*Division of Computing Science and Mathematics*

*Faculty of Natural Sciences*

*University of Stirling*

Predicting Appliance Energy Use in Residential Buildings

Collins Ayidan

**Dissertation submitted in partial fulfilment for the degree of   
Master of Science in *Mathematics and Data Science***

**September 2024**

Abstract

This study intends to investigate how to predict appliance energy usage in residential buildings by employing various machine learning approaches. To achieve this, eight machine learning models—Support Vector Regression, Extra Trees, Decision Tree, Random Forest, Linear Regression, Gradient Boosting, Neural Networks and XGBoost—were trained and evaluated using the Energy usage Prediction dataset from the Machine Learning Repository of UCI. Extensive feature engineering was conducted to generate new variables. The study also utilized HalvingRandomSearchCV for hyperparameter tuning to enhance the models’ performance. Several metrics, including RMSE, R-squared (R²), MAE, and MAPE were employed to evaluate the machine learning models. The results indicated that ensemble models like Extra Trees and Random Forest provided superior RMSE and R-Squared in predicting energy consumption, outperforming other traditional models. The Extra Trees model achieved a testing RMSE of 63.28 and an R² of 0.60, while the Random Forest model had a testing RMSE of 65.51 and an R² of 0.57, significantly outperforming other traditional models.

Attestation

I understand the nature of plagiarism, and I am aware of the University’s policy on this.

I certify that this dissertation reports original work by me during my university project except for the following:

* **Several works on GeeksforGeeks were consulted for:**
* Implementing the Gradient Boosting algorithm in Section 4.2.
* Replicating the pairplots in Appendix 1.

**Signature: Collins Ayidan** **Date: 29th August 2024**

Acknowledgements

This dissertation would not have been possible without the guidance and the help of individuals who in one way or another contributed and extended their valuable assistance in preparation and completion of this study, it a pleasure to thank those who made it possible.

First, I would like to express my sincere gratitude to Prof. Kevin Swingler, Dr Nora Tanner and all staff of the standard, structured dissertation team for their invaluable patience, expertise knowledge, and feedback throughout the dissertation project.

Completion of this project could not have accomplished without the support of my cohort members, especially Enock Hagan for their knowledge and understanding of concepts that empowered me to overcome difficulties I encountered, am grateful.

I would be remiss in not mentioning my family, especially my sister Joyce Ankomah and his Husband Eric Danso, Rev. Kwarteng Boamah my pastor and Enock Boakye a friend for their immerse support and prayers. Their believe in me has kept my spirit and motivation high during this dissertation project.

Table of Contents

Abstract 1

Attestation 2

Acknowledgements 3

Table of Contents 4

List of Figures 5

List of Tables 6

1 Introduction 7

1.1 Background and Context 7

1.2 Scope and Objectives 7

1.3 Achievements 7

1.4 Overview of Dissertation 7

2 Literature Review 8

2.1 Factors Influencing Appliance Energy Use in Residential Buildings 8

2.2 Use of AI models for predicting energy use 8

2.3 Optimization and Tuning 10

2.3.1 Hyperparameter Tuning 10

2.3.2 Approaches to Feature Selection and Engineering 11

2.4 Conclusion 11

3 Methodology 12

3.1 Data Type and Source 12

3.2 Data Preprocessing 12

3.3 Modelling 12

3.4 Training and Testing Procedure 14

3.4.1 Model Evaluation 15

3.4.1.1 Root Mean Square Error (RMSE) 15

3.4.1.2 Coefficient of Determination (R-squared, *R2*) 15

3.4.1.3 Mean Absolute Error (MAE) 15

3.4.1.4 Mean Absolute Percentage Error (MAPE) 16

3.5 Technologies Used 16

4 Results 17

4.1 Exploratory Data Analysis 17

4.2 Models Discussion 18

4.2.1 Important Features of the Extra Trees model 20

4.2.2 Partial Dependence 21

4.2.3 Residual Distribution 22

4.2.4 Residual Plots 23

4.2.5 Cross-Validation RMSE Distribution 24

4.2.6 Bootstrapped RMSE Distribution 25

5 Conclusion 27

5.1 Summary 27

5.2 Evaluation 27

5.3 Future Work 27

References 28

Appendix 1 31

List of Figures

[Figure 1. Appliance Energy Consumption with Day of the Week and Hour of the Day 19](#_Toc175384224)

[Figure 2. Correlation Heatmap 19](#_Toc175384225)

[Figure 3. Feature Importance from the Extra Trees Model 23](#_Toc175384226)

[Figure 4. Partial dependence of top 6 features 24](#_Toc175384227)

[Figure 5. Residuals Distribution 24](#_Toc175384228)

[Figure 6. How the Residuals change with the NSM 25](#_Toc175384229)

[Figure 7. Distribution of Cross Validation Scores 26](#_Toc175384230)

[Figure 8. Bootstrapped RMSE distribution 27](#_Toc175384231)

[Figure 9. Distribution of Appliance Energy Consumption 33](#_Toc175384232)

[Figure 10. Appliance Energy Consumption: All time and Weekly plot 33](#_Toc175384233)

[Figure 11. Pair Plot Set 1 33](#_Toc175384234)

[Figure 12. Pair Plot Set 2 34](#_Toc175384235)

[Figure 13. Pair Plot Set 3 35](#_Toc175384236)

[Figure 14. Pair Plot Set 4 36](#_Toc175384237)

List of Tables

[Table 1: Model performance – Training and Testing sets. 19](#_Toc175386755)

[Table 2: Model performance from Original Study – Training and Testing sets. 19](#_Toc175386756)

