determination
RMSE	Root Mean Squared Error
SRM	Simple Regression Model
STLF	Short-Term Load Forecasting
TEC	Total Electricity Consumption
XGBoost	Extreme Gradient Boosting

1 INTRODUCTION

Accurate electricity demand forecasting is paramount for efficient energy management and grid stability, especially in the context of evolving energy landscapes influenced by renewable integration and changing consumption patterns. This chapter lays the groundwork for the thesis, which focuses on developing tailored forecasting models for the city of Vaasa. It will begin by establishing the background and motivations for this research, followed by a clear outline of the study's objectives. Subsequently, the chapter will detail the methodological approach adopted for model development and evaluation, and conclude by acknowledging the inherent limitations of the study.

1.1 Background

Powering both home and industrial operations worldwide, electricity is a fundamental component of contemporary life. Growing worldwide populations, fast industrial expansion, and the unpredictable consequences of climate change have made it more important than ever to provide a steady and uninterrupted supply of power. In the past, fundamental necessities like heating, lighting, and industrial machinery were the main drivers of energy consumption patterns, which were generally predictable. However, the rise in the use of electric vehicles (EVs), the spread of sophisticated household gadgets, and the growing need for industrial goods define the current environment. At the same time, the paradigm for electricity generation is changing due to the shift to renewable energy sources like solar and wind, which adds new dynamics to supply and demand forecasting. This shift, while environmentally beneficial, necessitates sophisticated forecasting models to manage the inherent variability of renewable energy generation. These evolving global trends underscore the acute need for robust forecasting methodologies. The city of Vaasa, a significant energy hub in Finland with its own dynamic energy landscape, serves as a pertinent and accessible case study for developing and evaluating such models.

1.2 Motivations

The primary motivation for this research is to address the critical gap in tailored electricity demand forecasting solutions that can support sustainable and efficient energy management, particularly within the unique context of Vaasa. In Vaasa, a city with a strong focus on renewable energy and a significant industrial presence, the ability to predict electricity demand accurately is crucial for both long-term strategic planning and short-term operational efficiency. The rapid growth of domestic EV charging, coupled with the increasing integration of intermittent renewable energy sources, presents unique challenges for grid stability and resource allocation. By developing and deploying robust forecasting models, this research aims to contribute to the city's ability to proactively address these challenges, ensuring a reliable and cost-effective electricity supply.

1.3 Objectives

The primary objective of this study is to develop both short-term and long-term electricity demand forecasting models tailored to the specific characteristics of Vaasa. Specifically:

- To develop a long-term forecasting model that will assist authorities, such as government bodies and electricity generation companies, in formulating policies and planning infrastructure to meet future electricity demand efficiently.

- To create a short-term forecasting model that will support operational tasks like daily load balancing, ensuring stable and cost-effective electricity distribution.
- To evaluate the effectiveness of machine learning algorithms in comparison to traditional statistical models for electricity demand forecasting in the context of Vaasa.
- To identify and analyze the key features influencing electricity demand within the city.

1.4 Method

This research will employ a quantitative approach, leveraging historical electricity consumption data, weather data, and other relevant information from the city of Vaasa. The methodology will involve following key stages:

- **Data Acquisition and Preprocessing:** Gathering and cleaning data from various sources, including energy providers and weather services.
- **Feature Engineering and Selection:** Identifying and selecting relevant features that influence electricity demand. Statistical analysis, including techniques such as Analysis of Variance (ANOVA) and Pearson correlation tests to assess feature significance, will be performed, utilizing tools such as JASP 0.19.3 for this purpose.
- **Model Development:** Training and evaluating both statistical and machine learning forecasting models. The development and implementation of these models will be carried out using the Python programming language, primarily leveraging the scikit-learn library for its comprehensive suite of tools for machine learning and statistical modeling.

- **Model Evaluation:** Using metrics such as Mean Absolute Percentage Error (MAPE) and R-squared to evaluate model performance and determine the most accurate and reliable forecasting models for Vaasa.

1.5 Limitations

This study is subject to certain limitations. Firstly, the accuracy of the forecasting models is dependent on the quality and availability of historical data. Secondly, the models may not fully capture unforeseen events or abrupt changes in electricity consumption patterns, such as sudden industrial expansions or significant shifts in consumer behavior. Finally, the localized nature of the study, focusing on Vaasa, means that while the methodological approach is transferable, the specific findings and model parameters may have limited generalizability to other regions with different energy landscapes. These limitations are acknowledged, and the research aims to provide robust insights within its defined scope.

2 LITERATURE REVIEW

Accurate electricity demand forecasting is crucial for both the operational efficiency and strategic planning of power systems. Short-term load forecasting (STLF), typically covering horizons from hours to a week, is essential for daily operations such as load balancing, scheduling, and market bidding (López et al., 2018; Kuo & Huang, 2018). Longer-term forecasting, spanning months to years, informs critical decisions regarding infrastructure investment, energy policy formulation, and resource planning (Kandananond, 2011; Rahman et al., 2018). This review synthesizes findings from four distinct studies to evaluate methodologies relevant to developing short-term and long-term electricity demand forecasting models specifically for Vaasa, Finland. It compares the forecasting approaches, the performance of machine learning versus traditional statistical models, and the key features identified as influencing electricity demand, thereby contextualizing the objectives of the proposed research for Vaasa.

Long-term forecasting, essential for policy and infrastructure planning, is addressed using various approaches in the reviewed literature, primarily focusing on national-level annual data. Kandananond (2011), forecasting Thailand's annual demand (1986-2010), utilized and compared Autoregressive Integrated Moving Average (ARIMA), Artificial Neural Network (ANN), and Multiple Linear Regression (MLR) models. The goal was explicitly to support national energy policy development. Similarly, Rahman et al. (2018), while labelling their work "short-term" forecast India's Total Electricity Consumption (TEC) up to 2030 using annual data, making it relevant for medium-to-long-term planning. They employed several methods including MLR, Simple Regression (SRM), and various exponential smoothing techniques (Holt's, Brown's, Damped Trend).

Comparing the methodologies, both studies rely heavily on regression techniques (MLR) and time-series approaches (ARIMA, exponential smoothing). Kandananond (2011) uniquely incorporates ANN into this long-term comparison. Key drivers identified for long-term forecasting in these national contexts are primarily socio-economic: Kandananond (2011) used population, Gross Domestic Product (GDP), stock index, and export revenues, while Rahman et al. (2018) focused on population, GDP, and GDP per capita. Rahman et al. (2018) explicitly concluded that GDP was a better predictor of TEC for India than population or GDP per capita, based on correlation and regression model fit (R-squared values). Both studies underscore the link between accurate forecasting and effective energy policy and planning.

Short-Term Load Forecasting (STLF) techniques are vital for daily operational tasks, and the literature reflects a strong focus on Artificial Intelligence (AI) and time-series models for this purpose. López et al. (2018) provide a detailed comparison of Neural Networks (NN) and Autoregressive (AR) models for STLF (one hour to several days ahead) within the context of the Spanish national grid operator. Their work highlights the use of STLF for system reliability, network efficiency, and optimizing market bidding. They emphasize that STLF is highly specific to each case and often involves hybrid models combining classical and modern techniques. Kuo & Huang (2018) focus specifically on advancing STLF accuracy using a Deep Neural Network (DNN) model, positioning it within the context of smart grid technology development. They argue that even small improvements in STLF accuracy can yield significant cost reductions.

Methodologically, López et al. (2018) compare NN (an AI approach) directly with AR model, measuring performance via Root Mean Square Error (RMSE) over a year. Kuo & Huang (2018) propose a specific DNN architecture and evaluate it against other AI algorithms using Mean

Absolute Percentage Error (MAPE) and Cumulative Variation of Root Mean Square Error (CV-RMSE). They report high accuracy for their DNN model. Both studies acknowledge the broader context of STLF, including essential stages like data pre-processing, input variable selection (especially handling calendar effects like special days, weather variables), and determining appropriate training time frames (López et al., 2018; Kuo & Huang, 2018).

The comparison between machine learning (ML)/AI techniques (like ANN, NN, DNN) and traditional statistical models (like ARIMA, AR, MLR, Smoothing) yields nuanced results across the reviewed studies. Kandananond (2011), comparing ANN with ARIMA and MLR for longer-term forecasting, found that while ANN achieved the lowest MAPE, the difference was not statistically significant compared to the simpler, more parsimonious ARIMA and MLR models. For STLF, López et al. (2018) found the traditional AR model performed slightly better than NN under ideal conditions, but the NN proved more accurate under certain "stress situations," suggesting NNs may offer more robustness when conditions deviate from the norm.

Kuo & Huang (2018) implicitly advocate for the superiority of advanced ML, specifically DNNs, over traditional methods like time series and linear regression for STLF, framing deep learning as a recent advancement integrated with ML techniques. They position their high-accuracy DNN model as an improvement within the AI forecasting landscape. Rahman et al. (2018) focus solely on traditional regression and smoothing methods for their long-term forecast for India.

Collectively, these studies suggest that while ML/AI models, particularly NNs and DNNs, show significant promise and can outperform traditional methods, especially in complex STLF tasks or under specific conditions (López et al., 2018; Kuo & Huang, 2018). Traditional methods remain competitive, relevant, and sometimes preferable due to their simplicity, especially for longer-term horizons where statistical significance over ML may not be established (Kandananond, 2011). The choice often depends on the specific context, data availability, desired accuracy, robustness requirements, and the forecasting horizon.

The identification and analysis of key features driving electricity demand are critical for model accuracy. For longer-term national forecasts, socio-economic variables are paramount. Kandananond (2011) included population, GDP, stock index, and industrial export revenue. Rahman et al. (2018) used population, GDP, and GDP per capita, finding GDP to be the most influential single predictor for India's TEC.

For STLF, while socio-economic factors have less immediate impact, other variables become crucial. Historical load data itself is a primary input for time-series based models like AR (López et al., 2018) and sequence-learning models like NNs/DNNs (López et al., 2018; Kuo & Huang, 2018). Exogenous variables, particularly meteorological factors like temperature (Kuo & Huang, 2018) and calendar effects (time of day, day of week, holidays, seasons), are widely acknowledged as critical for STLF accuracy (López et al., 2018). Proper selection and processing of these input features are highlighted as key stages in the forecasting process (López et al., 2018).

The reviewed literature presents a range of methodologies, from traditional statistical models (ARIMA, AR, Regression, Smoothing) to various AI approaches (ANN, NN, DNN), applied to both short-term and long-term electricity forecasting challenges in diverse geographical contexts (Thailand, India, Spain). There is a clear trend towards using AI/ML, especially for STLF, due to its ability to model complex, non-linear patterns (López et al., 2018; Kuo & Huang, 2018). However, traditional methods remain relevant, particularly for longer horizons or when parsimony is valued (Kandananond, 2011), and comparative studies show that ML does not always offer a statistically significant advantage (Kandananond, 2011) or may only excel under specific conditions (López et al., 2018). Feature selection is context-dependent, with socio-economic factors dominating longer-term national forecasts and weather/calendar effects being critical for STLF.

A significant gap exists concerning the specific needs of Vaasa. None of the reviewed studies:

- Focus on a Nordic city context like Vaasa, which likely has unique consumption patterns related to climate, heating needs, and industrial profile.
- Simultaneously develop both tailored short-term and long-term models for a specific city.
- Provide a direct comparison of modern ML approaches versus traditional statistical methods using data specifically representative of Vaasa for both time horizons.
- Conduct an in-depth analysis aimed at identifying the key local features driving electricity demand specifically within Vaasa.

The proposed research directly addresses these gaps. By aiming to develop both short-term and long-term models tailored to Vaasa, evaluate ML against traditional methods within this specific context, and identify locally relevant features. This study will provide valuable, actionable insights currently lacking in the reviewed literature.

3 DATA COLLECTION AND PREPARATION

This chapter details the process of acquiring and preparing the necessary data to develop and evaluate the short-term and long-term electricity demand forecasting models for Vaasa, as outlined in the research objectives. The selection of variables was guided by established practices in electricity load forecasting literature (e.g., Kandananond, 2011; López et al., 2018; Rahman et al., 2018) and the specific context of Vaasa.

3.1 Data Sources and Variables

To construct the forecasting models, time series data for electricity demand and several potentially influential exogenous variables were collected. The primary dependent variable and the selected independent variables, along with their sources and temporal resolution, are listed below:

Electricity Demand:

Variable: Historical hourly electricity demand (consumption) for the Vaasa region.

Source: Vaasan Sähkö.

Period: January 1, 2020, to December 31, 2024.

Resolution: Hourly.

Meteorological Data:

Variable: Ambient temperature in Vaasa.

Source: Finnish Meteorological Institute.

Period: January 1, 2020, to December 31, 2024.

Resolution: Hourly (average temperature collected).

Astronomical Data:

Variable: Sunrise and sunset times for Vaasa.

Source: Time and Date AS.

Period: January 1, 2020, to December 31, 2024.

Resolution: Daily.

Demographic Data:

Variable: Population statistics for Vaasa.

Source: Statistics Finland.

Period: January 2020 to December 2024.

Resolution: Monthly.

Economic Data:

Variable: Price of electricity by type of consumer in Finland.

Source: Statistics Finland.

Period: January 2020 to December 2024.

Resolution: Monthly.

Transportation Data:

Variable: Number of electric vehicles (EV) in Ostrobothnia.

Source: Finnish Transport and Communications Agency Traficom.

Period: January 2020 to December 2024.

Resolution: Quarterly.

3.2 Data Preparation and Feature Engineering

Raw data collected from the sources listed above required processing and structuring to be suitable for input into the forecasting models. A custom script developed in Python was used for all data preparation tasks. The key steps involved were:

Dataset Aggregation: The raw hourly data were aggregated to create consistent hourly, daily, and monthly datasets. This facilitates the development of models tailored to different forecasting horizons (short-term typically using hourly/daily data, long-term using monthly data).

Temperature Feature Extraction: From the hourly temperature data, daily and monthly statistics including average, minimum, and maximum temperatures were calculated and included in the respective datasets. These derived features can capture different aspects of temperature's influence on heating and cooling demand.

Sunrise/Sunset Feature Engineering: The daily sunrise and sunset times were transformed into an hourly categorical feature representing distinct periods of natural light availability. This was done using a three-flag system:

Flag 1: Hours between the previous day's sunset and the current day's sunrise (night time).

Flag 2: The first two hours after sunrise and the last two hours before sunset (transition periods).

Flag 3: Remaining daytime hours between the Flag 2 periods. This feature aims to capture variations in electricity usage related to lighting needs and potentially human activity patterns influenced by daylight.

Data Merging and Alignment: All datasets (demand, temperature, sunrise/sunset flags, population, price, electric vehicles) were aligned based on timestamps to create final datasets for model training and evaluation. Handling of different resolutions (e.g., monthly population/price data, quarterly automobile data for hourly/daily models) involved appropriate mapping (e.g., repeating the value for all hours/days within that period).

Handling Missing Data: During the data preparation phase, the datasets were checked for missing values using a Python script. Few missing data points were identified, exclusively within the hourly Vaasa average temperature data obtained from the Finnish Meteorological Institute. These occurrences typically involved only a single missing data point between two available data points. To ensure data continuity, these missing values were imputed using linear interpolation, specifically by calculating the average of the adjacent data points.

This systematic data collection and preparation process resulted in structured datasets ready for the subsequent stages of model development, training, and evaluation, ensuring that relevant temporal patterns and exogenous factors influencing electricity demand in Vaasa are appropriately represented.

4 FEATURE SELECTION ANALYSIS

Prior to model development, an exploratory data analysis was conducted to identify relevant predictor variables and understand their relationship with electricity demand in Vaasa. This involved visual inspection of time series plots and relationship plots (raincloud plots, flex plots), as well as statistical tests including Analysis of Variance (ANOVA) for categorical features and Pearson correlations for continuous features.

4.1 Time Series Analysis

Figure 1 displays the hourly electricity demand (MWh) in Vaasa over the five-year period from January 1, 2020, to December 31, 2024. The plot clearly reveals strong cyclical variations corresponding to seasonal changes, a pattern commonly observed in electricity consumption data, particularly in regions with distinct temperature variations.

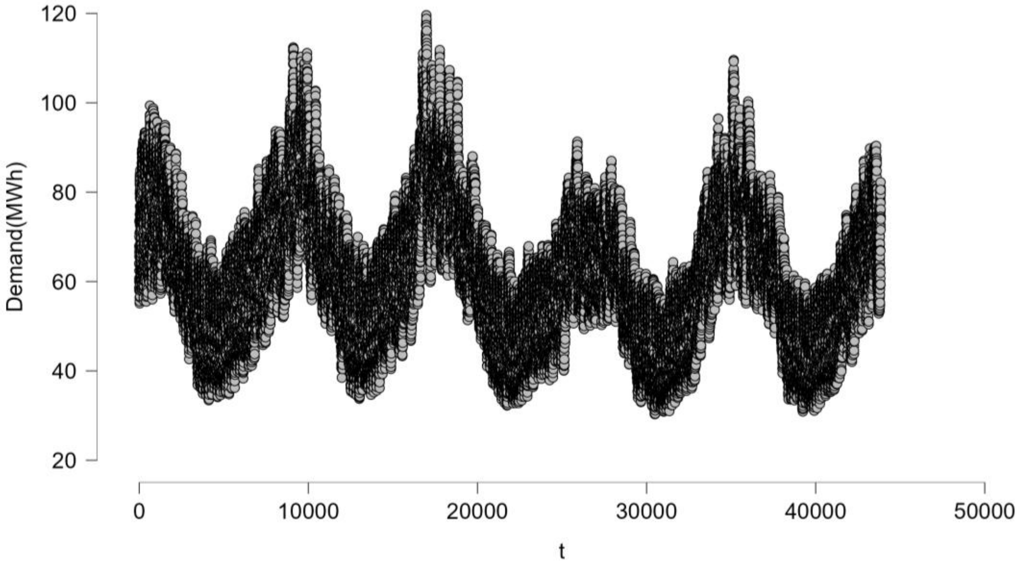


Figure 1. Hourly Electricity Demand Time Series for Vaasa (MWh), 2020-2024.

Higher demand peaks are visible during winter months, likely driven by heating needs, while lower demand troughs occur during summer months. Over this specific five-year timeframe, no prominent long-term upward or downward trend in overall demand is visually apparent. This observation underscores the critical importance of capturing seasonality within the forecasting models.

4.2 Relationship with Categorical Features

Figure 2 illustrates the distribution of hourly electricity demand across the twelve months of the year. It visually confirms the strong seasonality noted in the time series plot.

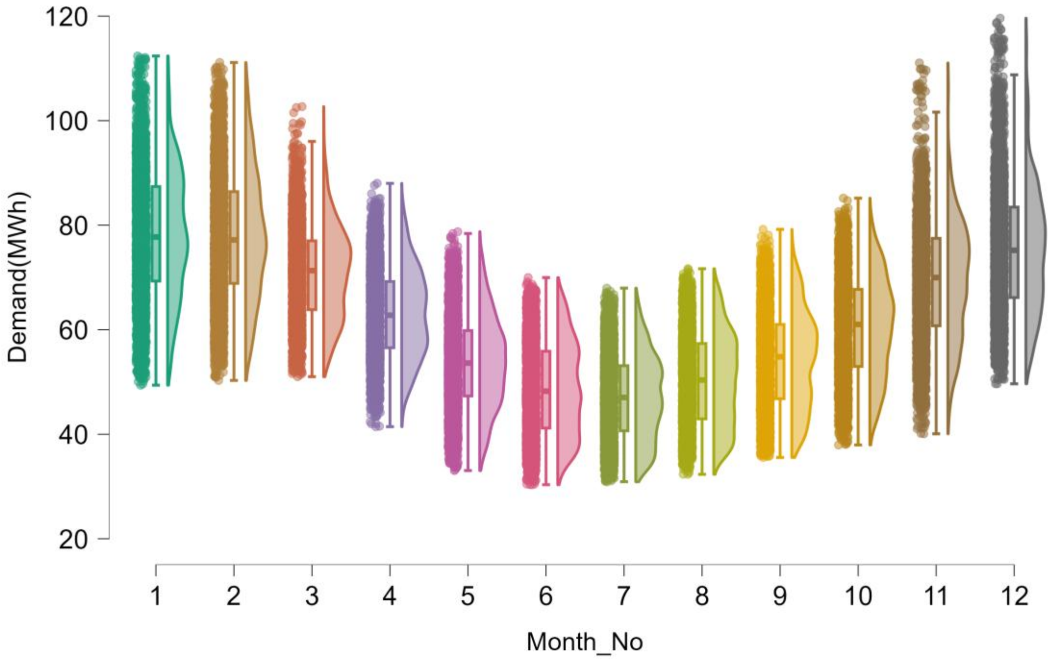


Figure 2. Distribution of Hourly Electricity Demand (MWh) by Month in Vaasa, 2020-2024.