# Introduction

## Background and Context

Energy use in residential buildings represents a major portion of total energy consumption, impacting both economic costs and environmental consequences. Therefore, accurately predicting appliance energy use is vital for developing effective energy management strategies and enhancing energy efficiency. Building on prior research [1] [2], this project utilizes machine learning techniques to forecast appliance energy usage in residential settings, with a focus on identifying the algorithm that best predicts this consumption. This study also aims to deepen the understanding of how various machine learning models can predict energy consumption, offering insights into which algorithms are most effective in this context. The findings could inform the design of smarter, more energy-efficient residential buildings and contribute to the promotion of sustainable living practices. project employs eight machine learning techniques to predict appliance energy use in residential buildings with the objective on determining the algorithm that best predicts this usage. The study is also significant as it enhances the understanding of how different machine learning models can predict energy consumption, providing insights into which algorithms are most effective in this context. The findings can inform the design of smarter, more energy-efficient residential buildings and add to the advancement of energy-friendly living practices.

## Scope and Objectives

The basic objective of this study is to predict appliance energy consumption accurately with emphasis on model improvement. The study specifically aims to:

1. To find the model that best predicts appliance energy usage.
2. To introduce new models – Decision Trees and Random Forest models – and find out if they can outperform the models machine models used in [1], [2].
3. To introduce additional feature engineering in the UCI Appliance energy Prediction data and determine if they are important in influencing appliance energy consumption.

## Achievements

This study offers several valuable contributions to the domain of energy consumption prediction in residential buildings. It not only introduces new machine learning models but also evaluates their performance against existing models. Additionally, it explores the impact of feature engineering on model accuracy, providing valuable insights into factors that influence appliance energy consumption.

## Overview of Dissertation

The dissertation starts with an exploration of existing literature on energy consumption prediction and the application of AI techniques. It then details the methodology used in this study, including the data preprocessing, model selection, and feature engineering processes. The subsequent chapters present the results of the model evaluations, followed by a discussion on the implications of these findings. The dissertation wraps up with directions for future research.

# Literature Review

## Factors Influencing Appliance Energy Use in Residential Buildings

Building characteristics and household compositions significantly influence residential energy consumption. Xie and Noor [3] conducted a comprehensive analysis using multiple regression models to investigate factors that affect residential energy usage, focusing on building features, household compositions, lifestyles, and equipment used at home. Their study identified the area of the floor and the total number of members in the family as key factors affecting energy use for various purposes such as cooling and appliances. Similarly, Rickwood [4] and Kavousian et al. [5] utilized detailed analyses controlling for household demographic and income variables to demonstrate that dwelling type, household size, and household income significantly influence energy consumption.

Another important determinant of residential energy consumption in the literature is appliance ownership and usage patterns. Leahy and Lyons [6] employed logit and OLS analyses on household level dataset to show how household characteristics explain appliance ownership and its effect on energy demand. They noted that the type and number of appliances owned by a household directly impacts overall energy use. Additionally, studies by Cetin et al. [7] and Kavousian et al. [8] analyzed disaggregated energy use data and energy efficiency frontiers to highlight that usage patterns, such as the time of day and frequency of use, significantly affect energy consumption. These studies found that user-dependent appliances, like washers and dryers, vary greatly in their energy use patterns, which can influence peak demand and overall energy consumption.

Behavioral factors of household occupants are also found to be important in influencing residential energy consumption. Rouleau et al. [9] conducted a case study of a high-performance Canadian social housing building, using regression analysis and monitoring systems to assess energy use and occupant behavior. They found that occupant behavior, such as the frequency of opening windows or the use of electrical appliances, significantly impacts energy use, often outweighing structural factors. Households with higher education levels and those that actively track their energy usage tend to be more energy-efficient, demonstrating the importance of awareness and proactive energy management. This was supported by Kavousian et al. [8] through their analysis of smart meter data and energy efficiency rankings.

Technological advancements and environmental conditions have also been found to play significant roles in residential energy consumption. Iwayemi et al. [10] discussed the potential of smart consumption strategies and energy management systems to enhance energy efficiency without compromising living standards. They highlighted the benefits of advanced technologies and energy-efficient appliances in reducing energy consumption. Reyna and Chester [11] conducted a study using climate change projections to measure the demand for electricity and natural gas in residential homes at Los Angeles. They found that change in the climate could increase residential energy demand substantially, and therefore recommended the need for aggressive energy efficiency policies to mitigate this impact.

## Use of AI models for predicting energy use

Candanedo et al. [1] and Assadian & Assadian [2] applied various machine learning approaches, such as gradient boosting machines (GBM), random forest, and extra trees regressor, to predict appliance energy usage based on data-driven approaches. The studies demonstrated that Gradient Boosting Machines (GBM) and Extra Trees Regressor achieved high accuracy in forecasting appliance energy consumption. The study employed metrics such as RMSE, R-Squared (R²), MAE, and MAPE to assess how best these models perform. Ullah et al [12] studied how to predict energy consumption of appliance in low-energy homes using machine learning algorithms. Their study benchmarked eight models, including linear regression, ridge regression, LASSO regression, nearest neighbour regression, support vector machine, multilayer perceptron, extra trees, and XGBoost models. The study evaluated these algorithms based on metrics such as the RMSE, coefficient of determination, MAE, and the time taken to train the models. Extra-Trees and XG-Boost emerged as the top-performing algorithms while demonstrating efficient performance across the specified error metrics. Duarte et al. [13] estimated various machine learning algorithms, such as ANN, Random Forests, and Support Vector Regression (SVR) and conducted an evaluation of these models' effectiveness. The study concluded that ANN combined with interaction variables offered the best prediction accuracy. This study also employed metrics such as MAE, RMSE, R-Squared (R²), and MAPE to evaluate the performance of the machine learning models. Rambabu et al. [14] trained several machine learning algorithms including Linear Regression, Lasso Regression, Random Forest, Extra Trees Regressor, and XGBoost, for predicting household energy consumption. Evaluation was conducted using R-squared to assess predictive accuracy based on time-series data. The study concluded that tree-based models offer superior performance in forecasting household energy consumption patterns influenced by factors like temperature, humidity, and time of day.

Edwards et al. [15] evaluated seven machine learning algorithms for predicting hourly residential energy consumption using 15-minute interval data. Among the methods evaluated, Least Squares Support Vector Machines (LS-SVM) demonstrated superior performance compared to other algorithms, including Neural Networks. Jain et al. [16] utilized SVR to predict energy consumption in residential buildings of several families and found that the floor level with intervals per hour is the most effective monitoring granularity. Long Short-Term Memory (LSTM) networks showed superior performance with the highest R2 (0.97) and lowest RMSE (21.36) over traditional machine learning models in testing set [17].

Studies have also utilized Deep learning models to study appliance energy usage. [18] employed Conditional Restricted Boltzmann Machine (CRBM) and Factored CRBM to predict energy usage in time series data. The study found that, the CRBM models outperformed traditional machine learning methods like SVM and ANN in predicting energy consumption in time series. The research used RMSE and the spearman correlation coefficient as the primary evaluation metrics. WaveNet models have been effective for energy disaggregation, outperforming other deep learning methods in terms of error measures and computational cost [19]. Bourhnane et al. [20] utilized Artificial Neural Networks (ANN) along with Genetic Algorithms to predict energy consumption in smart buildings. The ANN model demonstrated a modest prediction accuracy. Mocanu et al. Ngo et al. [21] proposed an ensemble machine learning model combining ANN, SVR, and M5Rules for predicting the amount of energy used in non-residential buildings. Their ensemble approach significantly improved prediction accuracy compared to individual models. A study by Al-Rakhami et al. [22] employed XGBoost to predict the intensity of heating and cooling loads in residential buildings, achieving high prediction performance.

Nambiar et al. [23] investigated the use of deep learning algorithms for predicting household appliance electricity consumption. Their study applied models like SVR, k-Nearest Neighbour (kNN), Decision Tree Regression (DTR), Fully Connected Neural Network, and Long Short-Term Memory (LSTM). The effectiveness of the algorithms was measured using RMSE and MAE to assess prediction accuracy. Similarly, Sajjad et al. [24] introduced a hybrid form of CNN-GRU model specifically for immediate residential load prediction. The CNN-GRU model was evaluated using RMSE, MAE, and MAPE metrics, showing superior performance over traditional models such as Gradient Boosting Regression (GBR), ANNs, Extreme Learning Machine (ELM), and SVM.

Ibrahim et al. [25] utilized multiple linear regression (MLR) and multilayer perceptron (MLP) techniques to investigate the amount of heating load (HL) and that of cooling load (CL) in residential buildings. Their findings indicated that MLP produced highly accurate estimates, with low MAE and RMSE, and high R² values. Moradzadeh et al. [26] introduced a hybrid model that combines Group Method of Data Handling (GMDH) with SVR to predict Cooling Load as well as Heating Load. These models showed superior performance, outperforming traditional regression models. The study assessed the models' performance using the R-Squared (R²), MAE, MSE, and RMSE. Wu et al. [27] conducted a comparative study on various machine learning models, including Gradient Boosting Regressor and HB-Regressor, for predicting heating and cooling loads. To determine which model performed best, the study used the R², MAE, and MSE as evaluation metrics. In addition to these metrics, the Accuracy and the Error rate were also employed to assess the overall performance. The study found that Gradient Boosting Regressor provided high prediction accuracy with efficient fitting speeds. [28] assessed how effective several machine learning models are, for predicting the amount of energy used in smart homes. Several evaluation metrics such as RMSE, R-Squared (R²), and the MAE were employed to compare the models’ performance. Based on the findings, the random forest model demonstrated superior performance and therefore emerged as the best algorithm for estimating energy consumption of appliances in smart homes.

Tran et al. [29] introduced the Evolutionary Neural Machine Inference Model (ENMIM) hybrid model to forecast appliance energy usage and the maximum energy demand. Their hybrid model demonstrated superior predictive accuracy compared to other benchmark models. Hybrid methods combining clustering (e.g., k-medoids) and machine learning models (e.g., SVM, ANN) have achieved high accuracy (99.2%) in forecasting appliance consumption and peak demand [30]. Zaini [31] discussed the use of a Feature Optimization Prediction Framework (FOPF) and KNN models for estimating the amount of appliance energy usage in low-energy buildings. The study underscores the effectiveness of K-Nearest Neighbors (KNN) models, which demonstrated the highest accuracy and the lowest error (RMSE = 0.0078) among the machine learning algorithms tested.

Ahmed Al-Adaileh and S. Khaddaj [32] proposed a comprehensive smart energy management system that leverages various machine learning r

egression techniques to forecast and schedule the operating times of appliances in households. The system integrates data from the surrounding environment to enhance energy efficiency, achieving up to a 36% reduction in energy consumption. Iram et al. [33] proposed a decision algorithm model utilizing machine learning-based data mining and picture fuzzy operators. The study assessed the performance of machine learning algorithms using accuracy metrics and employed a decision matrix with fuzzy operators to aggregate and rank the algorithms based on their predictive capabilities. The approach integrates Lasso Regression to analyze weather patterns and features affecting smart home micro-climates, providing insights into appliance electricity consumption and overall energy usage.