The distributions, median values (indicated by the horizontal line within the 'cloud'), and density shapes (the 'cloud' itself) vary significantly between months. Winter months (e.g., January, February, December) exhibit higher median demand and often wider distributions compared to summer months (e.g., June, July, August), which show lower median demand. This distinct monthly variation confirms that 'Month' is a highly relevant categorical predictor for demand.

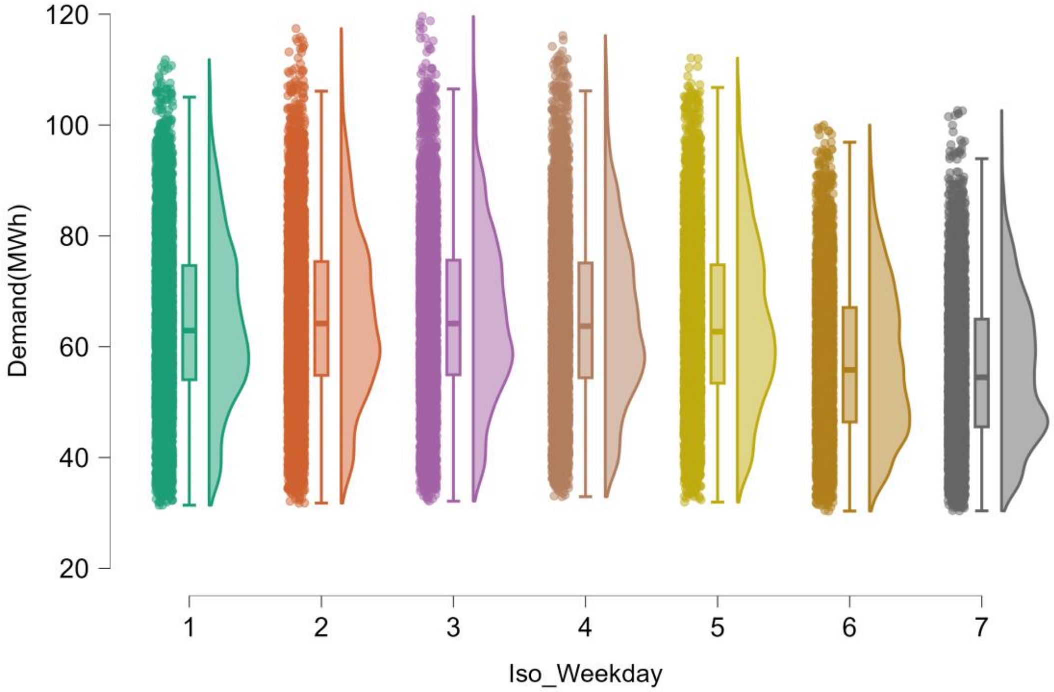


Figure 3. Distribution of Hourly Electricity Demand (MWh) by Day of Week (ISO Standard) in Vaasa, 2020-2024.

Figure 3 presents the distribution of hourly electricity demand across the days of the week, using the ISO weekday standard (1 = Monday, 7 = Sunday). A distinct weekly pattern is evident. Demand characteristics (median, distribution) are broadly similar for working days (Monday to Friday), reflecting regular commercial and societal activity schedules. In contrast, Saturday (6) and especially Sunday (7) show markedly lower median demand and potentially altered distribution shapes, consistent with reduced activity on weekends. This highlights 'Weekday' as another essential categorical feature for capturing the weekly demand cycle.

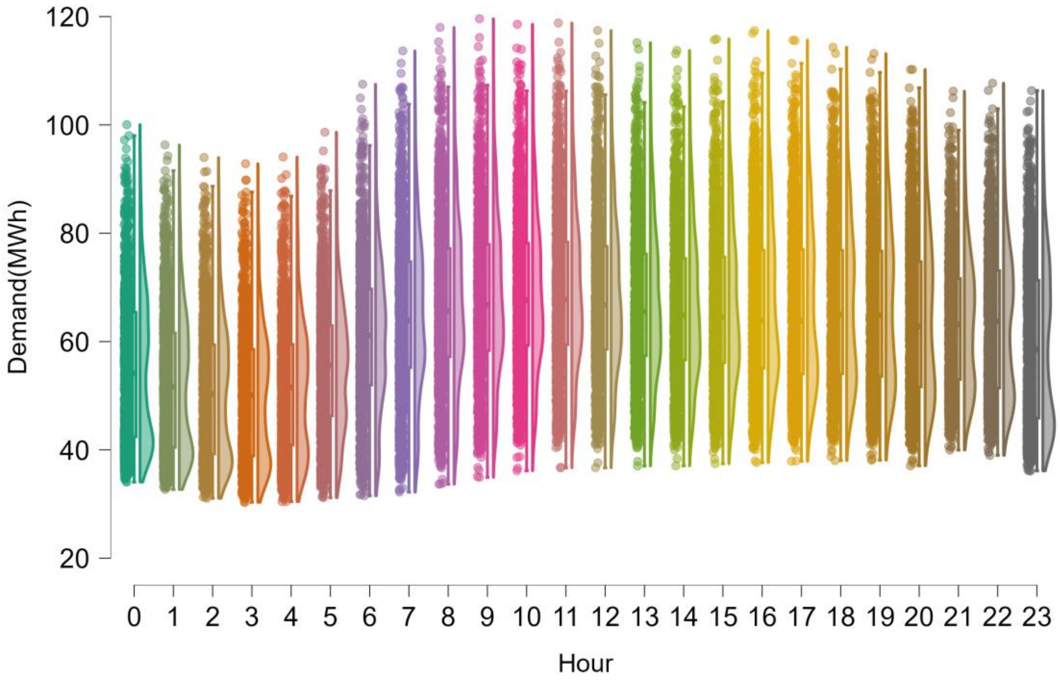


Figure 4. Distribution of Hourly Electricity Demand (MWh) by the Hour of Day in Vaasa, 2020-2024.

Figure 4 illustrates the distribution of hourly electricity demand across the 24 hours of the day in Vaasa (2020-2024). A distinct diurnal pattern is clearly visible. Demand is lowest during the late night and early morning hours (approx. 00:00-05:00), begins to ramp up significantly in the morning (around 06:00-08:00), generally stays high during daytime working hours with potential peaks, and often shows another peak in the evening before declining overnight. The distributions and median demand levels vary substantially between different hours, confirming that the 'Hour' of the day is a critical predictor variable capturing the fundamental daily rhythm of electricity consumption.

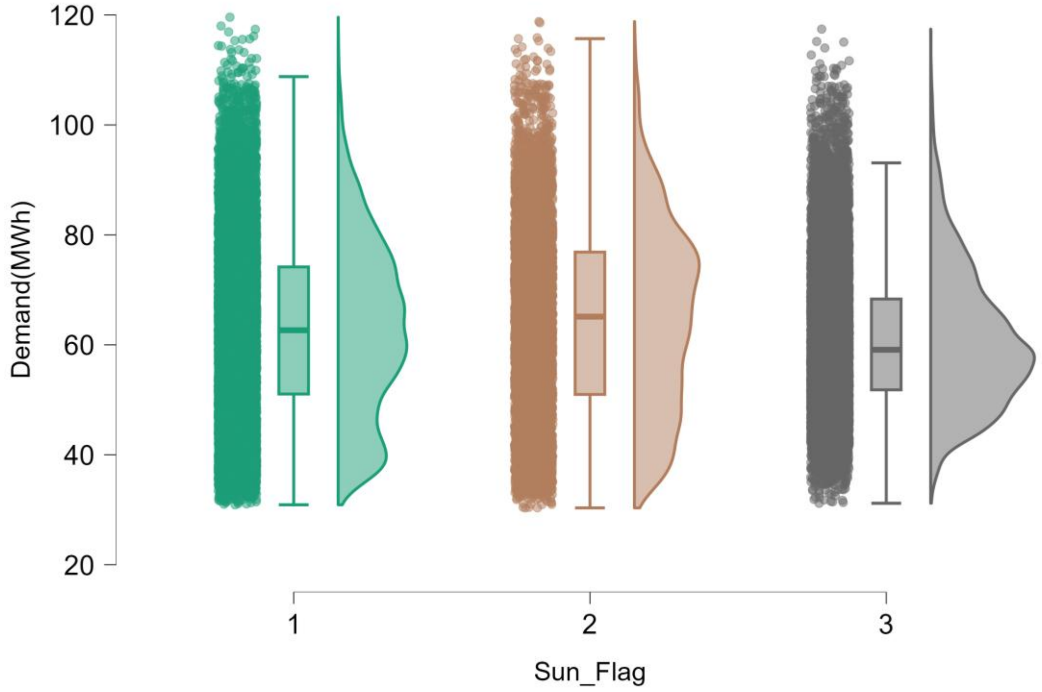


Figure 5. Distribution of Hourly Electricity Demand (MWh) by the engineered Sun_Flag feature in Vaasa, 2020-2024.

Figure 5 displays the distribution of hourly electricity demand based on the engineered 'Sun_Flag' feature (where Flag 1 = night, Flag 2 = sunrise/sunset transition, Flag 3 = full daylight). The plot shows clear differences in demand levels associated with these daylight periods. Demand is generally lowest and has a narrower distribution during the night period (Flag 1). Demand is highest during the transition periods (Flag 2) and the full daylight period (Flag 3), which show similar, higher median demand levels compared to the night. This suggests that the Sun_Flag feature effectively captures distinct demand regimes related to natural light availability and associated activities, making it a potentially valuable categorical predictor.

4.3 Relationship with Continuous Features

Figure 6 explores the relationship between average hourly electricity demand and average hourly temperature. The plot demonstrates a characteristic non-linear relationship, often referred to as U-shaped or V-shaped in load forecasting literature (López et al., 2018).

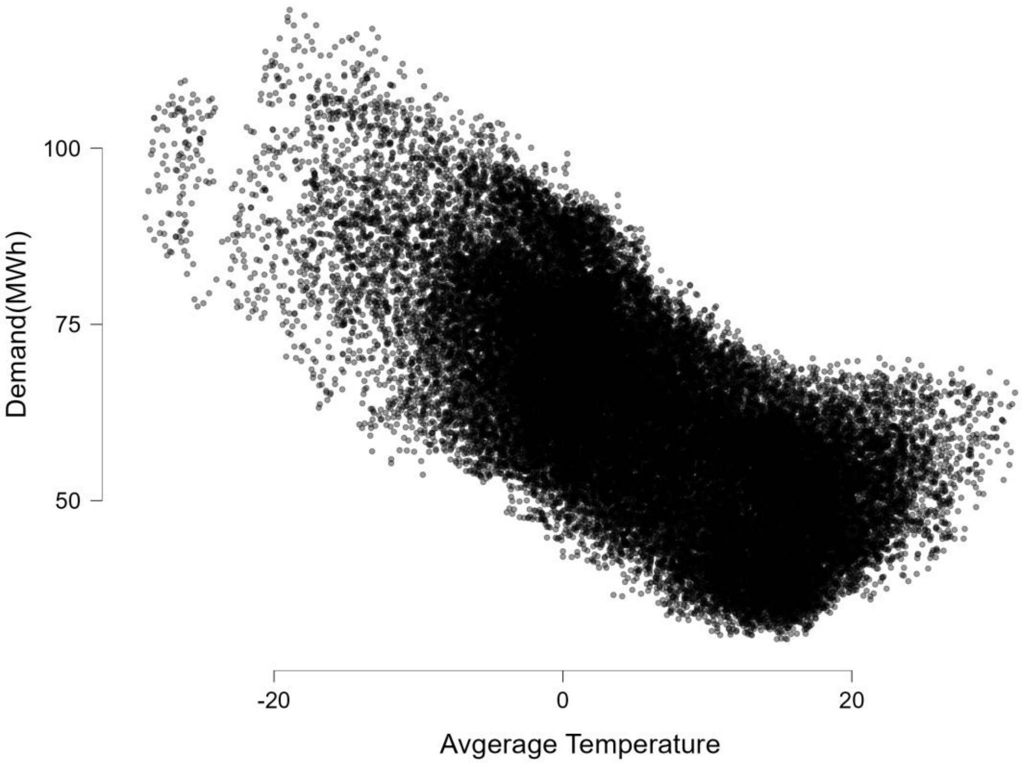


Figure 6. Relationship between Hourly Average Electricity Demand (MWh) and Hourly Average Temperature (°C) in Vaasa, 2020-2024.

Demand is highest at low temperatures (below approximately 0°C), decreases as temperature rises to a moderate range (roughly 5°C to 15°C), and then tends to increase again slightly at higher temperatures (above 20°C). The strong negative correlation at lower temperatures is likely dominated by heating load, while the increase at higher temperatures may reflect cooling demand. This non-linearity confirms temperature as a critical predictor and suggests that models must be capable of capturing this behavior.

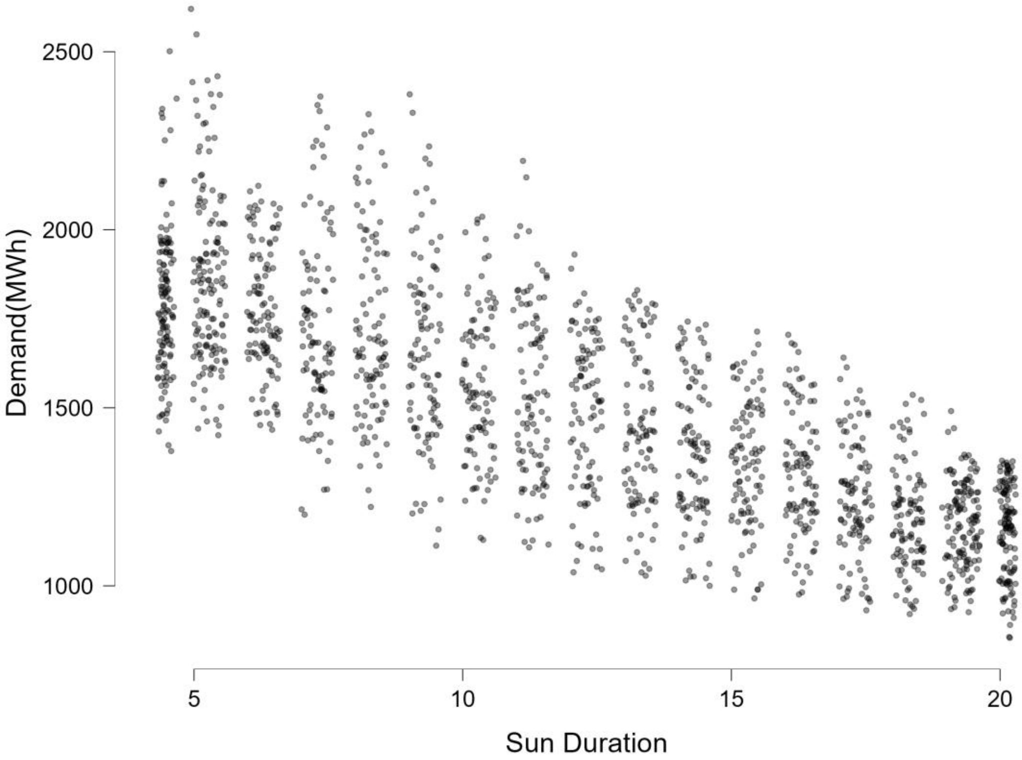


Figure 7. Relationship between Daily Average Electricity Demand (MWh) and Daily Sun Duration (Hours) in Vaasa, 2020-2024.

Figure 7 investigates the relationship between average daily electricity demand and the duration of sunlight. The plot suggests a reasonably clear negative linear trend: as sun duration increases (characteristic of spring and summer months), the average daily electricity demand tends to decrease. This is likely due to factors associated with longer daylight hours, such as reduced heating requirements and potentially less need for artificial lighting. The strength of this relationship will be further quantified by the Pearson correlation analysis.

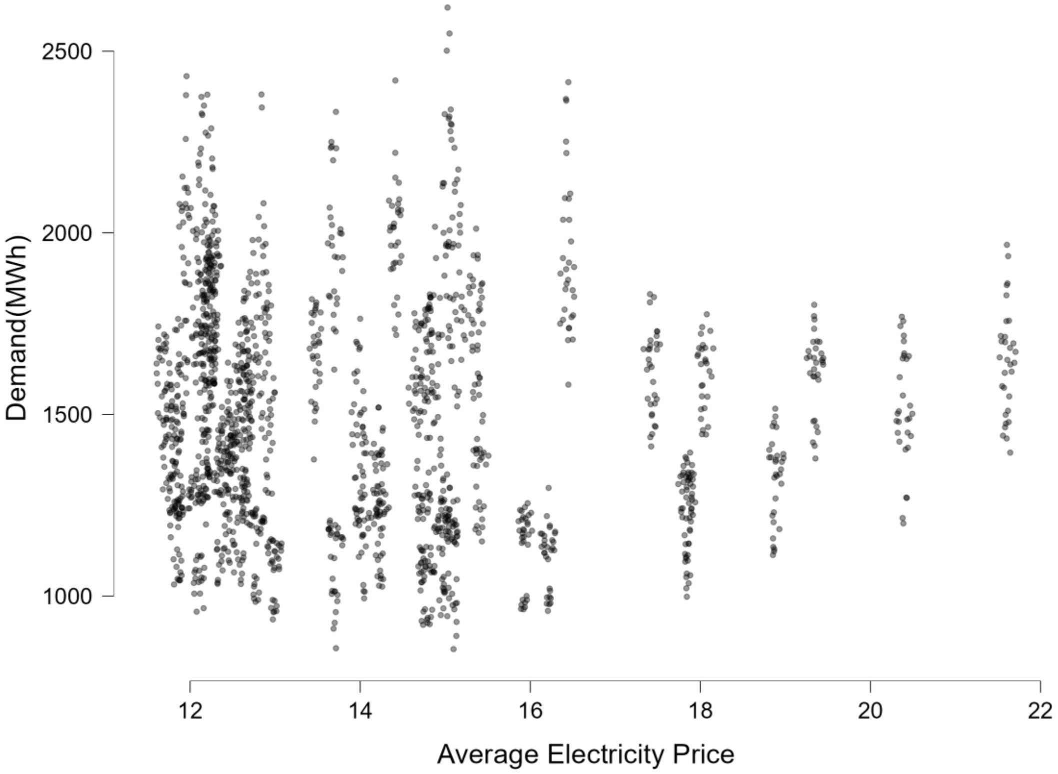


Figure 8. Relationship between Monthly Average Electricity Demand (MWh) and Average Consumer Electricity Price (c/kWh) in Finland, 2020-2024.

Figure 8 examines the relationship between average monthly electricity demand and the average price of electricity for consumers in Finland. Visual inspection of the scatter plot does not reveal an obvious or strong linear correlation between these two variables within the observed range. While economic theory suggests price influences demand, this effect might be weak at this aggregation level, non-linear, subject to time lags, or overshadowed by more dominant factors like weather over the analyzed period.

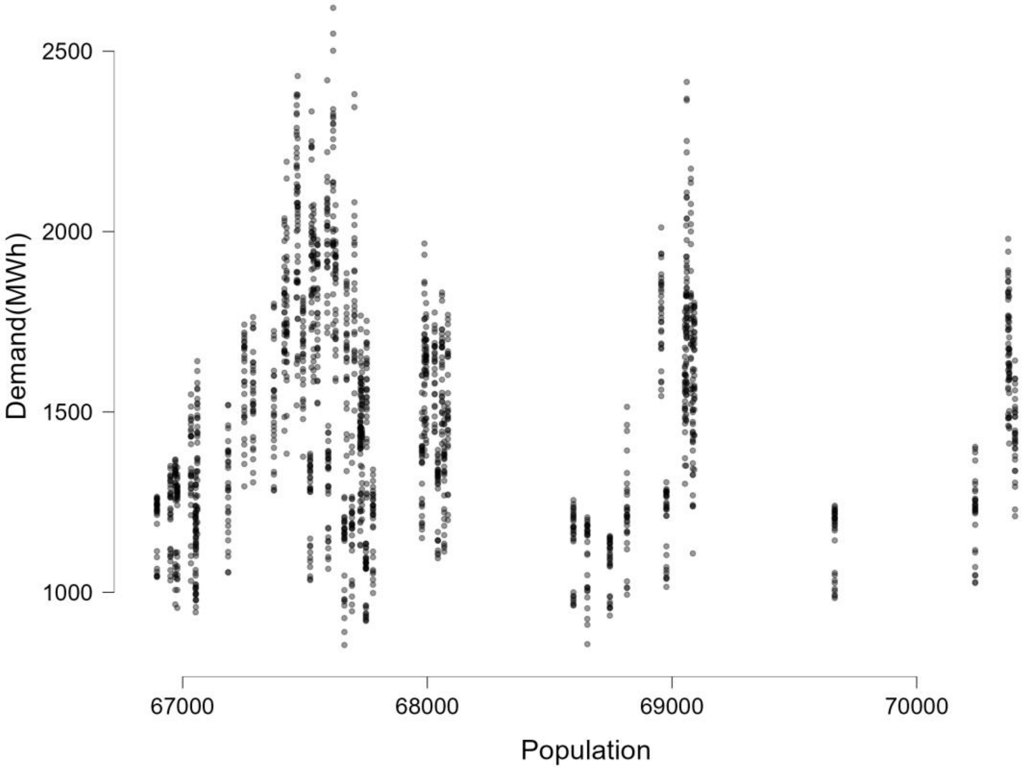


Figure 9. Relationship between Monthly Average Electricity Demand (MWh) and Vaasa Population, 2020-2024.