## Optimization and Tuning

### Hyperparameter Tuning

GridSearchCV systematically explores a specified range of hyperparameters by evaluating all possible combinations. For example, it has been used effectively in tuning models like Random Forests and Support Vector Machines [1], [2], [34], [35]. Although exhaustive, it can be computationally expensive [36]. Random Search samples hyperparameters randomly from a specified distribution, and it is often more efficient than grid search for identifying optimal configurations within a limited timeframe [36].

Bayesian optimization has been applied successfully in optimizing hyperparameters for neural networks and gradient boosting models [37]. Similarly, evolutionary algorithms have been employed to optimize models like Random Forests and neural networks, enhancing their performance through iterative tuning [38]. According to [39], evolutionary algorithms use mutation, crossover, and selection to optimize hyperparameters. Multi-objective optimization techniques, such as binary grey wolf optimization, have also been employed to enhance how models like Random Forest and K-NN perform by optimizing feature selection [40].

### Approaches to Feature Selection and Engineering

A crucial aspect of this study is selecting the best features and creating new ones from existing features to enhance model performance. This allows us to concentrate on the most relevant features and reduce dimensionality. A common technique in literature is recursive feature elimination (RFE). Recursive Feature Elimination (RFE) works by systematically removing less important features based on the model’s performance, thereby improving both efficiency and accuracy [41]. This technique is especially beneficial for linear regression and support vector machines, where selecting the most impactful features can significantly enhance model accuracy and efficiency.

**Another common approach is** the use of principal component analysis (PCA). **Principal Component Analysis (PCA)** reduces dimensionality by transforming features into orthogonal components, allowing it to capture the most variance with fewer features [42]. PCA has been used effectively in logistic regression and neural network models to improve computational efficiency and accuracy [43].

## Conclusion

The literature review provides several key insights that are more relevant for this study. Existing studies have highlighted the significant influence of building characteristics and household composition, appliance ownership and usage patterns, and environmental conditions on energy consumption. This informs the study of the relevance of including these variables or their proxies such as temperature and humidity across different rooms and the number of occupants as well as specific variables related to appliance usage, such as energy consumption by appliances and lights. External weather conditions, including temperature outdoor as well as humidity, and wind speed are also included to represent the role of environmental factors on energy consumption.

The use of machine learning approaches in predicting energy consumption is a significant topic in the literature. Several machine learning algorithms, such as Random Forest, GBM, SVR, and Neural Networks (NN), are recognized for their effectiveness in accurately predicting energy consumption. This study employs a wide array of models, such as GBM, RF, SVR, and NN, to predict energy consumption, which is consistent with the literature’s recommendations on the effectiveness of these models.

The literature emphasizes the importance of hyperparameter tuning and feature engineering for optimizing model performance. Techniques such as RandomSearchCV, and Bayesian optimization are highlighted as critical for enhancing model accuracy and efficiency. This study implements HalvingRandomSearchCV for efficient hyperparameter tuning and employs feature engineering to refine the input variables, directly informed by these best practices from the literature. Existing studies have also identified key evaluation metrics including RMSE, MAE, R2, and MAPE for assessing model performance particularly in this area of study.

# Methodology

This chapter details the methodologies and techniques used in the project to predict appliance energy use in residential buildings. The chapter covers the types and sources of data, a brief description of the variables involved, the estimation techniques and evaluation metrics employed, the technologies used, data preprocessing steps, and the modelling procedures.

## Data Type and Source

The study employed secondary dataset from the UCI Machine Learning Repository [44], specifically the "Appliances Energy Prediction" dataset. This dataset contains 19,735 instances of measurements related to energy consumption in residential buildings, collected over a period. The data includes various features such as temperature, humidity, weather conditions, and electrical energy usage of appliances.

## Data Preprocessing

The study initially explores the dataset to obtain knowledge about the distribution of the data. This included exploring the relationships between variables and validating the data. The study also conducted a correlation to help identify highly correlated variables, which was useful for feature selection and engineering.

In addition to the variables derived from the original data, which included the total number of seconds away from midnight, the classification of days as weekends or weekdays, and specific days of the week, further feature engineering was implemented using the timestamp data provided and following the outline by [2]. This approach aimed to extract additional insights from the temporal information that was not initially integrated into the modelling process. New features created included hour, month, and day features. This was extracted from the timestamp to capture temporal variations in energy consumption patterns. Another feature is the seasonal categorical data, which was created based on the month, categorizing data into autumn, winter, spring, or summer.

## Modelling

The study involved training and evaluating eight distinct machine learning models to predict appliance energy usage. These models included: the linear regression (LM) model, decision tree (DT), support vector regression (SVR), neural network (NN), random forest (RF), gradient boosting (GB), XGBoost (XGB), and the extra trees model (ET).

The Linear Regression (LM) is a straightforward but powerful statistical technique employed to determine the linear association between the dependent variable—here, appliance energy consumption—and one or more independent variables. LM helps to explain how changes in the features used influence the appliance energy usage by fitting a linear equation to the observed data. SVR, on the other hand, extends this concept by leveraging support vectors to map the data fed into the model into a higher-dimensional feature space. This approach allows SVR to capture more complicated associations that might not be clear in the dataset. The primary aim of SVR is to find a hyperplane that best separates the forecasted values from the real values, ensuring that the model provides an optimal fit with minimal error. The "support vectors" are essentially the data points that are nearer to this hyperplane, and they play an important role in defining how to place the hyperplane, thereby impacting the accuracy of the model.

SVR is particularly effective in scenarios where the association existing between the input features and the output is not linear, as it can adapt to non-linear data by applying kernel functions. These kernels enable SVR to find complex patterns and interactions within the data, making it a versatile tool in machine learning. The ability to balance model complexity with generalization is what makes SVR a popular choice for tasks involving energy consumption prediction.

Decision Tree (DT) models operate on the principle of breaking down the data repeatedly into increasingly smaller subsets according to input feature values. The goal is to partition the dataset in such a way that each subset is as homogenous as possible with respect to the target variable. This process continues until a predefined stopping condition is fulfilled, such as achieving the maximum tree depth or when additional splitting no longer significantly improves the model's accuracy. While Decision Trees are intuitive and easy to understand and interpret, they are likely to be overfitted, especially when they become too complicated. This is where Random Forest comes in. Random Forest (RF) is an ensemble learning technique that builds a multitude of decision trees during training, each using a randomly selected subset of features. This method capitalizes on the diversity of the trees, where each tree is trained on a unique subgroup of the dataset and features. The process begins by selecting random samples from the dataset, which are used to build individual decision trees. Within each tree, the best split is determined based on the gain in information, typically measured by metrics like Gini impurity or entropy. This helps the tree decide how to break the data at each point, optimizing the classification or regression task at hand. The important strength of the Random Forest algorithm is in its potential to limit the possibility of overfitting, which is a likely problem with individual decision trees. Random forest computes the average prediction from the aggregates of the individual decision trees. This aggregation of predictions leads to more accurate and stable predictions compared to relying on a single decision tree. Additionally, because each tree is exposed to a subset of the dataset that differs from each other, the model becomes less sensitive to noise, further enhancing its predictive power.

Gradient Boosting (GB) builds a series of decision trees. The model aims to minimize a loss function, such as mean squared error (MSE), by progressively adding more decision trees, each one correcting the errors of its predecessors, until the loss is reduced to an acceptable level. This iterative process ensures that the model becomes increasingly accurate as more trees are added, based on the correct predictions with each iteration. XGBoost (XGB), an advanced implementation of gradient boosting, enhances this process by incorporating several regularization methods to limit the likelihood of overfitting and therefore improve generalization. These techniques include L1 and L2 regularization, which provide a penalty factor for highly complicated models by adding constraints to the parameters of the model. This helps in maintaining a balance between model complexity and its potential to predict unseen data. Moreover, XGBoost includes tree pruning, which removes branches that add little to no predictive power, thereby streamlining the model and making it more efficient. Early stopping is another critical feature of XGBoost; it stops the training process if the performance of the model on a validation set ceases to improve, to prevent the model from overfitting the training dataset. Additionally, XGBoost supports parallel tree boosting to enhance computational efficiency.

*Extra Trees (ET)* are more related to random forests but select splits randomly without considering split quality. This method accelerates the training process and increases tree diversity, improving model generalization. Neural Network (NN) models are composed of interconnected layers of nodes, commonly referred to as neurons. Each connection between neurons has a corresponding weight that allows for the identification of the strength and direction of the signal transmitted. In the event of training the algorithm, these weights are systematically adjusted to reduce the errors that are likely to happen between the forecasted values and the real values. This iterative adjustment is typically guided by optimization algorithms such as gradient descent, which fine-tunes the weights to reduce the overall loss function. Neural networks are particularly powerful because they can model complicated relationships in the data including non-linear associations, something that simpler models like linear regression might struggle with. The layers of neurons in a neural network allow it to learn multiple levels of abstraction, capturing complex patterns and dependencies within the data.

## Training and Testing Procedure

The dataset was divided into distinct training and testing sets to assess how well the model performs on seen and unseen data. Each of the eight regression models underwent 10-fold cross-validation with 3 iterations, a technique designed to ensure comprehensive evaluation. In this approach, the dataset is split into 10 distinct groups, with each group serving as the validation set in one iteration while the others are utilized as training set. This continuous process is iterated 10 times, allowing every subset to act as the validation set at least once [2], [45]. The final performance is averaged across all 3 iterations. This offers a robust estimate of the performance of the model by leveraging diverse data rather than relying on a single randomized split.

Hyperparameter tuning was performed using *HalvingRandomSearchCV*. Unlike a standard RandomSearchCV, where combinations are randomly selected evaluated exhaustively, *HalvingRandomSearchCV* is more efficient. It starts by testing all the randomly selected hyperparameter combinations on a small subset of the data, then it continuously narrows down the pool of combinations by eliminating the worst-performing ones while increasing the amount of data used in each step. This halving process continues until only the best-performing combinations remain, which are then evaluated more thoroughly. This approach reduces the computational cost and provides a more efficient path to finding the optimal hyperparameters for the model. In addition to the hyperparameters, the study employed the Bootstrapping approach to visualize a histogram of how the RMSE changes in different subsets of the data using the best performing model. The shape of this distribution reveals the model's sensitivity to changes in the data. A narrow distribution, where most RMSE values cluster around a central point, indicates that the model is stable and consistently performs well across different samples. In contrast, a wider distribution suggests that the efficiency of the algorithm may fluctuate depending on the specific features of the dataset it is being trained on. This variability could highlight areas where the model is less robust, potentially guiding further refinements to improve its generalizability. The bootstrapping approach, therefore, provides a better understanding of model reliability, complementing traditional cross-validation techniques and offering greater confidence in the potential of the model to perform well in diverse real-world scenarios.