Figure 9 plots average monthly electricity demand against the monthly population estimate for Vaasa. Similar to electricity price, there is no apparent strong linear relationship visible in the scatter plot over the 2020-2024 timeframe. Monthly population changes in Vaasa during this period appear too gradual to exhibit a clear direct correlation with the more volatile monthly demand fluctuations, which are strongly influenced by seasonality and weather. While population is a known long-term driver (Rahman et al., 2018), its direct predictive power for monthly variations in this dataset seems limited based on visual inspection.

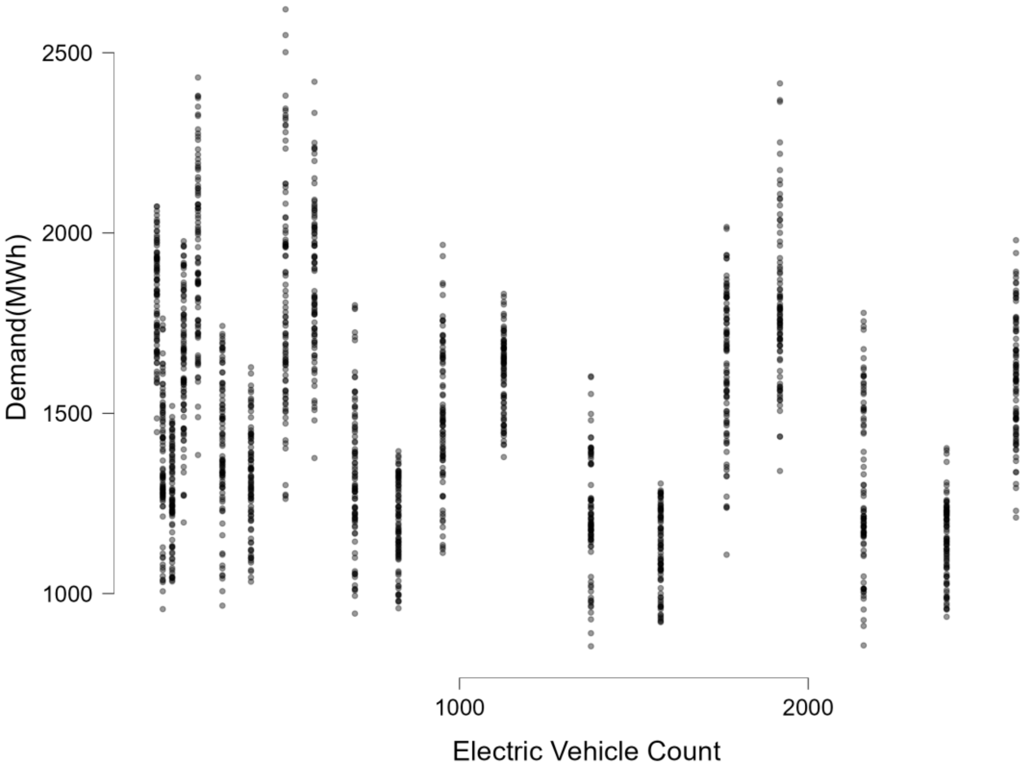


Figure 10. Relationship between Quarterly Average Electricity Demand (MWh) and Electric Vehicle Count in Ostrobothnia, 2020-2024.

Figure 10 explores the potential link between average quarterly electricity demand and the number of registered electric vehicles (EVs) in the Ostrobothnia region. The scatter plot does not show a clear, discernible relationship between these variables during the 2020-2024 period. Although increasing EV adoption is expected to contribute to higher electricity demand, the current penetration level, the quarterly aggregation, or other stronger influencing factors might be masking this effect in the analyzed data. Further statistical analysis (correlation) is needed to detect any potential weaker linear association.

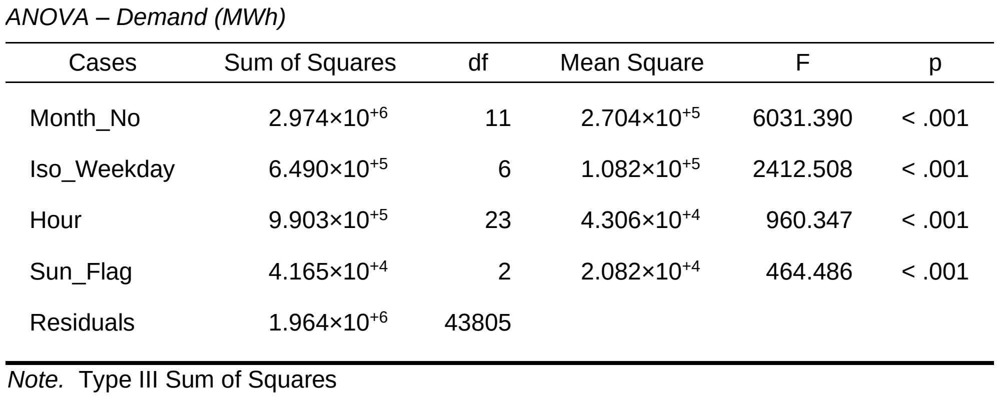
4.4 Statistical Significance Testing

To complement the visual analysis and provide quantitative evidence for feature selection, statistical tests were performed on both the daily and monthly aggregated datasets.

4.4.1 Analysis of Categorical Variables (Hourly Data)

Table 1 presents the results of an ANOVA examining the influence of key categorical variables ('Month_No', 'Iso_Weekday', 'Hour', 'Sun_Flag') on hourly electricity demand (MWh). The results confirm the strong effects observed visually.

Table 1. ANOVA Results: Effects of Month, Iso_Weekday, Hour and Sun_Flag on Daily Electricity Demand (MWh) in Vaasa, 2020-2024.



ANOVA – Demand (MWh)

Cases	Sum of Squares	df	Mean Square	F	p
Month_No	$2.974\times10^{+6}$	11	$2.704\times10^{+5}$	6031.390	< .001
Iso_Weekday	$6.490\times10^{+5}$	6	$1.082\times10^{+5}$	2412.508	< .001
Hour	$9.903\times10^{+5}$	23	$4.306\times10^{+4}$	960.347	< .001
Sun_Flag	$4.165\times10^{+4}$	2	$2.082\times10^{+4}$	464.486	< .001
Residuals	$1.964\times10^{+6}$	43805			

Note. Type III Sum of Squares

Month: Shows a highly statistically significant effect (F(11, 43805) = 6031.39, p < .001), confirming substantial variation in hourly demand across months.

Weekday: Also exhibits a highly statistically significant effect (F(6, 43805) = 2412.508, p < .001), validating the strong weekly pattern in hourly demand.

Hour: The hour of the day demonstrates a highly statistically significant influence (F(23, 43805) = 960.347, p < .001), confirming the distinct diurnal pattern observed in Figure 4.

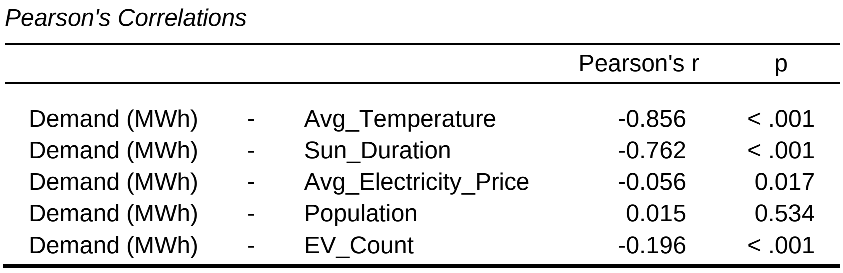
Sun_Flag: The engineered sun flag feature also shows a highly statistically significant effect (F(2, 43805) = 464.486, p < .001), supporting its relevance as seen in Figure 5.

These ANOVA results strongly justify the inclusion of Month, Weekday, Hour, and Sun_Flag as predictors in models using hourly data.

4.4.2 Analysis of Continuous Variables (Daily & Monthly Data)

Table 2 provides Pearson's correlation coefficients (r) quantifying the linear association between daily average electricity demand and key predictor variables aggregated or mapped to a daily level, along with their statistical significance (p-values).

Table 2. Pearson Correlation Coefficients between Daily Average Electricity Demand (MWh) and Daily Predictor Variables in Vaasa, 2020-2024.



Pearson's Correlations

			Pearson's r	p
Demand (MWh)	-	Avg_Temperature	-0.856	< .001
Demand (MWh)	-	Sun_Duration	-0.762	< .001
Demand (MWh)	-	Avg_Electricity_Price	-0.056	0.017
Demand (MWh)	-	Population	0.015	0.534
Demand (MWh)	-	EV_Count	-0.196	< .001

Average Temperature: A very strong, statistically significant negative linear correlation (r = -0.856, p < .001) is confirmed. This aligns with the dominant pattern in Figure 6 (higher demand at colder temperatures). However, the correlation coefficient alone does not capture the non-linear upturn at higher temperatures.

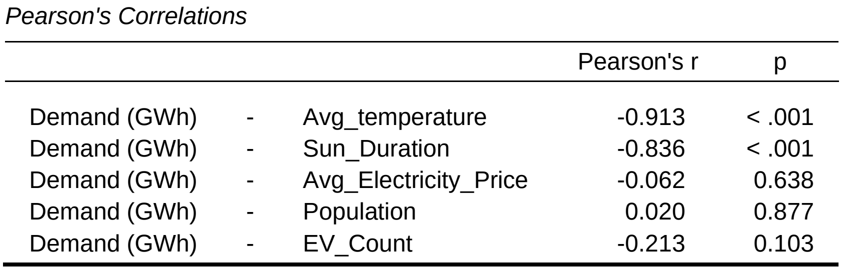
Sun Duration: A strong, statistically significant negative linear correlation (r = -0.762, p < .001) is observed, supporting the visual trend in Figure 7 where longer sun duration corresponds to lower demand.

Avg Electricity Price: Reveals a statistically significant (p = 0.017) but very weak (r = -0.056) negative linear correlation. While statistically significant, the extremely low correlation coefficient suggests price has minimal linear explanatory power for daily demand variations in this dataset.

Population: The correlation is very weak and not statistically significant (r = 0.015, p = 0.534), confirming the lack of a discernible linear relationship seen in Figure 9 over this period.

EV Count: A statistically significant but weak negative correlation (r = -0.196, p < .001) is found. The statistical significance suggests a non-random linear association, but the negative direction is counter-intuitive (more EVs should increase, not decrease, demand). This weak, unexpected correlation might arise from confounding factors within the quarterly aggregated data or other complex dynamics not captured by simple linear correlation. Given its weakness and counter-intuitive sign, the direct predictive value of EV count, based on this analysis, appears questionable for the current models despite its statistical significance.

Table 3. Pearson Correlation Coefficients using Monthly Average Electricity Demand (GWh) and Monthly Predictor Variables in Vaasa 2020-2024.



Pearson's Correlations

			Pearson's r	p
Demand (GWh)	-	Avg_temperature	-0.913	< .001
Demand (GWh)	-	Sun_Duration	-0.836	< .001
Demand (GWh)	-	Avg_Electricity_Price	-0.062	0.638
Demand (GWh)	-	Population	0.020	0.877
Demand (GWh)	-	EV_Count	-0.213	0.103

Table 3 presents Pearson's correlation coefficients calculated using the monthly aggregated dataset, assessing linear relationships with monthly average electricity demand (GWh):

Average Temperature: The correlation remains very strong, negative, and highly significant (r = -0.913, p < .001).

Sun Duration: The strong, negative, and significant correlation persists at the monthly level (r = -0.836, p < .001).

Avg Electricity Price: The correlation is very weak (r = -0.062) and not statistically significant (p = 0.638) at the monthly level. This differs from the daily result and further suggests price is not a strong linear predictor in this context.

Population: Consistent with the daily analysis, the correlation remains very weak and non-significant (r = 0.020, p = 0.877).

EV Count: Also consistent with previous daily findings, the correlation is weak and not statistically significant (r = -0.213, p = 0.103).

4.5 Summary of Feature Selection

Integrating findings from visual inspection (Figures 1-10) and statistical tests (ANOVA Table 1 based on hourly data; Correlation Tables 2 & 3 based on daily and monthly data), the following conclusions guide the feature selection for the Vaasa electricity demand forecasting models.

Strong Predictors:

Categorical: 'Month', 'Weekday', 'Hour', and the engineered 'Sun_Flag' consistently demonstrate strong and statistically significant relationships with hourly electricity demand (ANOVA Table 1). They are essential predictors for capturing seasonality, weekly cycles, diurnal patterns, and daylight effects in short-term model.

Continuous: 'Average Temperature' and 'Sun Duration' show strong, significant correlations with demand at both daily and monthly levels (Tables 2 & 3). They are critical predictors. The non-linear relationship with temperature, evident visually, requires appropriate handling (e.g., non-linear model capability or specific feature engineering).

Weak/Inconsistent Predictors:

'Avg Electricity Price': Showed a statistically significant but practically negligible correlation with daily demand, and no significant correlation with monthly demand. Its predictive value appears limited.

'Population': Displayed no statistically significant linear correlation with demand at either daily or monthly levels during this period.

'EV Count': Exhibited a weak, counter-intuitive negative correlation significant at the daily level but not significant at the monthly level. Its reliability as a predictor is questionable.

Therefore, the primary features selected for model development based on this comprehensive analysis are Month, Weekday, Hour, Sun_Flag, Average Temperature (and derived features like min/max), and Sun Duration. Price, Population, and EV Count will be excluded due to weak, inconsistent, or non-significant findings in this dataset and timeframe.

5 LONG-TERM FORECASTING

A primary objective of this research is the development of a robust long-term electricity demand forecasting model. Such a model is intended to serve as a decision-support tool for relevant authorities, including governmental bodies and electricity generation companies, aiding in the formulation of effective energy policies and the strategic planning of necessary infrastructure to meet future electricity demand.

5.1 Model Fitting using Multiple Linear Regression Approach

The investigation into a long-term electricity predictive model for Vaasa by a MLR approach. This method was implemented using the scikit-learn library in Python. The primary goal is to establish a robust and interpretable baseline for forecasting monthly electricity demand by utilizing essential predictors identified during the feature selection analysis (Chapter 4). The analysis employs monthly data for the period spanning January 2020 to December 2024. Based on prior feature selection, Average Temperature (and its derivatives like minimum and maximum temperature) and Sun Duration were identified as potentially significant predictors.

An initial MLR model was constructed incorporating all identified predictive features: Average Temperature, Minimum Temperature, Maximum Temperature, and Sun Duration. The coefficients for this initial model are presented in Table 4. Observation of Table 4 reveals that the magnitudes of the coefficients for Min_temperature, Max_temperature, and Sun_Duration are relatively small compared to that of Avg_temperature.

Table 4. Multiple Linear Regression Model Coefficients with All Predictive Features

	coef	std err	z	P>|z|	[0.025	0.975]
const	52.4435	3.157	16.614	0.000	46.078	58.810
Avg_temperature	-0.4089	0.394	-1.039	0.305	-1.203	0.385
Min_temperature	-0.2540	0.208	-1.220	0.229	-0.674	0.166
Max_temperature	0.0246	0.280	0.088	0.930	-0.540	0.589
Sun_Duration	-0.0161	0.008	-2.108	0.041	-0.032	-0.001

This observation prompted an iterative process of model refinement, exploring different combinations of these features to identify a more parsimonious and potentially more robust model.

Table 5. Model Diagnostics for Multiple Linear Regression with different Feature Combinations.

Model No	Included Features	MSE	RMSE	MAE	MAPE	R²
1	Avg_temperature, Min_temperature, Max_temperature, Sun_Duration	4.251	2.062	1.598	0.038	0.943
2	Avg_temperature, Min_temperature	4.489	2.119	1.687	0.041	0.940
3	Avg_temperature, Max_temperature	2.770	1.664	1.409	0.033	0.963
4	Avg_temperature, Sun_Duration	3.042	1.744	1.452	0.035	0.959
5	Avg_temperature	3.270	1.808	1.428	0.035	0.956

Table 5 summarizes the diagnostic metrics for five different MLR model configurations, each utilizing a different subset of the initially selected features. The models were evaluated based on Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R²).

Based on the comparative diagnostics in Table 5, Model 3, which includes Avg_temperature and Max_temperature as predictor variables, demonstrates superior performance. It achieved the lowest MSE

(2.770), RMSE (1.664), MAE (1.409), and MAPE (0.033), alongside the highest R² value (0.963). The R² value indicates that approximately 96.3% of the variance in monthly electricity demand can be explained by the variations in average and maximum temperatures within this model. Therefore, Model 3 is selected as the preferred MLR model for further analysis and interpretation.

5.2 Residual Analysis of the Selected Multiple Linear Regression Model

To validate the adequacy of the selected MLR model (Model 3: predictors Avg_temperature, Max_temperature), a thorough analysis of its residuals was conducted.

Figure 11 plots the standardized residuals against the fitted (predicted) values. Ideally, this plot should exhibit a random scatter of points around the horizontal line at zero, with no discernible patterns (e.g., funnel shape, curvature).

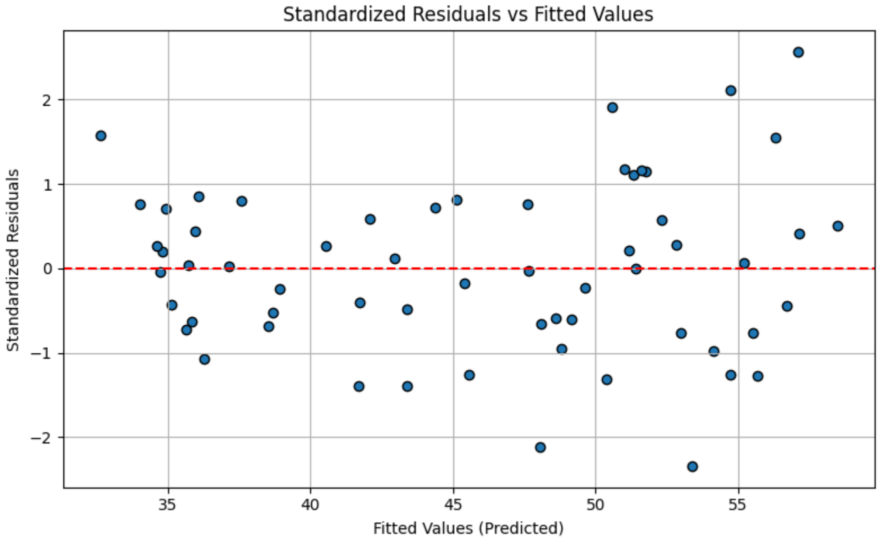


Figure 11. Standardized Residuals versus Fitted (Predicted) Values from Multiple Linear Regression Model 3.

The plot in Figure 11 shows a relatively random distribution of residuals, suggesting that the assumption of homoscedasticity (constant variance of errors) is reasonably met and that the model's linearity is appropriate. There are no obvious systematic trends in the residuals as the fitted values change. Some individual points lie further from zero (e.g., around +2 and -2 standard deviations), indicating specific months where the model's prediction was less accurate, but these do not appear to form a systematic pattern.

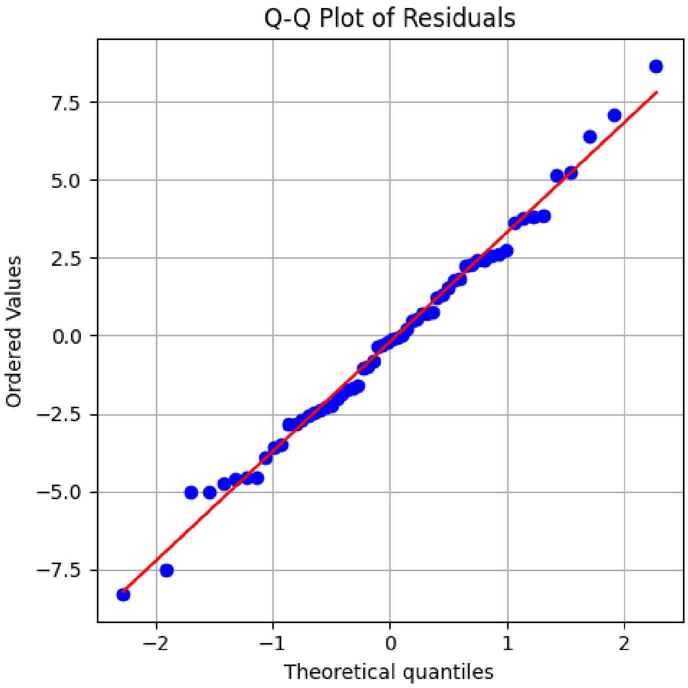


Figure 12. Quantile-Quantile (Q-Q) Plot of the Standardized Residuals from Multiple Linear Regression Model 3.