For Support Vector Regression (SVR), the critical hyperparameters gamma, cost and kernel were tuned. The kernel type (rbf or linear) in Support Vector Regression (SVR) determines how the data is transformed into a higher-dimensional space, making it easier to separate into classes. Gamma influences the shape of the decision boundary, while cost manages the trade-off between training error and testing error. For SVR, the optimal values determined were 1 for gamma and 1 for cost, with the linear kernel proving to be the best. For the Random Forest (RF) model, the tuning process focused on the number of trees (estimators) and the number of features considered for splits (max features). The optimal configuration for Random Forest model was determined to be 500 estimators and a max depth of None. In addition to the number of estimators and the max depth parameters, the minimum samples split was also fine-tuned for the Extra Trees (ET) model to optimize its performance. The optimal number of estimators were found to be 100, minimum samples split of 5 and maximum depth of 20 for ET model.

Like the Random Forest model, the Gradient Boosting (GB) model also tuned the number of estimators and the max\_depth parameters. The optimal values identified were 500 for the number of estimators and 2 for max\_depth.

Linear Regression (LM) does not typically require hyperparameter tuning but was included as a baseline model for comparison. The model fits a linear equation to the observed data, using all available features.

The additional models added to the work of [2] are the DT and the NN models. For the DT model, hyperparameters were carefully tuned to optimize performance while avoiding overfitting. The optimal configuration was determined to be a maximum depth of 20, with a criterion of 'absolute\_error,' a minimum of 4 samples per leaf, and a minimum of 2 samples per split. For the Neural Network model, several key hyperparameters were adjusted to achieve the best performance. The optimal parameters included an activation function of 'logistic', an alpha value of 0.0001 for regularization, and a single hidden layer with 100 neurons. Additionally, the learning rate was set to 'adaptive', with an initial learning rate of 0.01. These settings were selected to ensure the neural network could effectively learn from the data without overfitting, while also adapting the learning rate as training progressed.

### Model Evaluation

To assess and compare the performance of each model in predicting appliance energy consumption, several metrics were employed:

#### Root Mean Square Error (RMSE)

RMSE quantifies the magnitude of the mean error between the predicted and real values. This throws light on the accuracy of the model. A lower RMSE value suggests better model performance, as it reflects smaller discrepancies between the predictions and the actual outcomes. The equation for RMSE is stated as follows:

#### Coefficient of Determination (R-squared, *R2*)

R-squared measures the proportion of the variations occurring in the dependent variable (appliance energy usage) that can be predicted by the independent variables (X) in the model. A high R-squared value indicates a better fit. Specifically, values closer to 1 suggest that a greater part of the variation in the dependent variable is explained by the model, indicating a more accurate fit. The R-squared formula is defined as follows:

#### Mean Absolute Error (MAE)

MAE measures the average magnitude of the errors in a set of predictions, without considering their direction. Like RMSE, a lower MAE indicates better predictive accuracy, as it reflects the average size of the errors, irrespective of whether they are positive or negative.

#### Mean Absolute Percentage Error (MAPE)

MAPE, or Mean Absolute Percentage Error, represents the average absolute percentage error between predicted and actual values, expressed relative to the actual values. It provides insight into the accuracy of predictions in terms of percentage. MAPE provides a measure of prediction accuracy in terms of relative error percentage. A lower MAPE indicates that the model's predictions are closer to the actual values. The formula for MAPE is defined as follows:

## Technologies Used

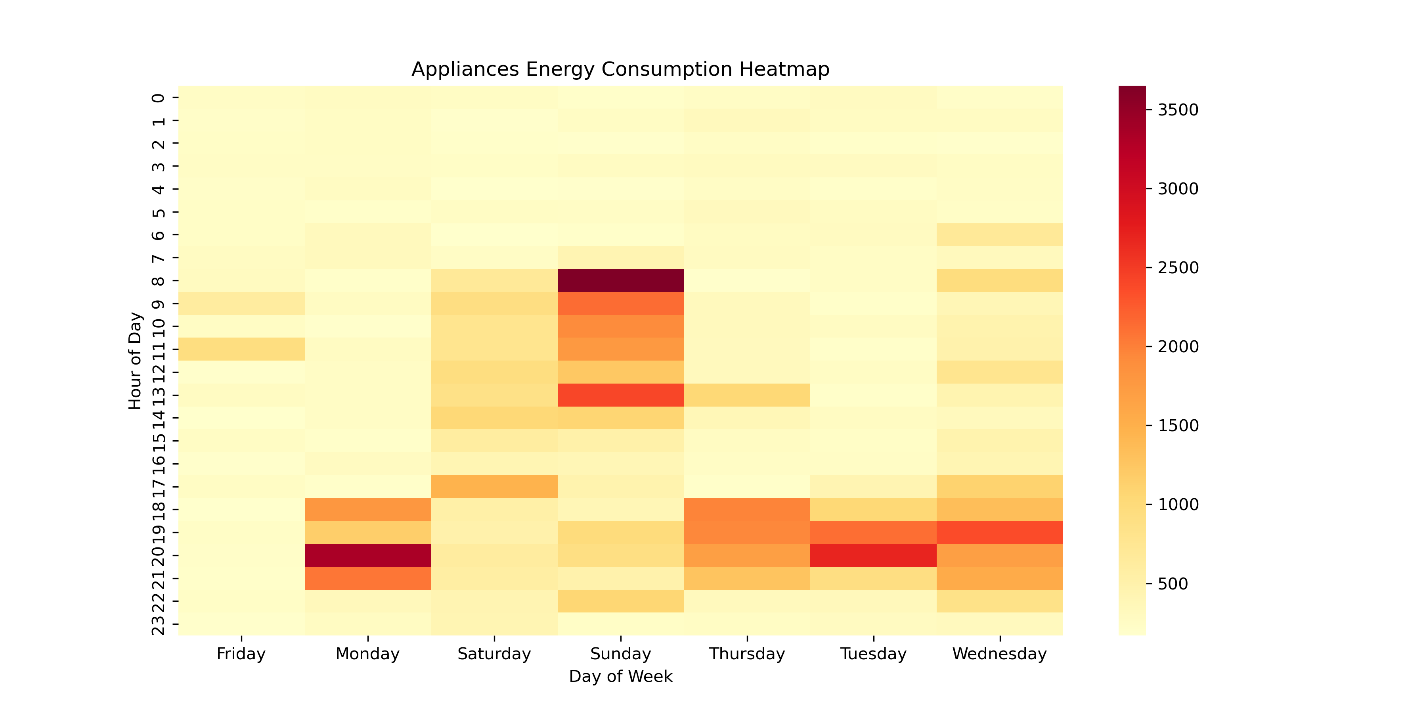
The project utilizes Python and its robust libraries to conduct thorough data analysis and implement models. Essential libraries include Pandas, used for efficient data manipulation and analysis, and Matplotlib alongside Seaborn, which are employed for visualizing the data. Scikit-learn plays a significant role in implementing various machine learning algorithms, enabling tasks like model training, evaluation, and fine-tuning of hyperparameters. Furthermore, XGBoost is specifically applied for implementing gradient boosting models.