Figure 12 presents the Quantile-Quantile (Q-Q) plot of the residuals, which is used to assess the normality of the error terms. If the residuals are normally distributed, the points should fall approximately along the red diagonal reference line. The plot shows that the points generally align closely with the reference line, especially in the central region. There are slight deviations at the extreme tails, with the observed quantiles being slightly more extreme than the theoretical quantiles, suggesting that the distribution of residuals might have slightly heavier tails than a perfect normal distribution. However, the overall adherence to the line indicates that the normality assumption is largely satisfied, which is important for the validity of statistical inference related to the model.

The residual analysis suggests that the selected MLR model (Model 3) is reasonably well-specified, with no major violations of the underlying assumptions of linearity, homoscedasticity, and normality of errors.

5.3 Interpreting the Selected Multiple Linear Regression Model

The selected MLR model utilizes Avg_temperature (°C) and Max_temperature (°C) to predict monthly electricity demand. The estimated coefficients for this model are presented in Table 6.

Table 6. Selected Multiple Linear Regression Model's Coefficients.

	coef	std err	z	P>|z|	[0.025	0.975]
const	53.7245	2.352	22.845	0.000	48.988	58.461
Avg_temperature	-0.6754	0.265	-2.549	0.014	-1.209	-0.142
Max_temperature	-0.2549	0.236	-1.079	0.286	-0.731	0.221

The MLR equation can be expressed as:

$$\begin{array}{c}\text{Demand}\\\text{GWh}\end{array} = 53.7245 - 0.6754\,\text{X}\;\begin{array}{c}\text{Average}\\\text{Temperature}\end{array} - 0.2549\,\text{X}\;\begin{array}{c}\text{Maximum}\\\text{Temperature}\end{array}$$

Intercept (53.7245): The intercept represents the estimated average monthly electricity demand in GWh when both Avg_temperature and Max_temperature are equal to zero. While a temperature of zero degrees Celsius is plausible, this interpretation should be made with caution, as it may represent an extrapolation beyond the typical range of observed temperature data. It primarily serves as a baseline for the model.

Avg_temperature: The coefficient for Avg_temperature is -0.6754. This indicates that for each one-unit (degree Celsius) increase in the average monthly temperature, the monthly electricity demand is predicted to decrease by approximately 0.6754 GWh, holding Max_temperature constant. This negative relationship is expected in Vaasa's climate, where higher average temperatures likely reduce heating requirements, a dominant factor in electricity consumption.

Max_temperature: The coefficient for Max_temperature is -0.2549. This suggests that for each one-unit (degree Celsius) increase in the maximum monthly temperature, the monthly electricity demand is predicted to decrease by approximately 0.2549 GWh, holding Avg_temperature constant. This might seem counterintuitive if one expects higher maximum temperatures to increase cooling demand. However, in the context of a multiple regression model, this coefficient represents the effect of maximum temperature after accounting for average temperature. It's possible that months with a high maximum temperature, given a certain average, might correspond to specific weather patterns (e.g., clear skies, less overall energy needed for prolonged heating/lighting even if peak cooling is slightly up) or that the effect of average temperature already captures the primary temperature-driven load. The negative sign indicates that, in this specific model configuration, higher maximums (once average temperature is controlled for) are associated with slightly lower demand. This could also be due to multicollinearity if average and maximum temperatures are highly correlated.

5.4 Model Performance and Forecast Visualization

Figure 13 provides a visual representation of the selected MLR model's performance. It plots the actual training data (January 2020 - December 2023), the model's fitted values on this training data, the actual test data (January 2024 - December 2024), and the model's forecasts for this test period.

The fitted values (dashed green line) closely track the actual training data (solid blue line). The model effectively captures the strong seasonal pattern inherent in the monthly electricity demand, including the peaks (typically in winter) and troughs (typically in summer).

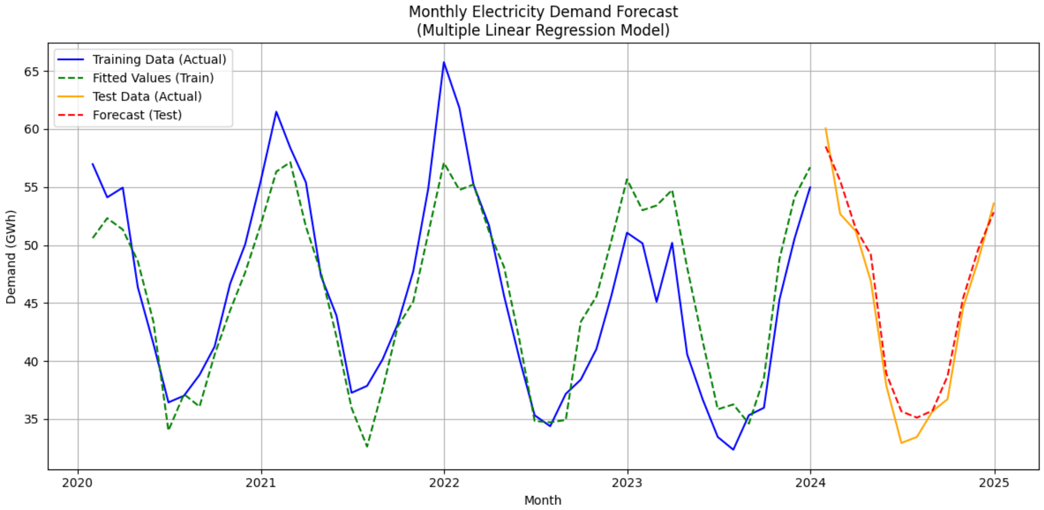


Figure 13. Actual Monthly Electricity Data and Forecasted Monthly Electricity (GWh) for 2020-2024 using Multiple Linear Regression Model.

The forecasts for the 2024 test period (dashed red line) also follow the expected seasonal path. When compared to the actual test data (solid orange line), the model demonstrates a good predictive capability, capturing the general trend and magnitude of demand. While not a perfect match for every month, the forecasts are consistently close to the actual values, indicating the model's ability to generalize to unseen data.

The high R² value of 0.963, coupled with the visual evidence from Figure 13 and the satisfactory residual diagnostics, suggests that this MLR model, despite its simplicity, provides a strong and reliable baseline for understanding and forecasting monthly electricity demand in Vaasa based on average and maximum temperatures. While it may not capture all the nuances of time series dynamics (like autocorrelation, which SARIMA models attempt to address), its interpretability and good predictive performance on this dataset make it a valuable tool.

6 SHORT-TERM (DAILY) LOAD FORECASTING

Having established a foundation in long-term monthly electricity demand forecasting, this section marks a transition to the more granular realm of short-term load forecasting (STLF). The subsequent sections will detail the development and evaluation of models for these shorter forecast horizons, initially employing a MLR approach, to be followed by an exploration of machine learning methodologies.

In machine learning methodologies two widely adopted machine learning algorithms, Random Forest Regressor and XGBoost (Extreme Gradient Boosting) Regressor, are employed to model and forecast daily electricity demand. The selection of these models is motivated by their proven efficacy in handling complex regression tasks and their ability to capture non-linear interactions between features.

6.1 Daily Electricity Demand Forecasting – Multiple Linear Regression Approach

Given the reasonable performance of the MLR model in the context of monthly demand forecasting (as discussed in Chapter 5), this approach is adopted as the initial methodology for daily demand forecasting. The models are implemented using the scikit-learn library in Python, a widely recognized toolkit for machine learning and statistical modeling.

The foundation for feature inclusion in the daily demand models is derived from the Feature Selection Analysis conducted in Chapter 4. This analysis identified several continuous and categorical features as significantly correlated with daily electricity demand in Vaasa:

- Continuous Features: Average Temperature (and its derivatives, minimum and maximum temperature) and Sun Duration.

- Categorical Features: Month and Day of the Week.

6.1.1 Feature Engineering and Transformations

To prepare the identified features for inclusion in the MLR models, specific transformations were applied:

- Temperature and Sun Duration: These continuous variables (Average Temperature, Minimum Temperature, Maximum Temperature, Sun Duration) were initially considered in their original scale.
- Month: The categorical feature Month was transformed using one-hot encoding. This process creates binary (0 or 1) columns for each month, with one month typically omitted to serve as a reference category and avoid perfect multicollinearity.
- Seasonality (Alternative to Month): An alternative approach to capturing monthly variations involved categorizing months into meteorological seasons:
  - Winter: December, January, February
  - Spring: March, April, May
  - Summer: June, July, August
  - Autumn: September, October, November

  This Season variable was also one-hot encoded for inclusion in one of the model iterations.

- Day of the Week: Insights from the exploratory data analysis, specifically the raincloud plot presented in Figure 3 (Distribution of Hourly Electricity Demand (MWh) by Day of Week), indicated a distinct pattern, relatively consistent and higher demand from Monday to Friday, and a noticeably lower, similar level of demand during the weekend (Saturday and Sunday). Consequently, the Day of the Week feature was engineered into a binary variable, Is_workday, taking a value of 1 for weekdays (Monday-Friday) and 0 for weekend days (Saturday-Sunday).

6.1.2 Multiple Linear Regression Model Development

An iterative approach was adopted for developing the daily MLR model. This involved fitting several model configurations and evaluating their performance to arrive at the most suitable specification.

Model 1: Comprehensive Feature Set

The initial model incorporated all identified and transformed predictive features: Avg_temperature, Min_temperature, Max_temperature, Sun_Duration, Is_workday, and the one-hot encoded Month variables (Month_No_2 through Month_No_12, with January as the reference). The estimated coefficients for this Model 1 are presented in Table 7.

Table 7. Multiple Linear Regression Model 1 Coefficients for Daily Forecast (All Predictive Features).

	coef	std err	z	P>|z|	[0.025	0.975]
const	1760.7751	26.886	65.491	0.000	1708.036	1813.514
Avg_temperature	-27.6339	5.604	-4.931	0.000	-38.627	-16.641
Min_temperature	-0.2697	2.674	-0.101	0.920	-5.514	4.975
Max_temperature	9.7048	3.278	2.960	0.003	3.274	16.136
Sun_Duration	-18.4661	4.075	-4.532	0.000	-26.459	-10.473
Is_workday	207.4446	6.931	29.932	0.000	193.850	221.040
Month_No_2	45.2110	19.473	2.322	0.020	7.013	83.409
Month_No_3	-34.0761	28.095	-1.213	0.225	-89.187	21.034
Month_No_4	-120.4805	38.946	-3.094	0.002	-196.877	-44.083
Month_No_5	-145.4304	49.871	-2.916	0.004	-243.258	-47.603
Month_No_6	-95.5699	57.344	-1.667	0.096	-208.056	16.916
Month_No_7	-137.7661	52.985	-2.600	0.009	-241.702	-33.831
Month_No_8	-138.4363	42.262	-3.276	0.001	-221.339	-55.534
Month_No_9	-178.9472	31.279	-5.721	0.000	-240.305	-117.589
Month_No_10	-184.8981	21.478	-8.609	0.000	-227.030	-142.767
Month_No_11	-118.7138	15.699	-7.562	0.000	-149.510	-87.918
Month_No_12	-90.1292	15.792	-5.707	0.000	-121.107	-59.151

An examination of the coefficients and their statistical significance (P>|z|) in Table 7 revealed that Min_temperature was the most statistically insignificant continuous feature (P>|z| = 0.920). This suggests that, in the presence of Avg_temperature and Max_temperature, Min_temperature does not provide additional significant explanatory power for daily electricity demand.

Model 2: Excluding Minimum Temperature

Based on the findings from Model 1, a second model (Model 2) was fitted, excluding Min_temperature but retaining all other features: Avg_temperature, Max_temperature, Sun_Duration, Is_workday, and the one-hot encoded Month variables. The coefficients for Model 2 are detailed in Table 8.

Table 8. Multiple Linear Regression Model 2 Coefficients for Daily Forecast (Excluding Minimum Temperature).

	coef	std err	z	P>|z|	[0.025	0.975]
const	1760.9466	26.823	65.651	0.000	1708.331	1813.562
Avg_temperature	-28.1532	2.212	-12.727	0.000	-32.492	-23.814
Max_temperature	9.9544	2.149	4.632	0.000	5.739	14.170
Sun_Duration	-18.4671	4.073	-4.534	0.000	-26.457	-10.477
Is_workday	207.4241	6.925	29.952	0.000	193.840	221.008
Month_No_2	45.2471	19.463	2.325	0.020	7.068	83.426
Month_No_3	-34.0383	28.082	-1.212	0.226	-89.125	21.048
Month_No_4	-120.5213	38.931	-3.096	0.002	-196.888	-44.155
Month_No_5	-145.3946	49.853	-2.916	0.004	-243.186	-47.603
Month_No_6	-95.4437	57.310	-1.665	0.096	-207.864	16.976
Month_No_7	-137.7091	52.964	-2.600	0.009	-241.603	-33.815
Month_No_8	-138.4544	42.248	-3.277	0.001	-221.328	-55.581
Month_No_9	-178.9818	31.267	-5.724	0.000	-240.315	-117.649
Month_No_10	-184.9291	21.469	-8.614	0.000	-227.042	-142.816
Month_No_11	-118.7266	15.693	-7.565	0.000	-149.511	-87.942
Month_No_12	-90.1244	15.787	-5.709	0.000	-121.092	-59.157

While Model 2 offers a more parsimonious representation by removing an insignificant variable, an inspection of its coefficients (Table 8) indicated that some of the one-hot encoded month variables, specifically Month_No_3 (P>|z| = 0.226) and Month_No_6 (P>|z| = 0.096), still exhibited relatively high p-values, suggesting potential statistical insignificance at conventional levels (e.g., α=0.05).

Model 3: Incorporating Season Instead of Month

To address the complexity and potential insignificance of individual month dummy variables, and to explore a more aggregated representation of seasonality, a third model (Model 3) was developed. This model replaced the one-hot encoded Month variables with one-hot encoded Season variables (Season_spring, Season_summer, Season_winter, with Autumn as the reference category). The other features included were Avg_temperature, Max_temperature, Sun_Duration, and Is_workday. The coefficients for Model 3 are shown in Table 9.

Table 9. Multiple Linear Regression Model 3 Coefficients for Daily Forecast (Season Included Instead of Month).

	coef	std err	z	P>|z|	[0.025	0.975]
const	1619.8929	16.907	95.810	0.000	1586.728	1653.058
Avg_temperature	-29.2460	2.287	-12.787	0.000	-33.732	-24.760
Max_temperature	8.5630	2.232	3.836	0.000	4.184	12.942
Sun_Duration	-18.7348	1.772	-10.571	0.000	-22.211	-15.258
Is_workday	208.3458	7.265	28.679	0.000	194.095	222.596
Season_spring	58.2975	13.659	4.268	0.000	31.503	85.092
Season_summer	67.4203	15.336	4.396	0.000	37.337	97.504
Season_winter	120.4364	11.259	10.697	0.000	98.350	142.523

In Model 3, all included features, including the seasonal dummies, demonstrated high statistical significance (P>|z| = 0.000 for all). This suggests that the seasonal aggregation provides a robust representation of broad seasonal effects.

6.1.3 Model Evaluation and Selection

To determine the optimal MLR model for daily electricity demand forecasting, the three developed models were compared based on several standard performance metrics: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2). These metrics were calculated on the test dataset (year 2024). The comparative diagnostics are presented in Table 10.

Table 10. Model Diagnostics for Multiple Linear Regression with different Feature Combinations (Daily Forecast)

Model No	MSE	RMSE	MAE	MAPE	R^2
1	7055.743	83.998	65.145	0.047	0.928
2	7062.158	84.037	65.178	0.047	0.928
3	7762.966	88.108	66.537	0.049	0.921

Analyzing the performance metrics in Table 10, Model 1 (utilizing Avg_temperature, Min_temperature, Max_temperature, Sun_Duration, Is_workday, and one-hot encoded Month) exhibits the lowest MSE (7055.743), RMSE (83.998), and MAE (65.145), along with the highest R^2 value (0.928). Model 2 (which excluded Min_temperature) performs very similarly to Model 1, with only marginal differences in these metrics. Model 3 (which used aggregated Season dummies instead of individual Month dummies) shows slightly worse performance across all metrics compared to Models 1 and 2 (e.g., higher RMSE of 88.108 and lower R^2 of 0.921).

While Model 2 is more parsimonious than Model 1 by excluding the statistically insignificant Min_temperature variable, Model 1 demonstrates a marginally superior fit according to the error metrics and R^2. Given that the primary goal is predictive accuracy for short-term forecasting, and the difference in complexity between Model 1 and Model

2 is minimal (one variable), Model 1 is selected as the best MLR model for daily electricity demand forecasting in Vaasa based on the model diagnostics. It explains approximately 92.8% of the variance in daily electricity demand and provides the lowest forecast errors on the test set among the multiple linear models considered.

6.1.4 Residual Analysis of the Selected Multiple Linear Regression Model

Following the model selection process outlined in section 6.1.2, MLR Model 1 (which incorporates Avg_temperature, Min_temperature, Max_temperature, Sun_Duration, Is_workday, and one-hot encoded Month variables) was chosen for further evaluation based on its superior performance metrics. To validate the assumptions of this MLR model and assess its adequacy, a thorough analysis of its residuals was conducted.

Figure 14 presents a scatter plot of the standardized residuals against the fitted (predicted) values of daily electricity demand. Ideally, this plot should exhibit a random scatter of points horizontally dispersed around the zero line, with no discernible systematic patterns (such as a curve or a funnel shape).

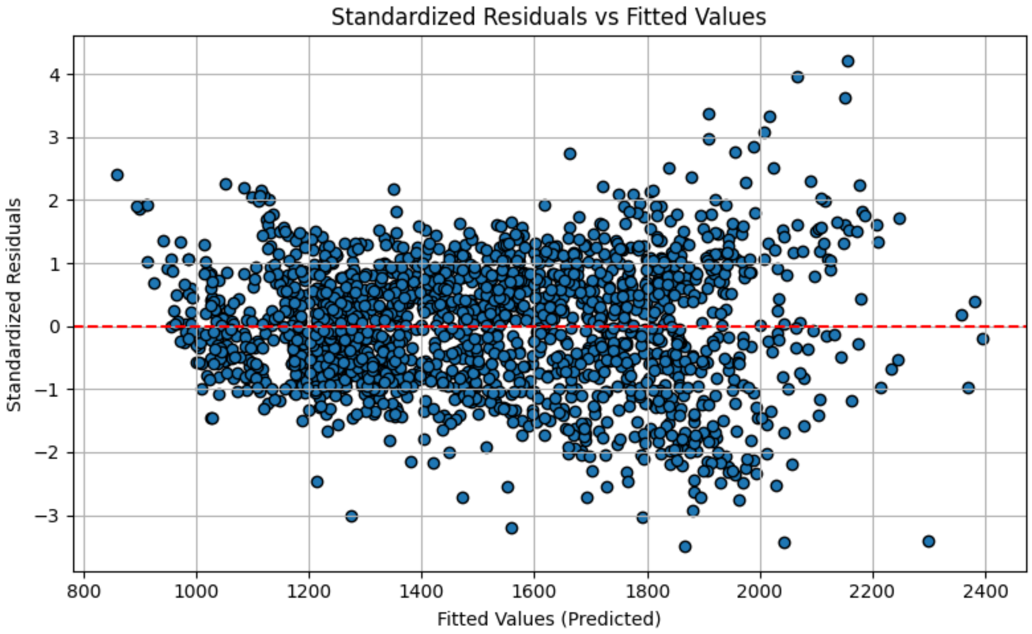


Figure 14. Standardized Residuals versus Fitted Values for the Selected Daily Multiple Linear Regression Model (Model 1).

The observed plot shows that while many points are clustered around the zero line, there is some evidence of heteroscedasticity, particularly as the fitted values increase. For instance, the spread of residuals appears slightly wider for higher fitted values (e.g., above 1800 MWh) compared to lower fitted values. There are also several points that lie beyond ±2 standard deviations, and a few beyond ±3, indicating the presence of outliers or days where the model's predictions were less accurate. However, there is no strong evidence of a non-linear pattern, suggesting the linearity assumption is reasonably met.

Figure 15 displays the Q-Q plot of the residuals. This plot compares the quantiles of the standardized residuals to the theoretical quantiles of a standard normal distribution.

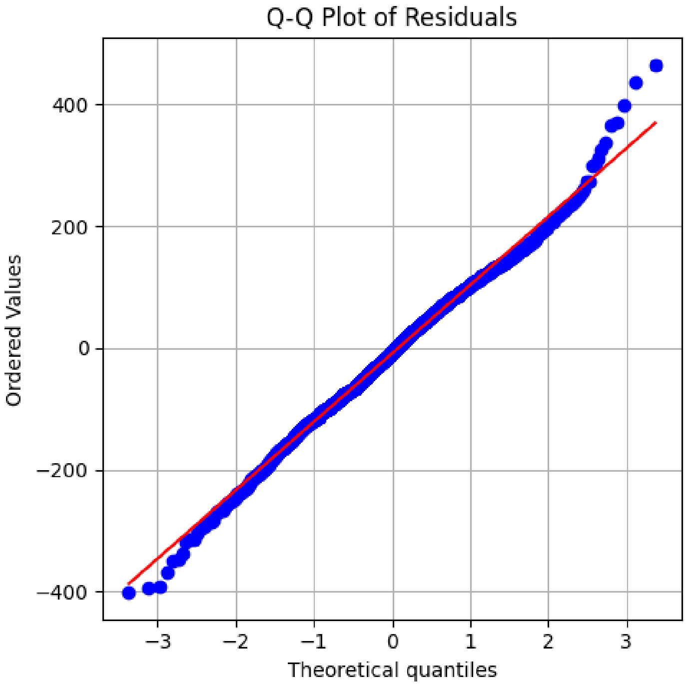


Figure 15. Normal Q-Q Plot of Residuals for the Selected Daily Multiple Linear Regression Model (Model 1).