# Results

## Exploratory Data Analysis

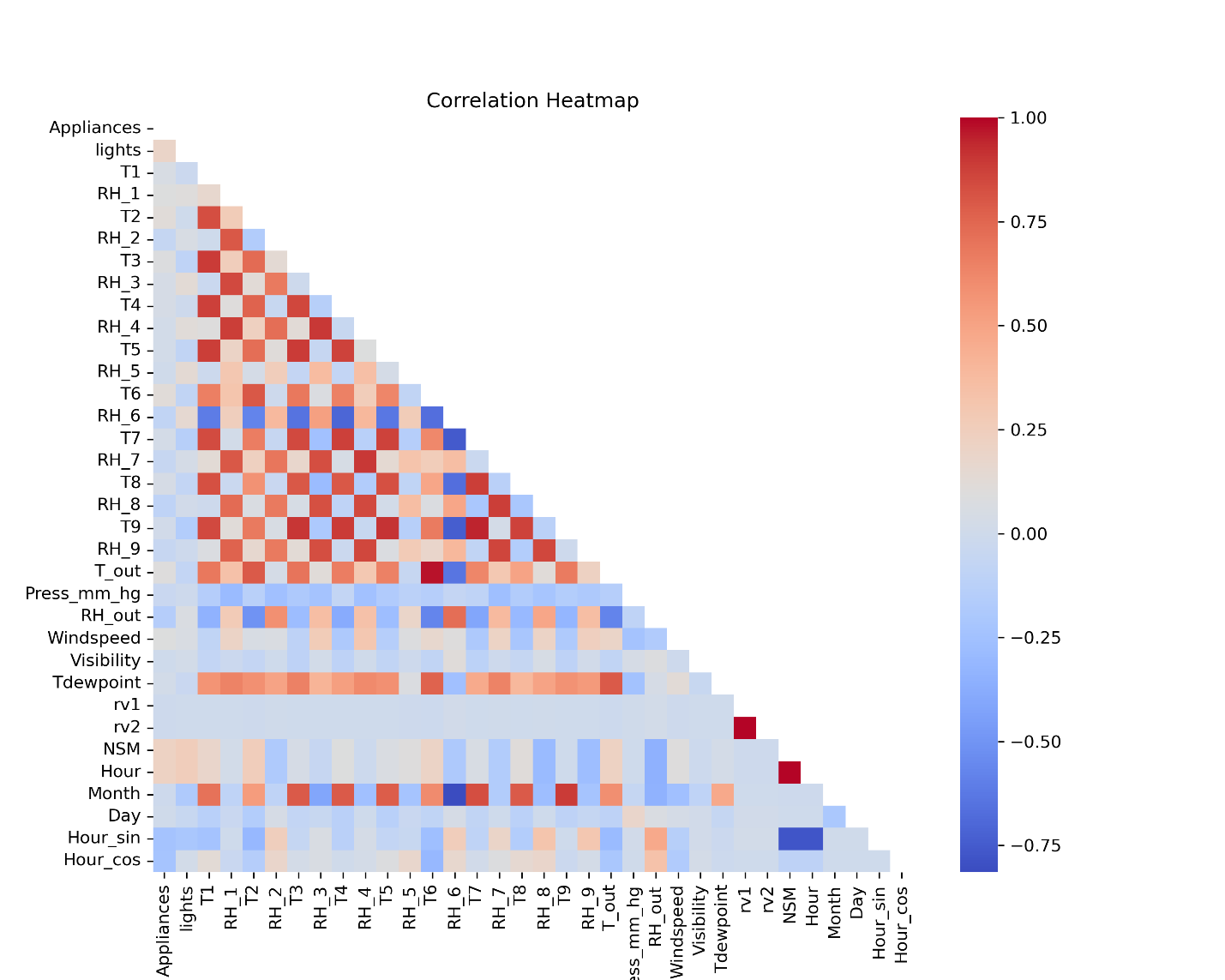
While most of the figures replicating the visuals of [1] are presented in APPENDIX 1, few relevant figures are discussed here. Figure 1 shows how hour of the day and day of the week determines the amount of appliance energy used. The heatmap shows peak energy consumption during specific hours on certain days, particularly in the evenings of Sunday, Tuesday, and Thursday, as well as mid-morning on Sunday. These peaks suggest that household activities requiring significant energy usage, such as cooking or laundry, are more concentrated during these periods.