If the residuals are perfectly normally distributed, all points should fall along the red diagonal reference line. The plot shows that the points generally follow the reference line, especially in the central region. However, there are deviations at the tails: the lower tail of the residuals extends slightly further to the left (more negative) and the upper tail extends slightly further to the right (more positive) than would be expected under a perfect normal distribution. This suggests that the distribution of residuals may be slightly leptokurtic (heavier tails) than a Gaussian distribution. Overall, while not perfectly normal, the deviation is not extreme, and the assumption of normality for the residuals is reasonably approximated for practical purposes in MLR.

6.1.5 Interpreting the Selected Multiple Linear Regression Model

The selected MLR model (Model 1) aims to predict daily electricity demand using a comprehensive set of features identified earlier. The estimated coefficients, standard errors, Z-statistics, and p-values for this model were presented in Table 7 (6.1.2).

The coefficients from Table 7 indicate the estimated change in daily electricity demand (in MWh) for a one-unit change in each predictor variable, holding all other variables constant.

Constant (Intercept): The intercept is 1760.7751 MWh. This is the predicted daily electricity demand when all continuous predictor variables are zero, and all categorical reference levels are active (i.e., for January, on a weekend). However, interpreting the intercept directly is often not meaningful if a value of zero for all predictors is outside the observed range of data or practically impossible.

Avg_temperature: A coefficient of -27.6339 (P>|z| = 0.000) indicates that, on average, for each one-degree Celsius increase in average daily temperature, the daily electricity demand is predicted to decrease by approximately 27.63 MWh, holding other factors constant. This strong negative relationship is expected, reflecting reduced heating needs as average temperatures rise.

Min_temperature: The coefficient is -0.2697 (P>|z| = 0.920). This variable is not statistically significant, suggesting that after accounting for average and maximum temperature, minimum temperature does not have a discernible independent linear impact on daily demand in this model.

Max_temperature: A coefficient of 9.7048 (P>|z| = 0.003) implies that for each one-degree Celsius increase in maximum daily temperature, demand is predicted to increase by about 9.70 MWh, ceteris paribus. This likely reflects increased demand for cooling or other activities associated with higher peak temperatures.

Sun_Duration: The coefficient is -18.4661 (P>|z| = 0.000). This suggests that for each additional hour of sun duration, daily electricity demand is predicted to decrease by approximately 18.47 MWh, holding other variables constant. This could be due to increased solar gain reducing heating needs, or increased natural lighting reducing artificial lighting needs.

Is_workday: The coefficient is 207.4446 (P>|z| = 0.000). This indicates that, on average, a workday is associated with approximately 207.44 MWh higher electricity demand compared to a weekend day, holding all other factors constant. This reflects the increased commercial and industrial activity on weekdays.

Month Variables (One-Hot Encoded, with January as the reference category):

- Month_No_2 (February): Demand is predicted to be 45.21 MWh higher than in January (P>|z| = 0.020), ceteris paribus.
- Month_No_3 (March): Demand is predicted to be 34.08 MWh lower than in January (P>|z| = 0.225, not statistically significant at α=0.05).
- Month_No_4 (April): Demand is predicted to be 120.48 MWh lower than in January (P>|z| = 0.002).
- Month_No_5 (May): Demand is predicted to be 145.43 MWh lower than in January (P>|z| = 0.004).
- Month_No_6 (June): Demand is predicted to be 95.57 MWh lower than in January (P>|z| = 0.096, borderline significant).
- Month_No_7 (July): Demand is predicted to be 137.77 MWh lower than in January (P>|z| = 0.009).

- Month_No_8 (August): Demand is predicted to be 138.44 MWh lower than in January (P>|z| = 0.001).
- Month_No_9 (September): Demand is predicted to be 178.95 MWh lower than in January (P>|z| = 0.000).
- Month_No_10 (October): Demand is predicted to be 184.90 MWh lower than in January (P>|z| = 0.000).
- Month_No_11 (November): Demand is predicted to be 118.71 MWh lower than in January (P>|z| = 0.000).
- Month_No_12 (December): Demand is predicted to be 90.13 MWh lower than in January (P>|z| = 0.000).

These coefficients collectively capture the seasonal pattern of demand relative to January, with lower demand generally observed in spring, summer, and autumn months.

6.1.6 Model Performance and Forecast Visualization

The performance of the selected MLR model (Model 1) on unseen data is crucial for assessing its practical utility. Figure 16 visualizes the model's forecasts for the test period (January 2024 to December 2024) against the actual observed daily electricity demand.

The plot shows that the model's forecasts (dashed red line) track the actual daily electricity demand (solid yellow line) for 2024 reasonably well. The model captures the overall seasonal trend, with higher demand in winter and lower demand in summer, as well as the weekly fluctuations (higher demand on workdays, lower on weekends, though this is not explicitly detailed in this specific plot, it's an underlying component of the model).

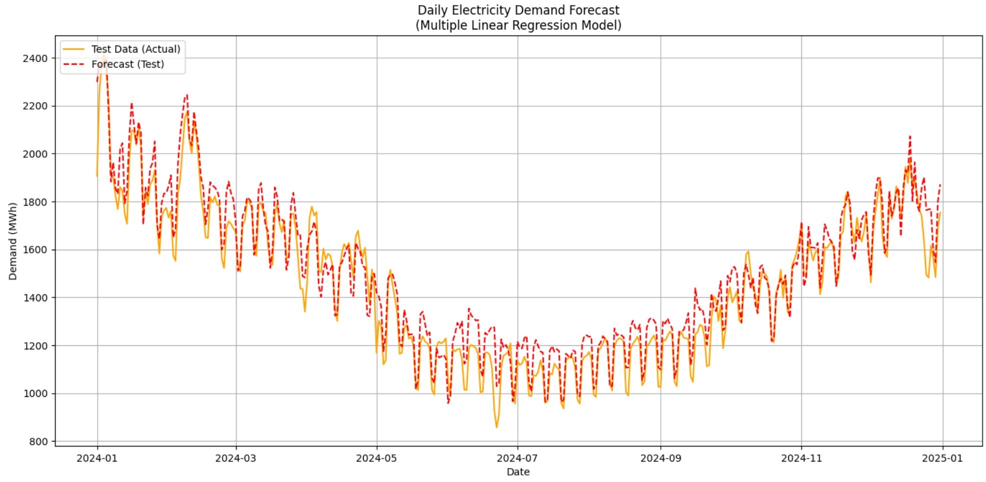


Figure 16. Actual vs. Forecasted Daily Electricity Demand (MWh) for Vaasa (2024) using the Selected Multiple Linear Regression Model (Model 1).

The forecasts generally align with the timing and magnitude of the peaks and troughs in actual demand, although there are periods where the model either overestimates or underestimates the actual consumption. For example, the model seems to capture the winter peaks at the beginning and end of 2024 quite well, and also the summer trough.

The model appears to replicate the day-to-day variability to a fair extent. The R^2 value of 0.928 (Table 10) for this model on the test set indicates that approximately 92.8% of the variance in daily electricity demand during the test period is explained by the model.

Overall, the MLR model, despite its relative simplicity compared to time series models like SARIMAX, demonstrates a strong capability in forecasting daily electricity demand for Vaasa when provided with relevant meteorological and calendar features. The residual analysis indicates a generally good fit, though with some potential for improvement regarding heteroscedasticity and normality of residuals. The model provides interpretable insights into the drivers of daily electricity demand.

6.2 Daily Electricity Demand Forecasting - Random Forest Regressor

The Random Forest algorithm is an ensemble learning method that operates by constructing a multitude of decision trees at training time and outputting the mean prediction (regression) of the individual trees. It is known for its robustness to overfitting and its ability to handle high-dimensional data.

The feature set utilized for these machine learning models is consistent with that identified in the Feature Selection Analysis (Chapter 4) and employed in the MLR approach.

A key advantage of tree-based models like Random Forest is their inherent ability to provide estimates of feature importance. This analysis helps in understanding which variables contribute most significantly to the model's predictions. Figure 17 presents the feature importances derived from the trained Random Forest Regressor.

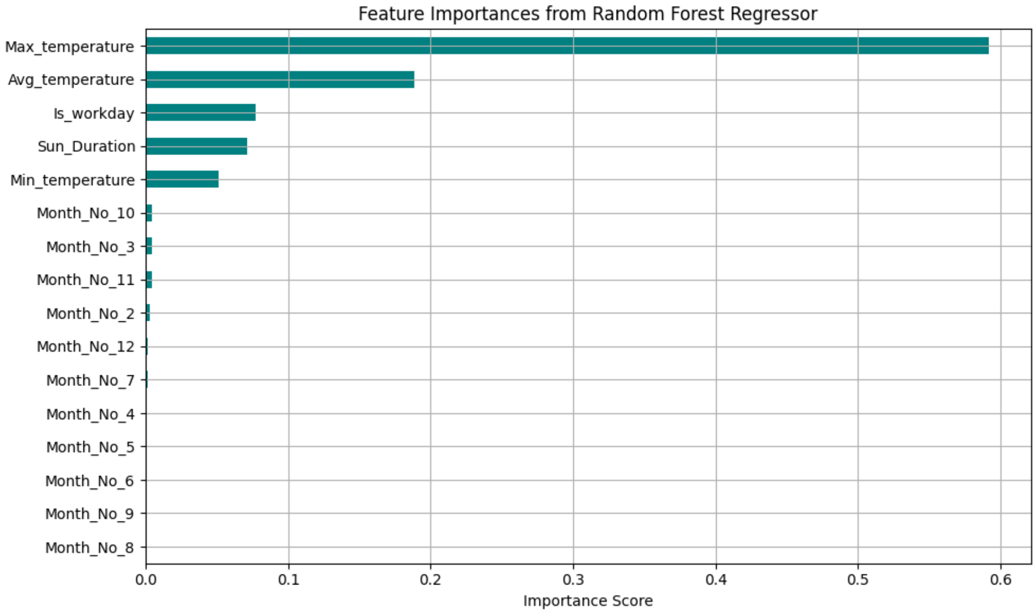


Figure 17. Feature Importance Scores from the Random Forest Regressor for Daily Electricity Demand.

As illustrated in Figure 17, Max_temperature emerges as the most influential predictor, followed by Avg_temperature. The binary variable Is_workday and Sun_Duration also exhibit considerable importance. Min_temperature shows moderate importance. The individual one-hot encoded month variables generally have lower importance scores, with Month_No_10 (October) being the most prominent among them. This ranking provides valuable insights into the primary drivers of daily electricity demand as perceived by the Random Forest model.

6.2.1 Model Performance (Random Forest)

The predictive performance of the Random Forest Regressor on the test dataset (year 2024) was evaluated using standard error metrics. These metrics are summarized in Table 6.2.

Table 11. Performance Diagnostics of the Random Forest Regressor for Daily Electricity Demand Forecasting (Test Set: 2024).

MSE	RMSE	MAE	MAPE	R^2
9827.670	99.130	74.680	0.052	0.900

The Random Forest model achieved an R^2 of 0.900, indicating that it explains 90% of the variance in the daily electricity demand on the test set. The RMSE of 99.130 MWh and MAE of 74.680 MWh quantify the average magnitude of the forecast errors.

6.2.2 Forecast Visualization (Random Forest)

Figure 18 provides a visual comparison of the actual daily electricity demand versus the forecasts generated by the Random Forest Regressor for the test period (2024).

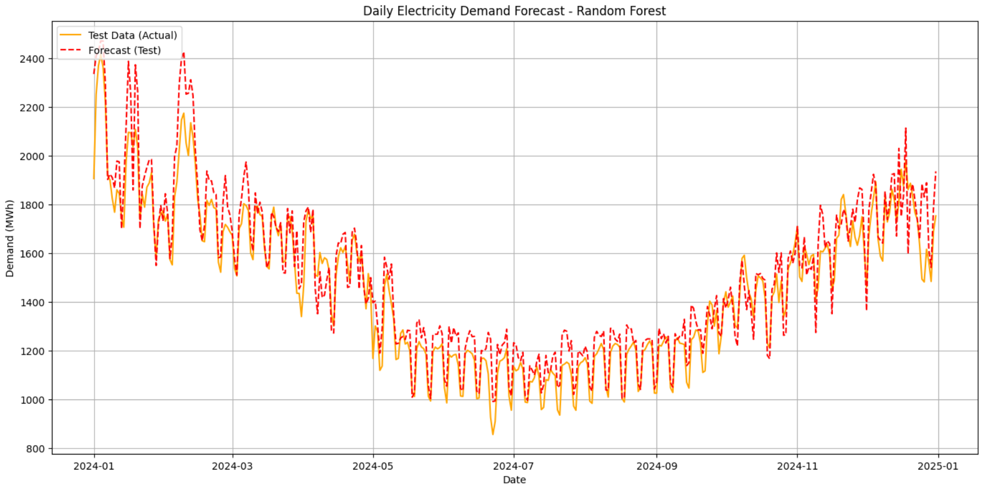


Figure 18. Actual vs. Forecasted Daily Electricity Demand (MWh) for Vaasa (Test Period: 2024) using the Random Forest Regressor.

The plot shows that the Random Forest forecasts (dashed red line) generally track the actual demand (solid yellow line) well, capturing the seasonal trends and weekly fluctuations. The model appears adept at following the overall pattern of demand throughout the year.

6.2.3 Residual Analysis (Random Forest)

To further assess the Random Forest model's adequacy, its residuals were analyzed. Figure 19 displays the standardized residuals plotted against the fitted (predicted) values.

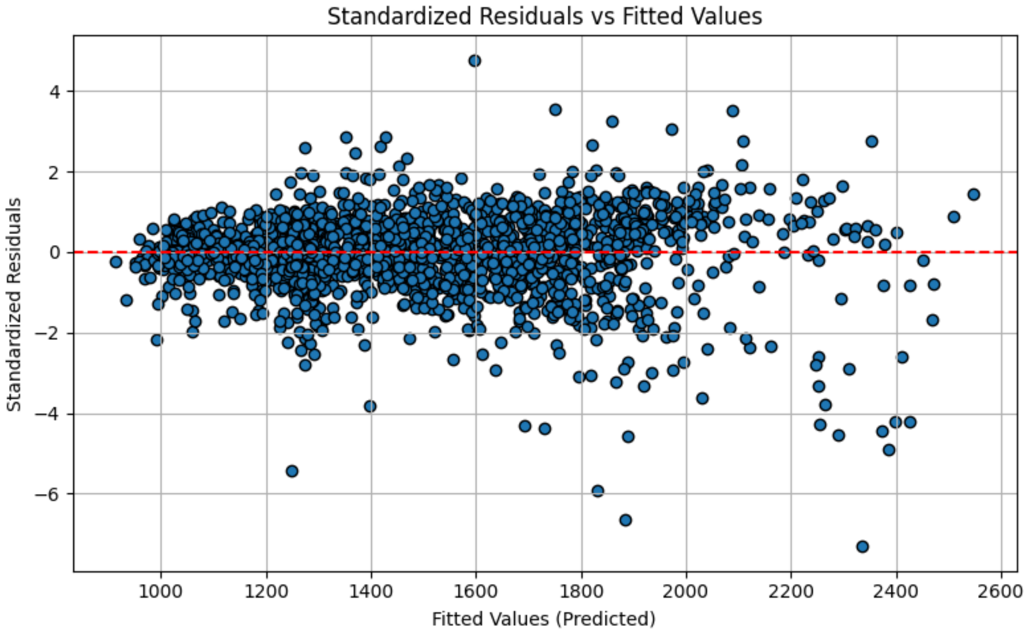


Figure 19. Standardized Residuals versus Fitted Values for the Random Forest Regressor (Daily Demand).

The residuals appear to be randomly scattered around the zero line, without a clear discernible pattern or obvious funnel shape, suggesting that the variance of the errors is relatively constant (homoscedasticity). However, there are several notable outliers, with some residuals extending beyond ±4 standard deviations, indicating specific days where the model's predictions were considerably off.

The Normal Q-Q plot of the residuals for the Random Forest model is presented in Figure 20.

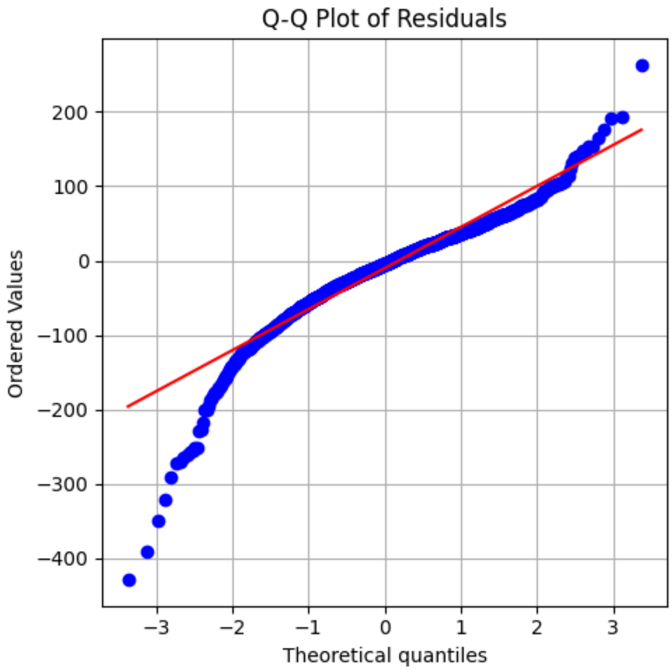


Figure 20. Normal Q-Q Plot of Residuals for the Random Forest Regressor (Daily Demand).

The plot shows significant deviations from the theoretical diagonal line, particularly at both tails. The S-shape indicates that the residuals are not normally distributed and possess much heavier tails than a Gaussian distribution, with a higher concentration of extreme values. This departure from normality is more pronounced than that observed for the MLR model.

6.3 Daily Electricity Demand Forecasting - XGBoost Regressor

XGBoost (Extreme Gradient Boosting) is another powerful ensemble learning algorithm that uses a gradient boosting framework. It has gained popularity due to its high predictive accuracy and efficiency in many machine learning competitions and real-world applications.

Similar to Random Forest, XGBoost can provide feature importance scores. Figure 21 illustrates the feature importances derived from the trained XGBoost Regressor.

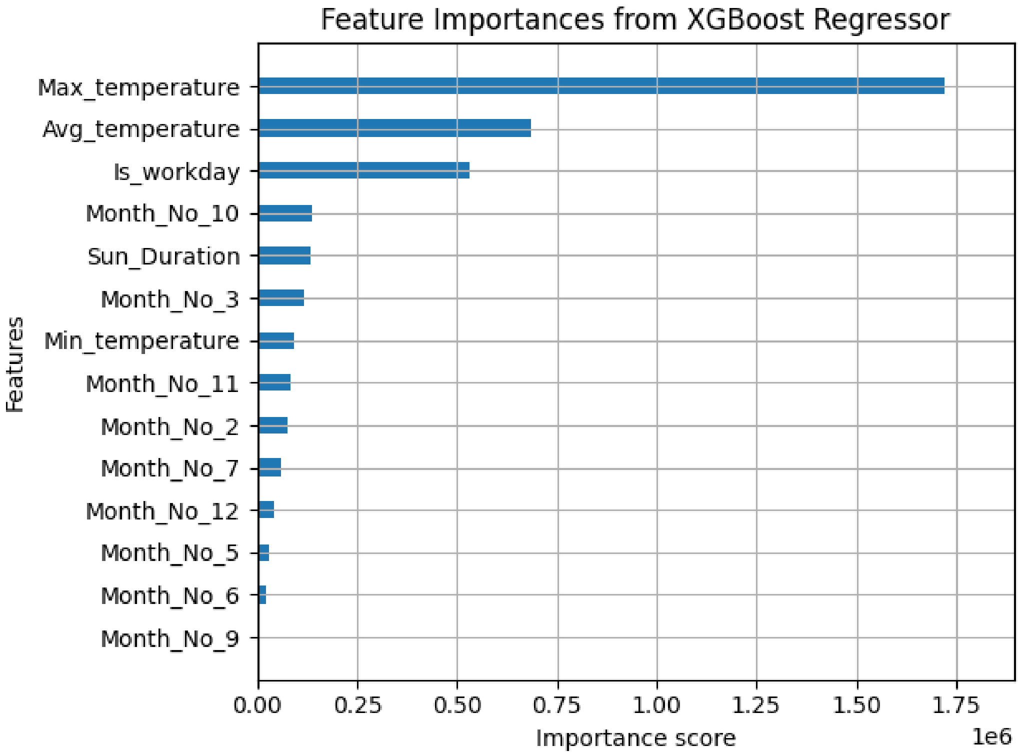


Figure 21. Feature Importance Scores from the XGBoost Regressor for Daily Electricity Demand.

The feature importance ranking from XGBoost (Figure 21) shows Max_temperature as the most dominant feature by a significant margin, followed by Avg_temperature and Is_workday. Month_No_10 (October) and Sun_Duration also show notable importance. The overall ranking is broadly similar to that of the Random Forest, emphasizing the strong influence of temperature and workday patterns.

6.3.1 Model Performance (XGBoost)

The performance diagnostics for the XGBoost Regressor on the test dataset are summarized in Table 12.

Table 12. Performance Diagnostics of the XGBoost Regressor for Daily Electricity Demand Forecasting (Test Set: 2024).

MSE	RMSE	MAE	MAPE	R²
9725.080	98.620	72.720	0.050	0.901

The XGBoost model achieved an R² of 0.901, explaining approximately 90.1% of the variance in daily electricity demand on the test set. Its RMSE is 98.620 MWh, and MAE is 72.720 MWh. These metrics are very similar to those of the Random Forest model.

6.3.2 Forecast Visualization (XGBoost)

Figure 22 visualizes the forecasts generated by the XGBoost Regressor against the actual daily electricity demand for the 2024 test period.

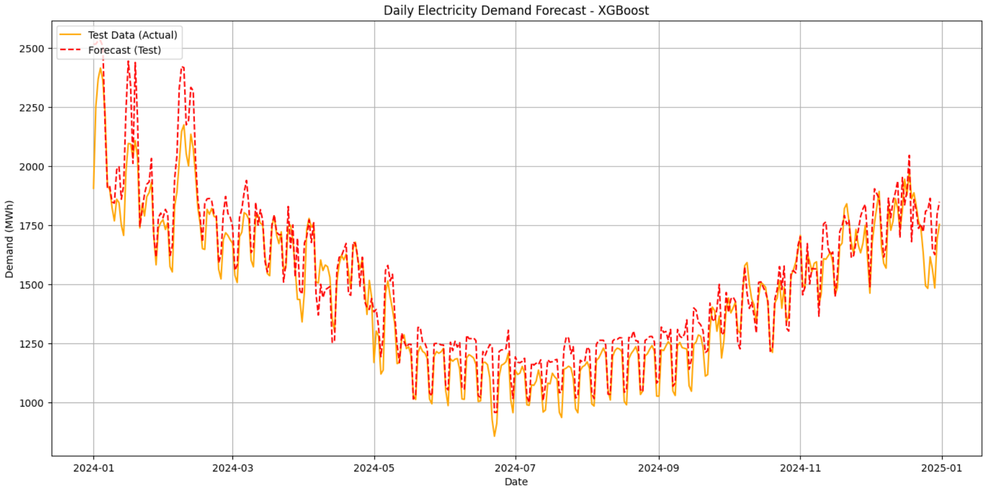


Figure 22. Actual vs. Forecasted Daily Electricity Demand (MWh) for Vaasa (Test Period: 2024) using the XGBoost Regressor.

Visually, the XGBoost forecasts (dashed red line) also demonstrate a good ability to track the actual demand (solid yellow line), capturing the seasonal variations and weekly cycles effectively, much like the Random Forest model.

6.3.3 Residual Analysis (XGBoost)

The residuals from the XGBoost model were also analyzed. Figure 23 shows the plot of standardized residuals against fitted values for the XGBoost model. Similar to the Random Forest, the residuals are generally scattered around zero, but there is evidence of some outliers, with residuals extending beyond ±4 and even ±6 standard deviations.

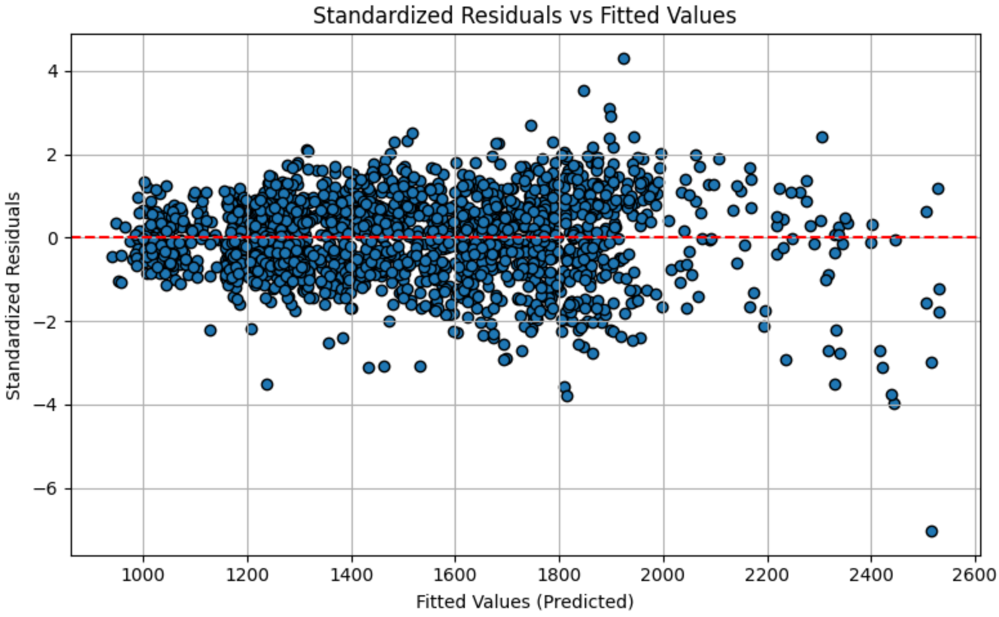


Figure 23. Standardized Residuals versus Fitted Values for the XGBoost Regressor (Daily Demand).

There isn't a strong systematic pattern, but the presence of these extreme residuals is notable.

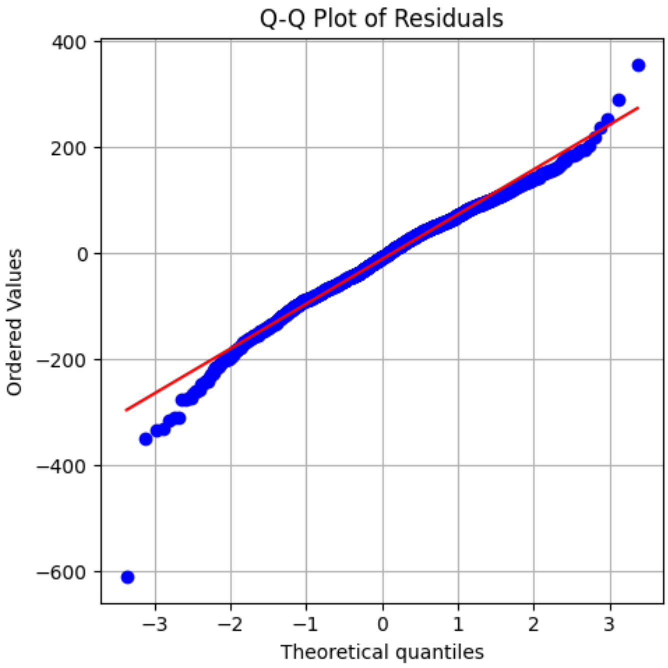


Figure 24. Normal Q-Q Plot of Residuals for the XGBoost Regressor (Daily Demand).

The Normal Q-Q plot for the XGBoost residuals is presented in Figure 24. The deviations from the normality line are even more pronounced than for the Random Forest model, particularly in the lower tail, indicating a significant departure from a normal distribution and the presence of extreme negative residuals.

6.4 Comparison of Daily Electricity Demand Forecasting Approaches

To determine the most effective approach for daily short-term load forecasting for Vaasa, the performance of the MLR model (Model 1, as selected in Section 6.1.3) is compared against the two machine learning models: Random Forest Regressor and XGBoost Regressor. The comparison is based on their respective performance metrics on the unseen test dataset (year 2024). Table 13 summarizes these key diagnostics.

Table 13. Comparative Performance Diagnostics of Multiple Linear Regression and Machine Learning Models for Daily Electricity Demand Forecasting (Test Set: 2024).

	MSE	RMSE	MAE	MAPE	R^2
Linear Regression (Model 1)	7055.743	83.998	65.145	0.047	0.928
Random Forest Regressor	9827.670	99.130	74.680	0.052	0.900
XGBoost Regressor	9725.080	98.620	72.720	0.050	0.901

Based on the comparative performance metrics presented in Table 13, the MLR (Model 1) demonstrates superior predictive accuracy for daily electricity demand forecasting on the 2024 test set compared to both the Random Forest and XGBoost Regressor models.

The MLR model achieved the lowest MSE (7055.743), RMSE (83.998 MWh), MAE (65.145 MWh), and MAPE (0.047 or 4.7%). It also yielded the highest R^2 value (0.928), indicating it explained a larger proportion of the variance in the test data.

While Random Forest and XGBoost are powerful algorithms capable of capturing complex non-linearities, in this specific application with the chosen feature set and daily aggregation level, they did not outperform the simpler MLR model. The machine learning models had higher error metrics (RMSE around 99 MWh) and slightly lower R^2 values (around 0.900-0.901).

Furthermore, the residual analysis for the machine learning models, particularly the Q-Q plots (Figures 20 and 24), indicated more significant deviations from normality compared to the MLR model's residuals (Figure 15). While normality of residuals is a less strict assumption for the predictive use of machine learning models compared to statistical inference in MLR, the presence of numerous extreme outliers in the ML models' residuals is also a point of attention.

Therefore, based on the empirical evidence from the test set performance, the MLR model (Model 1, incorporating Avg_temperature, Min_temperature, Max_temperature, Sun_Duration, Is_workday, and one-hot encoded Month variables) is selected as the best-performing model for daily short-term load forecasting in this study. Its combination of strong predictive accuracy, higher R^2, and more favorable residual characteristics (compared to the ML models in terms of normality, though with some noted heteroscedasticity) makes it the preferred choice among the models evaluated for this specific task. The interpretability of the MLR model is also an advantage in understanding the impact of different drivers on electricity demand.

7 SHORT-TERM (HOURLY) LOAD FORECASTING

Forecasting electricity demand at an hourly resolution presents distinct challenges and opportunities, as it requires capturing more granular temporal patterns and the immediate effects of changing conditions. This section outlines the application of a MLR (MLR) approach to model and predict hourly electricity demand. Subsequently, the analysis will be extended by exploring machine learning methodologies (Random Forest and XGBoost) to capture the potentially more complex and dynamic patterns inherent in hourly electricity consumption.

7.1 Hourly Electricity Demand Forecasting – Multiple Linear Regression Approach

The application of MLR techniques produced effective models for forecasting electricity demand at both monthly and daily resolutions. As a preliminary step, MLR was also explored for modeling hourly electricity demand.

The feature set for the hourly models is informed by prior exploratory data analysis and feature engineering (as described in Section 3.2 of this thesis). The key predictors include:

Continuous Features: Average Temperature (Avg_temperature).

Categorical Features:

- Month: To capture monthly seasonal variations.
- ISO weekday: To account for differences in demand across the days of the week (Monday=1 to Sunday=7).
- Hour of the Day: To model the distinct intra-day demand profile.
- Sun Flag: A categorical variable designed to represent different daylight phases, defined as:

  - Flag 1: Nighttime hours (between the previous day's sunset and the current day's sunrise).
  - Flag 2: Transition periods (the first two hours after sunrise and the last two hours before sunset).
  - Flag 3: Daytime hours (remaining hours between the Flag 2 periods).

These categorical features were one-hot encoded for inclusion in the regression models.

Three distinct MLR model configurations were developed and evaluated, varying primarily in their representation of day-of-week and monthly seasonality:

Model 1: Utilized Avg_temperature, one-hot encoded Month, one-hot encoded ISO weekday, one-hot encoded Hour of the Day, and one-hot encoded Sun Flag.

Model 2: Similar to Model 1, but replaced the one-hot encoded ISO weekday with the binary Is_workday feature (1 for workday, 0 for weekend).

Model 3: Used Avg_temperature, the binary Is_workday, one-hot encoded Season (instead of Month), one-hot encoded Hour of the Day, and one-hot encoded Sun Flag.

7.1.1 Model Selection for Hourly Forecasting

To identify the most effective MLR specification for hourly demand forecasting, the three model configurations were compared based on standard performance metrics. These metrics, calculated on a designated test set (year 2024), are presented in Table 14. Based on the evaluation metrics, Model 1 demonstrates the best performance among the three configurations. It achieved the lowest Mean Squared

Error (MSE) of 21.400, Root Mean Squared Error (RMSE) of 4.626 MWh, and Mean Absolute Error (MAE) of 3.666 MWh.

Table 14. Comparative Performance Diagnostics for Hourly Multiple Linear Regression Models.

Model No	Included Features Description	MSE	RMSE	MAE	MAPE	R²
1	Avg Temp, Month (OHE), ISO Weekday (OHE), Hour (OHE), Sun Flag (OHE)	21.400	4.626	3.666	0.066	0.901
2	Avg Temp, Month (OHE), Is_workday, Hour (OHE), Sun Flag (OHE)	21.577	4.645	3.680	0.072	0.900
3	Avg Temp, Season (OHE), Is_workday, Hour (OHE), Sun Flag (OHE)	25.296	5.030	3.957	0.072	0.883

It also yielded the highest R² value of 0.901, indicating that this model specification explains approximately 90.1% of the variance in hourly electricity demand on the test set. Model 2 performed very similarly, while Model 3, which used aggregated seasonal dummies instead of monthly ones, showed a discernible decrease in performance. Therefore, Model 1 is selected for detailed interpretation and further analysis.

7.1.2 Interpretation of the Selected Hourly Multiple Linear Regression Model (Model 1)

The selected MLR Model 1 predicts hourly electricity demand using Avg_temperature, and one-hot encoded representations of Month, ISO weekday, Hour of the Day, and Sun Flag. The estimated coefficients, standard errors, Z-statistics, and p-values for this model are presented in Table 15.

Table 15. Estimated Coefficients for the Hourly MLR Model (Model 1).

	coef	std err	z	P>|z|	[0.025	0.975]
const	68.7654	0.209	328.277	0.000	68.355	69.176
Avg_temperature	-0.7525	0.007	-103.028	0.000	-0.767	-0.738
Month_No_2	0.4938	0.161	3.072	0.002	0.179	0.809
Month_No_3	-4.5917	0.160	-28.688	0.000	-4.905	-4.278
Month_No_4	-10.0465	0.170	-59.169	0.000	-10.379	-9.714
Month_No_5	-13.0498	0.188	-69.447	0.000	-13.418	-12.681
Month_No_6	-12.2653	0.221	-55.404	0.000	-12.699	-11.831
Month_No_7	-13.5490	0.221	-61.358	0.000	-13.982	-13.116
Month_No_8	-11.5929	0.211	-54.929	0.000	-12.007	-11.179
Month_No_9	-11.5141	0.188	-61.236	0.000	-11.883	-11.146
Month_No_10	-10.0137	0.168	-59.697	0.000	-10.342	-9.685
Month_No_11	-5.4876	0.158	-34.634	0.000	-5.798	-5.177
Month_No_12	-2.8538	0.156	-18.239	0.000	-3.160	-2.547
Iso_Weekday_2	1.1566	0.120	9.618	0.000	0.921	1.392
Iso_Weekday_3	1.1748	0.120	9.779	0.000	0.939	1.410
Iso_Weekday_4	0.7768	0.120	6.465	0.000	0.541	1.012
Iso_Weekday_5	-0.1173	0.120	-0.976	0.329	-0.353	0.118
Iso_Weekday_6	-7.2809	0.120	-60.600	0.000	-7.516	-7.045
Iso_Weekday_7	-8.9240	0.120	-74.292	0.000	-9.159	-8.689
Hour_1	-3.0225	0.222	-13.592	0.000	-3.458	-2.587
Hour_2	-4.9927	0.222	-22.449	0.000	-5.429	-4.557
Hour_3	-5.6588	0.223	-25.428	0.000	-6.095	-5.223
Hour_4	-4.3007	0.224	-19.227	0.000	-4.739	-3.862
Hour_5	0.1554	0.225	0.690	0.490	-0.286	0.597
Hour_6	6.2577	0.228	27.400	0.000	5.810	6.705
Hour_7	10.6032	0.233	45.483	0.000	10.146	11.060
Hour_8	13.7300	0.239	57.380	0.000	13.261	14.199
Hour_9	15.7745	0.247	63.807	0.000	15.290	16.259
Hour_10	17.5396	0.257	68.199	0.000	17.036	18.044
Hour_11	18.6017	0.261	71.171	0.000	18.089	19.114
Hour_12	18.4207	0.267	68.968	0.000	17.897	18.944
Hour_13	17.4671	0.267	65.388	0.000	16.944	17.991
Hour_14	16.6032	0.262	63.441	0.000	16.090	17.116
Hour_15	16.3342	0.258	63.244	0.000	15.828	16.840
Hour_16	16.0228	0.250	63.990	0.000	15.532	16.514
Hour_17	15.3931	0.244	63.209	0.000	14.916	15.870
Hour_18	15.1494	0.240	63.215	0.000	14.680	15.619
Hour_19	14.3275	0.235	60.954	0.000	13.867	14.788
Hour_20	11.6433	0.231	50.387	0.000	11.190	12.096
Hour_21	10.4465	0.228	45.773	0.000	9.999	10.894
Hour_22	9.7983	0.225	43.464	0.000	9.356	10.240
Hour_23	5.2283	0.224	23.364	0.000	4.790	5.667
Sun_Flag_2	-1.4892	0.117	-12.726	0.000	-1.719	-1.260
Sun_Flag_3	-3.1790	0.148	-21.409	0.000	-3.470	-2.888

Interpreting Coefficients:

Constant (Intercept): The intercept is 68.7654 MWh. This represents the baseline predicted hourly electricity demand when Avg_temperature is zero, and all categorical features are at their reference levels (e.g., January, Monday (assuming ISO Weekday 1 is the reference), Hour 0 (midnight to 1 AM), and Sun Flag 1 (nighttime)). Direct interpretation of the intercept requires caution, as these specific conditions might be outside the typical observed range of the data. It primarily serves as the model's starting point.

Avg_temperature: The coefficient for average daily temperature is -0.7525, and it is highly statistically significant (P>|z| = 0.000). This indicates that for each one-degree Celsius increase in the average daily temperature, the hourly electricity demand is predicted to decrease by approximately 0.75 MWh, holding all other factors constant. This strong negative relationship is consistent with expectations, primarily reflecting reduced heating requirements as temperatures rise.

Month Dummies (Reference: January / Month_No_1):

The coefficients for the month dummy variables are all statistically significant and illustrate the annual seasonal pattern relative to January:

February (Month_No_2: +0.4938 MWh) shows a slightly higher demand.

Spring months (March Month_No_3: -4.59 MWh; April Month_No_4: -10.05 MWh; May Month_No_5: -13.05 MWh) exhibit progressively lower demand.

Summer months (June Month_No_6: -12.27 MWh; July Month_No_7: -13.55 MWh; August Month_No_8: -11.59 MWh) generally show the lowest demand levels compared to January, reflecting reduced heating and potentially some offsetting cooling load captured indirectly.

Autumn months (September Month_No_9: -11.51 MWh; October Month_No_10: -10.01 MWh; November Month_No_11: -5.49 MWh) show demand gradually increasing but still below January levels.

December (Month_No_12: -2.85 MWh) has lower demand than January, ceteris paribus.

These coefficients effectively map out the expected annual cycle of electricity consumption.

ISO weekday Dummies (Reference: Monday / Iso_Weekday_1):

The coefficients for the ISO weekday dummies (where Monday is assumed to be the reference category, Iso_Weekday_1) reveal the typical weekly demand pattern:

Tuesday (Iso_Weekday_2: +1.1566 MWh), Wednesday (Iso_Weekday_3: +1.1748 MWh), and Thursday (Iso_Weekday_4: +0.7768 MWh) show slightly higher electricity demand compared to Monday, and these effects are statistically significant.

Friday (Iso_Weekday_5: -0.1173 MWh) is not statistically significantly different from Monday (P>|z| = 0.329).

Saturday (Iso_Weekday_6: -7.2809 MWh) and Sunday (Iso_Weekday_7: -8.9240 MWh) exhibit substantially and statistically significantly lower electricity demand compared to Monday. This aligns with typical observations of reduced commercial and industrial activity during weekend days.

Hour of the Day Dummies (Reference: Hour 0 / Midnight-1 AM):

The coefficients for the Hour of the Day dummy variables trace a distinct and statistically significant diurnal (daily) electricity demand profile:

Demand is at its lowest during the early morning hours. Compared to Hour 0 (midnight-1 AM), demand is significantly lower for Hour_1 (-3.02 MWh), Hour_2 (-4.99 MWh), Hour_3 (-5.66 MWh), and Hour_4 (-4.30 MWh). Hour_5 (+0.16 MWh) is not statistically significantly different from Hour 0.

Demand begins to rise significantly from Hour_6 (+6.26 MWh) onwards, corresponding to the start of the typical workday and morning activities.

The demand peaks during the daytime and early evening hours. The highest demand relative to Hour 0 is observed around Hour_11 (11 AM - 12 PM), with an estimated increase of +18.60 MWh. Strong positive coefficients persist through Hour_18 (6 PM - 7 PM).

Demand then gradually decreases in the late evening hours. For example, Hour_20 (8 PM - 9 PM) is +11.64 MWh higher than Hour 0, and Hour_23 (11 PM - midnight) is +5.23 MWh higher than Hour 0.

These hourly coefficients effectively model the typical daily load curve with morning ramp-up, daytime/early evening peaks, and nighttime troughs.

Sun Flag Dummies (Reference: Sun Flag 1 / Nighttime):

The Sun Flag variables quantify the impact of broad daylight phases, after accounting for the specific hour of the day and other factors:

Sun_Flag_2 (Transition periods: the first two hours after sunrise and the last two hours before sunset): The coefficient is -1.4892 MWh, and it is highly statistically significant (P>|z| = 0.000). This indicates that, compared to nighttime hours (Sun Flag 1), these transition periods are associated with an average decrease in hourly demand of approximately 1.49 MWh, holding other factors constant.

Sun_Flag_3 (Daytime hours: remaining hours between the Flag 2 periods): The coefficient is -3.1790 MWh, also highly statistically significant (P>|z| = 0.000). This suggests that daytime hours (as defined by Sun Flag 3) have, on average, about 3.18 MWh lower electricity demand compared to nighttime hours (Sun Flag 1), ceteris paribus.

These negative coefficients for Sun_Flag_2 and Sun_Flag_3 (relative to nighttime) might seem counter-intuitive at first glance, as daytime often involves higher activity. However, it's important to remember these effects are after controlling for the specific Hour of the Day dummies (which capture the main diurnal shape), Avg_temperature, Month, and ISO weekday. Therefore, these Sun Flag coefficients might be capturing subtle residual effects. For instance, for a given hour that is already identified as a peak hour by its Hour_X dummy, the additional classification of it being "daytime" under Sun_Flag_3 might be associated with slightly lower demand due to factors like increased natural illumination reducing artificial lighting needs, or specific outdoor activities that reduce in-home consumption, which are not fully captured by the other variables.

7.1.3 Residual Analysis of the Selected Hourly Multiple Linear Regression Model

To assess the adequacy of the selected hourly MLR model (Model 1), an analysis of its residuals was performed.

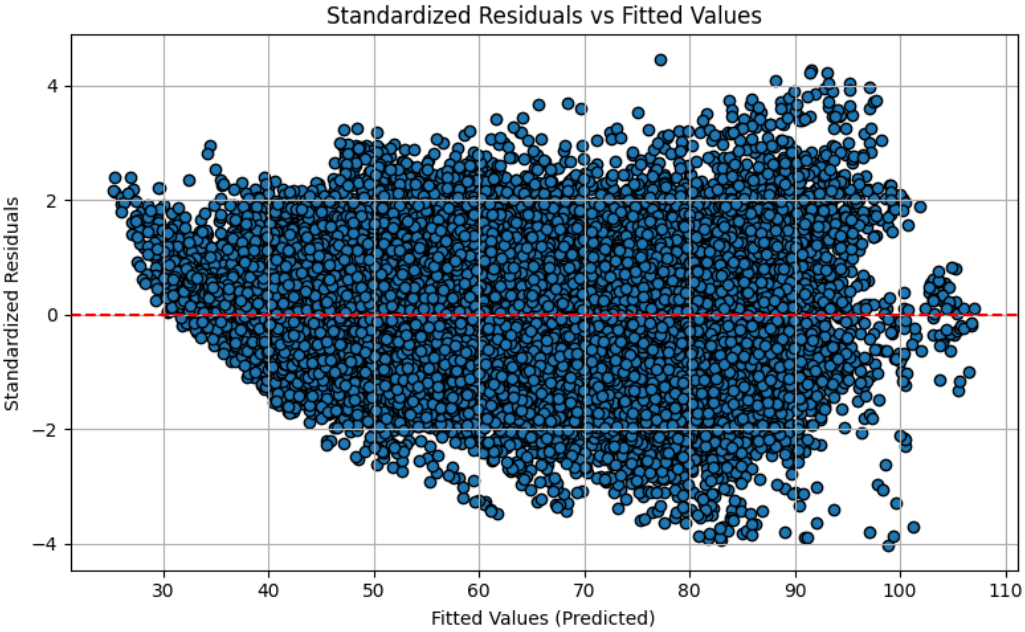


Figure 25. Standardized Residuals versus Fitted Values for the Selected Hourly Multiple Linear Regression Model (Model 1).

Figure 25 displays the standardized residuals plotted against the fitted (predicted) hourly electricity demand values. The plot shows a dense cloud of points, generally scattered around the horizontal zero line. There is no immediately obvious curvilinear pattern, suggesting the linearity assumption is acceptable. However, the spread of the residuals appears to widen as the fitted values increase (particularly for fitted values above approximately 60-70 MWh), indicating potential heteroscedasticity (non-constant variance of errors). This means the model's predictions might be less precise (i.e., have larger errors) for higher demand hours. Several outliers are visible, with residuals extending beyond ±4 standard deviations, signifying specific hours where the model's predictions were considerably inaccurate.

Figure 26 presents the Normal Q-Q plot of the residuals. The points deviate noticeably from the theoretical diagonal line, especially at the tails. Both the lower and upper tails of the residual distribution are heavier than what would be expected from a normal distribution (e.g., observed quantiles around -20 and +20 are more extreme than the theoretical quantiles around -3 and +3 respectively). This indicates leptokurtosis.

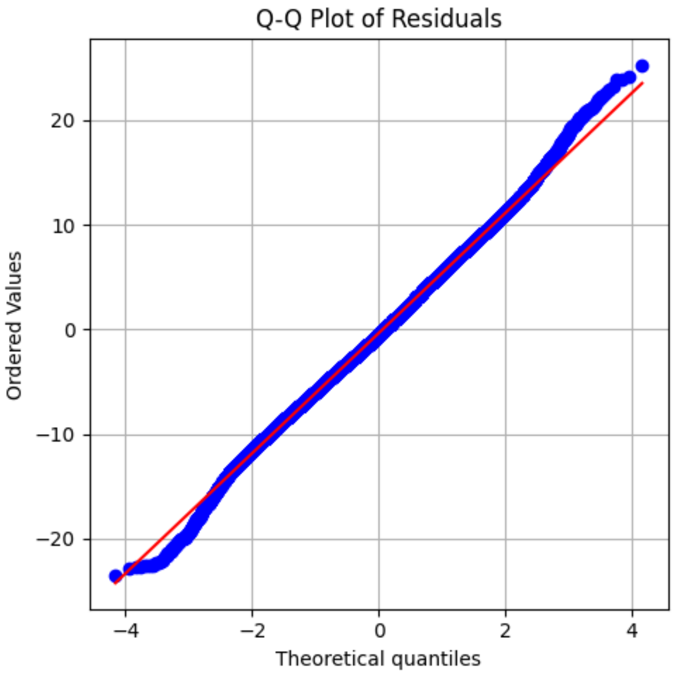


Figure 26. Normal Q-Q Plot of Residuals for the Selected Hourly Multiple Linear Regression Model (Model 1).

This departure from normality suggests that while the model captures the central tendency of the demand, it may underestimate the frequency or magnitude of extreme prediction errors. This can affect the reliability of confidence intervals for predictions.

7.1.4 Model Performance and Forecast Visualization (Hourly)

Figure 27 illustrates the performance of the selected hourly MLR model (Model 1) by comparing its forecasts for the test period (January 2024 - December 2024) against the actual observed hourly electricity demand.

Forecast Accuracy on Test Data: The plot demonstrates that the model's forecasts (dashed red line) effectively track the actual hourly electricity demand (solid yellow line) throughout 2024.

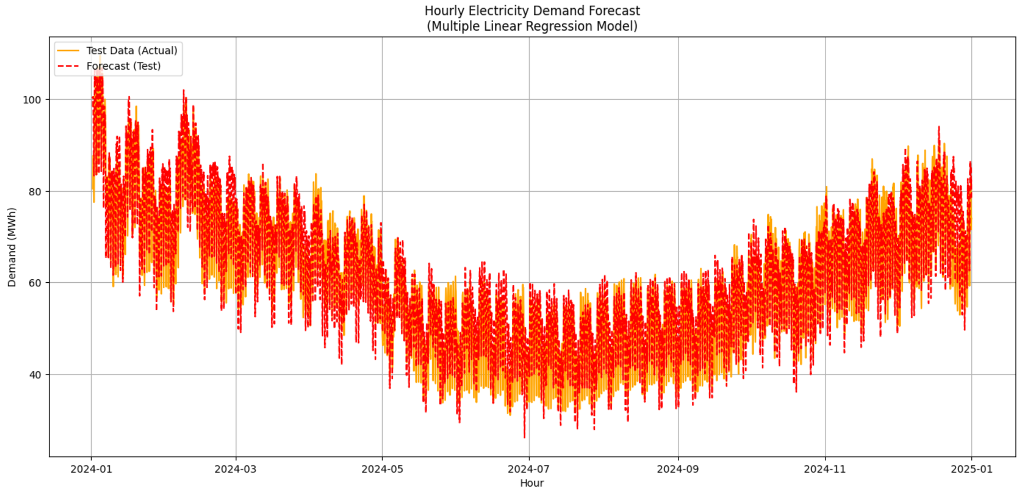


Figure 27. Actual vs. Forecasted Hourly Electricity Demand (MWh) for Vaasa (Test Period: 2024) using the Selected Multiple Linear Regression Model (Model 1).

The model successfully captures the main characteristics of the hourly demand, including:

- The overarching annual seasonal trend (visible as higher average demand in winter months and lower in summer months, influencing the baseline of the hourly fluctuations).
- The weekly cycle (differences between workdays and weekends, as captured by the ISO weekday dummies).
- The distinct diurnal (daily) pattern with peaks during active hours and troughs during nighttime/early morning hours.

The forecasts generally align well with the timing of peaks and troughs in actual demand on an hourly basis. The model explains a significant portion of the variance in hourly demand, as indicated by the R^2 value of 0.901 (from Table 14). While there are hour-to-hour deviations, as expected in any forecasting model, the overall consistency and directional accuracy of the forecasts at this granular hourly level are noteworthy.

The selected MLR model (Model 1) demonstrates a strong capability for forecasting hourly electricity demand in Vaasa. It effectively incorporates average daily temperature, monthly seasonality, day-of-week effects, distinct hourly patterns, and broad daylight phases (Sun Flags) to achieve a high coefficient of determination (R^2 of 0.901) on the test data. The interpretation of the model coefficients provides valuable quantitative insights into how each of these factors influences hourly electricity consumption.

However, the residual analysis highlighted some areas for potential improvement or considerations.

The presence of non-constant variance in the residuals, particularly an increase in error spread at higher demand levels, suggests that the model's precision may vary. Techniques such as weighted least squares or transforming the dependent variable could potentially address this.

The residuals exhibit heavier tails than a normal distribution, which can impact the reliability of prediction intervals.

While most coefficients align with expectations, the interpretation of some dummy variable effects (like Sun_Flag relative to Hour_of_the_Day) requires careful consideration of what other factors are already controlled for in the model.

Despite these points, the MLR model serves as a robust and interpretable baseline for hourly short-term load forecasting. Its ability to explain over 90% of the variance in hourly demand with readily available features is a significant finding. Further investigation with machine learning models in subsequent sections will explore whether these more flexible techniques can address the observed limitations (such as heteroscedasticity or capturing more complex non-linearities) and offer enhanced predictive performance for hourly load forecasting.

7.2 Hourly Electricity Demand Forecasting - Random Forest Regressor

The Random Forest algorithm constructs an ensemble of decision trees during training. For regression tasks, the final prediction is typically the average of the predictions from all individual trees. This approach is known for its robustness against overfitting and its capability to model non-linear relationships.

The feature set utilized for these machine learning models remains consistent with that employed in the selected MLR Model 1 (detailed in Section 7.1.1). This includes:

- Continuous Feature: Avg_temperature.
- Categorical Features (One-Hot Encoded): Month, ISO weekday, Hour of the Day, and Sun Flag.

The training period (January 2020 - December 2023) and test period (January 2024 - December 2024) are maintained for direct comparability with the MLR results.

7.2.1 Feature Importance Analysis (Random Forest)

An inherent benefit of tree-based ensemble models like Random Forest is the ability to estimate the relative importance of each feature in the prediction process. Figure 28 presents the feature importance scores derived from the trained Random Forest Regressor for hourly electricity demand.

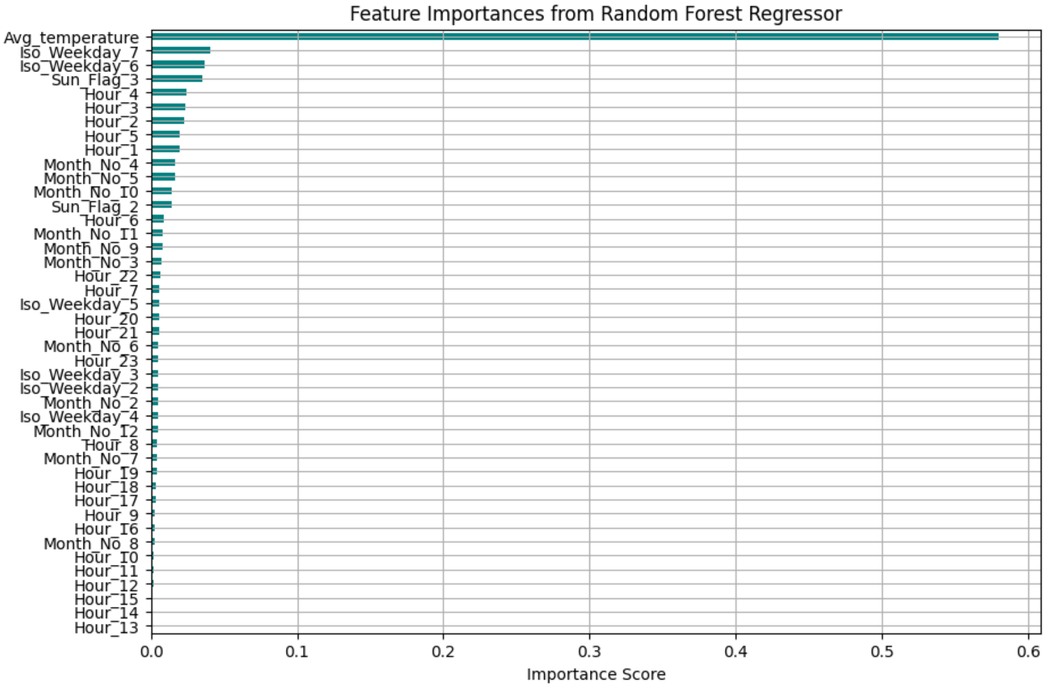


Figure 28. Feature Importance Scores from the Random Forest Regressor for Hourly Electricity Demand Forecasting.

As depicted in Figure 28, Avg_temperature is identified as the most influential feature by the Random Forest model. Following this, Iso_Weekday_7 (Sunday) and Iso_Weekday_6 (Saturday) show considerable importance, highlighting the distinct demand patterns on weekend days. Sun_Flag_3 (Daytime hours) and several early morning hour dummies (e.g., Hour_4, Hour_3, Hour_2) also rank highly. This suggests that temperature, day type (weekend vs. weekday implicitly through ISO weekday dummies), daylight phases, and specific early morning hours are key drivers as determined by this model. Interestingly, individual month dummies and later daytime hour dummies generally show lower importance scores in this specific Random Forest model configuration.

7.2.2 Model Performance (Random Forest Regressor)

The predictive performance of the Random Forest Regressor was evaluated on the test dataset (year 2024). The key performance metrics are summarized in Table 16.

Table 16. Performance Diagnostics of the Random Forest Regressor for Hourly Electricity Demand Forecasting (Test Set: 2024).

MSE	RMSE	MAE	MAPE	R^2
23.190	4.820	3.670	0.061	0.892

The Random Forest model achieved an R^2 of 0.892, indicating that it explains approximately 89.2% of the variance in hourly electricity demand on the test set. The Root Mean Squared Error (RMSE) was 4.820 MWh, and the Mean Absolute Error (MAE) was 3.670 MWh.

7.2.3 Forecast Visualization (Random Forest Regressor)

Figure 29 provides a visual comparison of the actual hourly electricity demand versus the forecasts generated by the Random Forest Regressor for the 2024 test period.

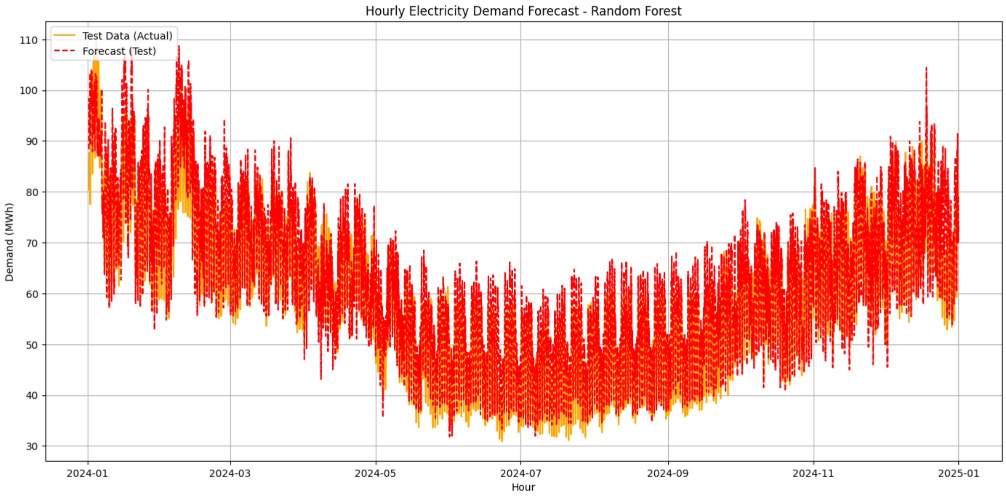


Figure 29. Actual vs. Forecasted Hourly Electricity Demand (MWh) for Vaasa (Test Period: 2024) using the Random Forest Regressor.

The plot illustrates that the Random Forest forecasts (dashed red line) closely follow the actual demand (solid yellow line). The model effectively captures the annual seasonal trend, the weekly cycle (evident in the repeating up-and-down patterns), and the pronounced diurnal (hourly) fluctuations in electricity demand.

7.2.4 Residual Analysis (Random Forest Regressor)

A diagnostic analysis of the Random Forest model's residuals was conducted. Figure 30 displays the standardized residuals plotted against the fitted (predicted) values. The residuals are predominantly scattered around the zero line. However, the plot exhibits a noticeable fanning-out pattern, where the variance of the residuals increases as the fitted values increase.

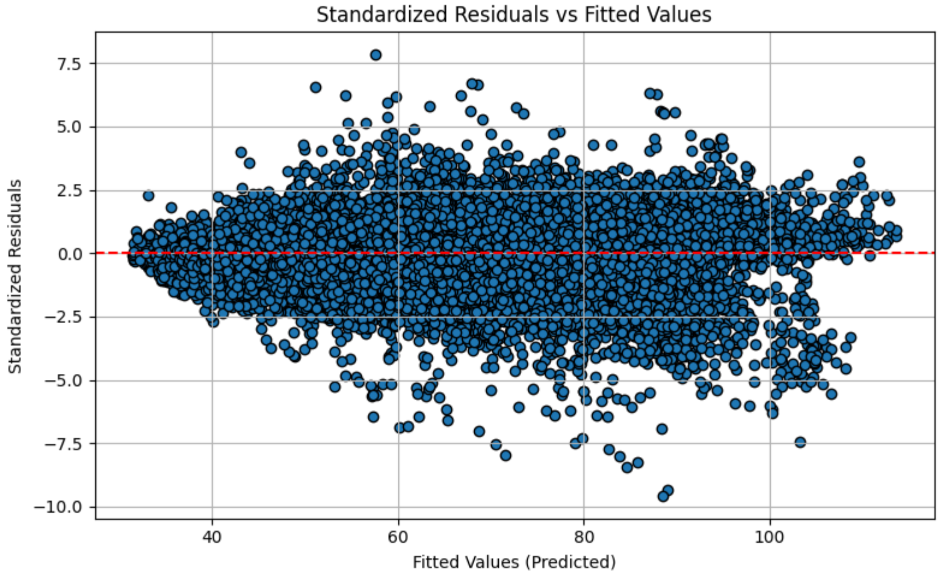


Figure 30. Standardized Residuals versus Fitted Values for the Random Forest Regressor (Hourly Demand).

This indicates heteroscedasticity, meaning the model's prediction errors are larger for higher demand levels. Numerous outliers are present, with some residuals extending beyond ±7.5 standard deviations, particularly at higher fitted values.

The Normal Q-Q plot of the residuals for the Random Forest model is presented in Figure 31. The plot shows substantial deviations from the theoretical diagonal line, especially at both tails.

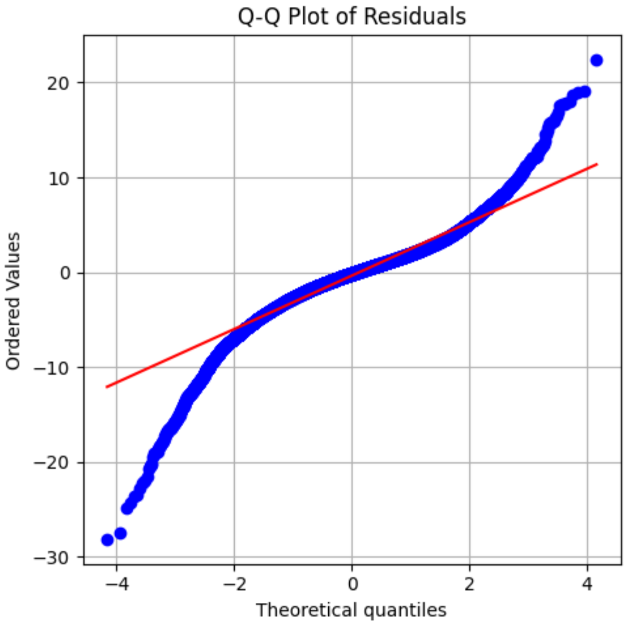


Figure 31. Normal Q-Q Plot of Residuals for the Random Forest Regressor (Hourly Demand).

This S-shaped pattern, with observed quantiles being more extreme than theoretical quantiles, indicates that the residuals are not normally distributed and possess significantly heavier tails than a Gaussian distribution. This departure from normality is quite pronounced.

7.3 Hourly Electricity Demand Forecasting - XGBoost Regressor

XGBoost (Extreme Gradient Boosting) is an advanced implementation of gradient boosted decision trees, renowned for its high performance and efficiency. It builds models in a stage-wise fashion and generalizes them by allowing optimization of an arbitrary differentiable loss function.

The feature set utilized for these machine learning models remains consistent with that employed in the selected MLR Model 1 (detailed in Section 7.1.1). This includes:

- Continuous Feature: Avg_temperature.
- Categorical Features (One-Hot Encoded): Month, ISO weekday, Hour of the Day, and Sun Flag.

The training period (January 2020 - December 2023) and test period (2024) are maintained for direct comparability with the MLR results.

7.3.1 Feature Importance Analysis (XGBoost Regressor)

Similar to Random Forest, XGBoost can provide feature importance scores. Figure 32 illustrates the feature importances derived from the trained XGBoost Regressor.

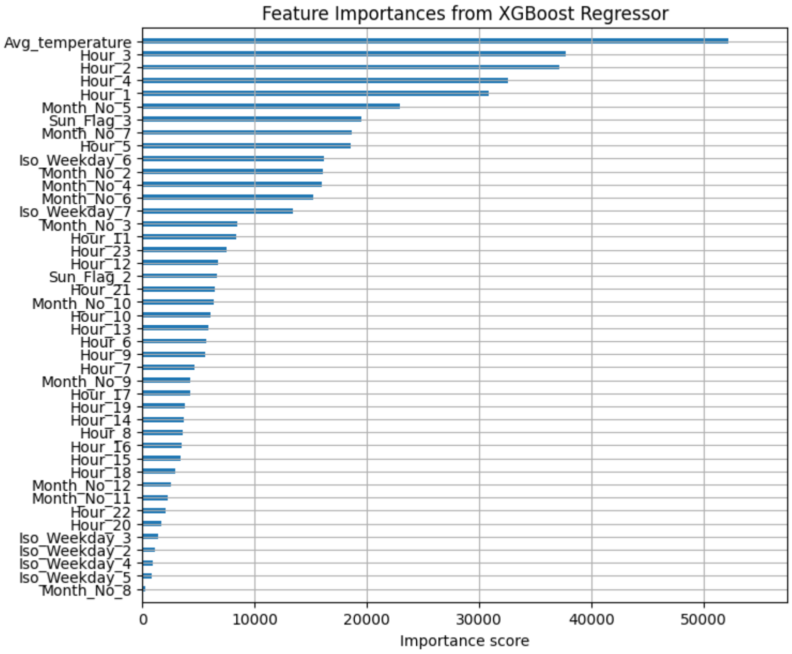


Figure 32. Feature Importance Scores from the XGBoost Regressor for Hourly Electricity Demand Forecasting.

Figure 32 reveals that Avg_temperature is the most dominant factor in predicting hourly short-term load, followed by specific early morning hours (like Hour_3, Hour_2, and Hour_4) and certain months (such as Month_No_5 - May). This indicates that temperature, daily cyclical patterns, and seasonal effects are primary drivers for the model. While features like Sun_Flag (indicating sunshine or day/night) and Iso_Weekday (day of the week) also contribute, their influence is less pronounced compared to temperature and specific times of the day and year. Overall, the model heavily relies on temperature and temporal features to make its forecasts.

7.3.2 Model Performance (XGBoost Regressor)

The performance diagnostics for the XGBoost Regressor on the 2024 test dataset are summarized in Table 17.

Table 17. Performance Diagnostics of the XGBoost Regressor for Hourly Electricity Demand Forecasting (Test Set: 2024).

MSE	RMSE	MAE	MAPE	R²
21.220	4.610	3.610	0.063	0.902

The XGBoost model achieved an R² of 0.902, explaining approximately 90.2% of the variance in hourly electricity demand on the test set. Its RMSE is 4.610 MWh, and MAE is 3.610 MWh. These metrics suggest a performance level slightly better than the Random Forest model and comparable to the MLR model.

7.3.3 Forecast Visualization (XGBoost Regressor)

Figure 33 visualizes the forecasts generated by the XGBoost Regressor against the actual hourly electricity demand for the 2024 test period. Visually, the XGBoost forecasts (dashed red line) also demonstrate a strong ability to track the actual demand (solid yellow line). The model effectively captures the seasonal, weekly, and diurnal patterns in the hourly electricity demand data, similar to the Random Forest model.

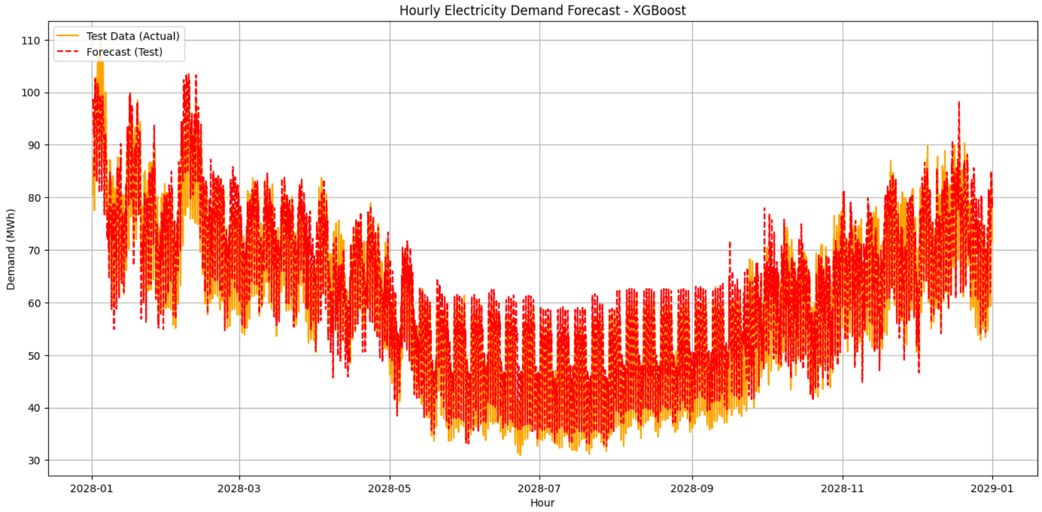


Figure 33. Actual vs. Forecasted Hourly Electricity Demand (MWh) for Vaasa (Test Period: 2024) using the XGBoost Regressor.

7.3.4 Residual Analysis (XGBoost Regressor)

Figure 34 shows the plot of standardized residuals against fitted values for the XGBoost model.

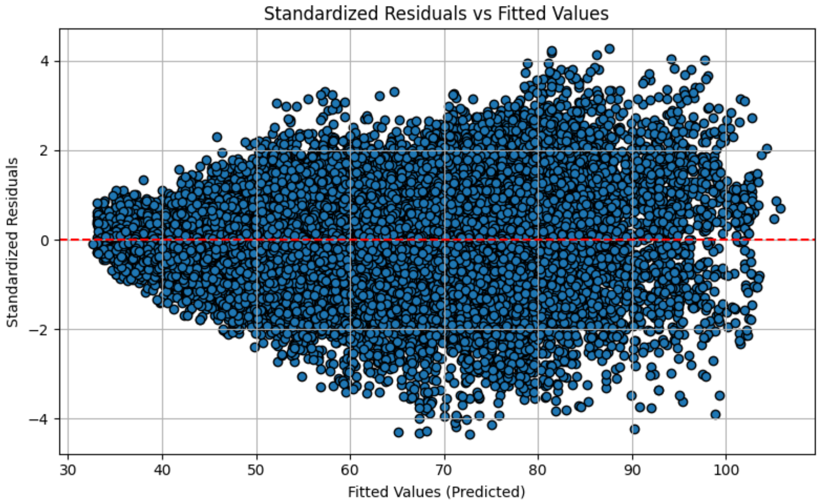


Figure 34. Standardized Residuals versus Fitted Values for the XGBoost Regressor (Hourly Demand).

Similar to the Random Forest model, the residuals are generally scattered around zero but exhibit a clear pattern of increasing variance with increasing fitted values (heteroscedasticity). A significant number of outliers are present, with some residuals extending beyond ±4.

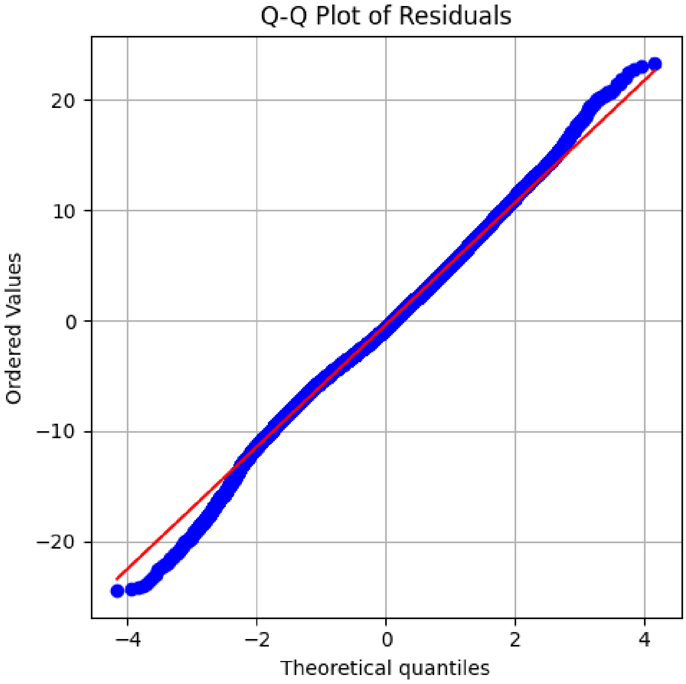


Figure 35. Normal Q-Q Plot of Residuals for the XGBoost Regressor (Hourly Demand).

The Normal Q-Q plot for the XGBoost residuals is presented in Figure 35. The deviations from the normality line are substantial, particularly at the tails, with the lower tail being notably heavier. This indicates a significant departure from a normal distribution for the residuals.

7.4 Comparison of Hourly Electricity Demand Forecasting Approaches

To determine the most effective methodology for hourly short-term load forecasting for Vaasa, the performance of the selected MLR model

(Model 1 from Section 7.1) is compared against the two machine learning models: Random Forest Regressor and XGBoost Regressor. The comparison is based on their respective performance metrics on the unseen test dataset for the year 2024. Table 18 consolidates these key performance diagnostics.

Table 18. Comparative Performance Diagnostics of Multiple Linear Regression and Machine Learning Models for Hourly Electricity Demand Forecasting (Test Set: 2024).

	MSE	RMSE	MAE	MAPE	R²
Linear Regression (Model 1)	21.400	4.626	3.666	0.066	0.901
Random Forest Regressor	23.190	4.820	3.670	0.061	0.892
XGBoost Regressor	21.220	4.610	3.610	0.063	0.902

Based on the comparative performance metrics presented in Table 18, the XGBoost Regressor emerges as the marginally best-performing model for hourly electricity demand forecasting on the 2024 test set.

The XGBoost Regressor achieved the lowest MSE (21.220), the lowest RMSE (4.610 MWh), and the lowest MAE (3.610 MWh). It also yielded the highest R² value (0.902), indicating it explained slightly more variance in the test data compared to the other models.

The MLR (Model 1) performed very competitively, with an R² of 0.901 and an RMSE of 4.626 MWh, making its performance almost identical to XGBoost in terms of these overall error metrics.

The Random Forest Regressor showed slightly higher errors (RMSE 4.820 MWh) and a lower R² (0.892) compared to both MLR and XGBoost. However, it achieved the lowest MAPE (0.061 or 6.1%).

While XGBoost shows a slight edge in most error metrics and R², the performance differences between XGBoost and MLR are minimal. The MLR model offers the advantage of greater interpretability of its coefficients. However, machine learning models like XGBoost are often favored for their potential to capture complex non-linearities without requiring explicit specification of these relationships.

The residual analyses for both machine learning models (Random Forest and XGBoost) indicated issues with heteroscedasticity and significant departures from normality, with many outliers. The MLR model also showed some heteroscedasticity and non-normality in its hourly residuals, though perhaps slightly less severe than the machine learning models in terms of the Q-Q plot's tail behavior.

Considering the slight numerical superiority in key error metrics (MSE, RMSE, MAE) and R^2, the XGBoost Regressor is selected as the best-performing model for hourly short-term load forecasting in this study. However, it is important to acknowledge that its performance is very closely matched by the MLR model. The choice in a practical application might also consider factors like model complexity, training time, and the need for direct interpretability of coefficients versus predictive power. The issues highlighted in the residual analysis for all models suggest that further refinements, such as advanced feature engineering, outlier treatment, or transformations of the target variable, could potentially enhance the performance and reliability of hourly forecasts.

8 DISCUSSION

This thesis aimed to develop and evaluate tailored long-term and short-term electricity demand forecasting models for Vaasa, Finland, identifying key influential features and comparing traditional statistical models with machine learning algorithms. The research successfully established robust models for different forecasting horizons, providing valuable insights into the drivers of electricity consumption in this Nordic city.

8.1 Achievements

The primary achievement of this research was the successful development of tailored forecasting models for Vaasa across multiple time horizons. For long-term monthly forecasting, a MLR model utilizing average and maximum temperatures demonstrated strong predictive performance, achieving an R^2 of 96.3% (Table 6). In the realm of short-term daily forecasting, a MLR model, which incorporated temperature variables, sun duration, a workday indicator, and month dummies, notably outperformed machine learning approaches like Random Forest and XGBoost, yielding an R^2 of 92.8% (Table 7). For granular hourly forecasts, an XGBoost Regressor showed a marginally superior performance with an R2 of 90.2% (Table 17), though a similarly structured MLR model performed very competitively R2 of 90.1%.

Across all models, temperature-related variables and various temporal features (such as month, day of week, hour, and sun duration/flag) were consistently identified as the most critical predictors of electricity demand in Vaasa. The successful application of engineered features like Is_workday and Sun_Flag also underscored the value of domain-specific feature creation in enhancing model accuracy. A significant finding was the strong performance of well-specified multiple linear models, challenging the common assumption of universal superiority for more complex machine learning algorithms in all contexts and aligning with literature that indicates context-dependent model efficacy. This research provides Vaasa's energy sector with practical forecasting tools for strategic planning and operational efficiency, directly addressing an identified gap for tailored models in a Nordic city context. Ultimately, this thesis successfully demonstrated that carefully constructed models can effectively forecast electricity demand in Vaasa, offering actionable insights for local energy management.

8.2 Challenges and Limitations

Despite the achievements, the research encountered certain challenges and is subject to limitations. The models' accuracy is inherently dependent on the 2020-2024 historical data used. This period may have unique characteristics, and the models may not fully capture the impact of unforeseen extreme events or long-term socio-economic shifts beyond this timeframe. As a localized study focused on Vaasa, the direct generalizability of specific model parameters to other regions with different climates or industrial structures may be limited, although the methodology remains transferable.

A notable challenge arose from model diagnostics; most models, particularly for hourly forecasts, exhibited residuals with some degree of heteroscedasticity and departures from normality. This indicates potential areas for model refinement to ensure more consistent error variance and reliable prediction intervals. Furthermore, the limited predictive power of socio-economic variables like Population, Electricity Price, and EV Count within this specific dataset posed a challenge in aligning findings with some broader long-term forecasting literature where these factors are often significant. This underscores the need for cautious interpretation if underlying conditions change or when comparing with national-level studies.

8.3 Future Work

Building on the findings and addressing the limitations of this study, several avenues for future research are recommended. Future work could focus on incorporating a wider array of features, such as more detailed weather variables (humidity, wind speed), specific local holiday indicators, or more granular data for EV charging patterns as it becomes available. Exploring advanced time series models like SARIMAX, LSTMs, or hybrid approaches may yield further improvements, particularly for the complex dynamics of hourly forecasts.

Efforts should also be directed towards addressing the observed residual patterns, possibly through data transformations, the use of robust regression techniques, or alternative error structures. Comprehensive hyperparameter optimization for the machine learning models could unlock further performance gains. Finally, developing dynamic model updating strategies to adapt to evolving consumption patterns and investigating probabilistic forecasting methods to quantify uncertainty would be valuable extensions, enhancing the practical utility and reliability of electricity demand forecasts for Vaasa.

REFERENCES

Kandananond, K. Forecasting electricity demand in Thailand with an artificial neural network approach. Energies 2011

Rahman, H., Selvarasan, I., & Begum, J.A. Short-term forecasting of total energy consumption for India-A black box based approach. Energies 2018

Kuo, P.H., & Huang, C.J. A high precision artificial neural networks model for short-term energy load forecasting. Energies 2018

López, M., Sans, C., Valero, S., & Senabre, C. Empirical comparison of neural network and auto-regressive models in short-term load forecasting. Energies 2018

Statistics Finland. (n.d.). Prices of electricity by type of consumer. Retrieved March 15, 2025, from https://pxdata.stat.fi/PxWeb/pxweb/en/StatFin/StatFin__ehi/statfin_ehi_pxt_13rb.px/sortedtable/tableViewSorted/

Statistics Finland. (n.d.). Population by month and area, 1990M01-2023M12. Retrieved March 15, 2025, from https://pxdata.stat.fi/PxWeb/pxweb/en/StatFin/StatFin__vaerak/statfin_vaerak_pxt_11s5.px/

Statistics Finland. (n.d.). Preliminary population structure by area, 2024M01-2025M02. Retrieved March 15, 2025, from https://pxdata.stat.fi/PxWeb/pxweb/en/StatFin/StatFin__vamuu/statfin_vamuu_pxt_11lj.px/

Finnish Meteorological Institute. (n.d.). Download observations. Retrieved March 17, 2025, from https://en.ilmatieteenlaitos.fi/download-observations

Time and Date AS. (n.d.). Sunrise and sunset times in Vaasa, Finland. Retrieved March 20, 2025, from https://www.timeanddate.com/sun/finland/vaasa

Finnish Transport and Communications Agency Traficom. (n.d.). Vehicles in use. Retrieved March 15, 2025, from https://trafi2.stat.fi/PXWeb/pxweb/en/TraFi/TraFi_Liikennekaytossa_olevat_ajoneuvot/040_kanta_tau_104.px/

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An introduction to statistical learning: With applications in R* (2nd ed.). Springer. https://doi.org/10.1007/978-1-0716-1418-1. Last visited May 11, 2025.

Goss-Sampson, M. A., Dr, University of Greenwich, SAGE Publications, & W H Freeman. (2020). JASP 0.14 (4th ed.). https://jasp-stats.org/wp-content/uploads/2020/11/Statistical-Analysis-in-JASP-A-Students-Guide-v14-Nov2020.pdf. Last visited April 2, 2025.

LinearRegression. (n.d.). Scikit-learn. https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html. Last visited May 11, 2025.

RandomForestRegressor. (n.d.). Scikit-learn. https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html. Last visited May 11, 2025.

Gradient Boosting regression. (n.d.). Scikit-learn. https://scikit-learn.org/stable/auto_examples/ensemble/plot_gradient_boosting_regression.html. Last visited May 11, 2025.