1. Appliance Energy Consumption with Day of the Week and Hour of the Day



The correlation heatmap presented in Figure 2 shows strong positive correlations between indoor temperatures and corresponding humidity levels within the same rooms. Additionally, variables such as Appliances and lights exhibit positive correlations with time-related features like Hour and NSM, suggesting that energy usage tends to peak during specific periods of the day. Additionally, indoor temperature variables in certain rooms, such as the kitchen (T1) and living room (T2), also show a positive correlation with Appliances, implying that as these rooms get warmer, there may be increased usage of appliances, possibly due to heating or cooling systems. Conversely, external weather factors, particularly RH\_out (humidity outside), show a negative correlation with some indoor temperature variables.

1. Correlation Heatmap



## Models Discussion

The performance of eight machine learning models—Multiple Linear Regression (LM), Support Vector Regression (SVR), Random Forest (RF), Decision Tree (DT), Gradient Boosting (GB), XGBoost (XGB), Extra Trees (ET), and Neural Network (NN)—is detailed in Table 2. These models were evaluated using key evaluation such as RMSE, MAE, and the R-Squared (R²).

The outcomes from the performance of the various models used in this study are presented in Table 2. The results are compared to the original findings in [1], [2] in Table 3. The results show some differences in the models’ performance which can be associated with several factors such as data processing techniques, hyperparameter tuning, and the specific implementations of the various models estimated. Models in the original study such as the Linear Regression (LM), SVR, RF, and GBM models are compared with [1]. Similarly, the results from XGBoost (XGB) and Extra Trees (ET) are compared to [2] and finally, the newly added models — Neural Network (NN) and Decision Tree (DT) — are compared to the individual results. This comparison allows us to evaluate their performance and effectiveness in predicting outcomes, using key metrics to determine which model offers the best predictive accuracy.

Table 1: Model performance – Training and Testing sets.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Training RMSE | Training R2 | Training MAE | Training MAPE % | Testing RMSE | Testing R2 | Testing MAE | Testing MAPE % |
| LM | 91.95 | 0.21 | 51.89 | 58.99 | 89.39 | 0.20 | 51.95 | 61.74 |
| SVM | 58.49 | 0.68 | 17.35 | 11.75 | 75.14 | 0.43 | 33.91 | 30.38 |
| GBM | 83.39 | 0.35 | 45.45 | 49.82 | 83.92 | 0.29 | 46.48 | 52.84 |
| RF | 25.31 | 0.94 | 12.02 | 11.80 | 65.51 | 0.57 | 31.37 | 31.44 |
| XGB | 2.14 | 1.00 | 1.55 | 2.46 | 64.36 | 0.58 | 31.84 | 32.51 |
| ET | 21.15 | 0.96 | 9.91 | 9.89 | 63.28 | 0.60 | 29.86 | 29.68 |
| DT | 86.46 | 0.30 | 34.63 | 27.31 | 90.53 | 0.18 | 38.69 | 31.84 |
| NN | 74.59 | 0.48 | 40.93 | 44.25 | 78.51 | 0.38 | 43.41 | 47.89 |

Table 2: Model performance from Original Study – Training and Testing sets.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Training RMSE | Training R2 | Training MAE | Training MAPE % | Testing RMSE | Testing R2 | Testing MAE | Testing MAPE % |
| LM | 93.21 | 0.18 | 53.13 | 61.32 | 93.18 | 0.16 | 51.97 | 59.93 |
| SVM | 39.35 | 0.85 | 15.08 | 15.6 | 70.74 | 0.52 | 31.36 | 29.76 |
| GBM | 17.56 | 0.97 | 11.97 | 16.27 | 66.65 | 0.57 | 35.22 | 38.29 |
| RF | 29.61 | 0.92 | 13.75 | 13.43 | 68.48 | 0.54 | 31.85 | 31.39 |
| XGB\*\* |  |  |  |  | 63.86 | 0.61 | 30.24 | 29.78 |
| ET\*\* |  |  |  |  | 59.61 | 0.66 | 26.62 | 25.37 |

\*\* From [2]. Only presented testing performance.

The linear regression model in the original studies showed a training RMSE of 93.21 and a testing RMSE of 93.18, with a corresponding training R-Squared of 0.18 and testing R-Squared of 0.16. This study found an improvement in these metrics, with a better training RMSE of 91.95 but a significantly improved testing RMSE of 89.39. The R² values also saw an improvement, with a training R-Squared of 0.21 and a testing R-Squared of 0.20. These improvements suggest that the modifications made, such as feature engineering or model optimization techniques, had a positive effect on the efficiency of the LM model.

For the SVR model, there is a notable divergence between the original results and that of this work. The original model performed exceptionally well with a training RMSE of 39.35 and a testing RMSE of 70.74, along with a high training R² of 0.85 and a testing R² of 0.52. In contrast, the results of this work indicated a much higher training RMSE of 58.49 and a testing RMSE of 75.14, with a significantly lower training and testing R² values of 0.68 and 0.43 respectively. This stark difference could be associated with various factors, such as variations in kernel selection, hyperparameter settings (like gamma and C values), or even the scale of the data. It appears that the SVM model struggled to generalize well, possibly due to overfitting during training or inadequate feature scaling, which often plays a crucial role in SVM performance. This result highlights the need for more refined hyperparameter tuning and possibly exploring different kernels to enhance model performance.

When comparing the Gradient Boosting Machine (GBM), the original results indicated strong performance with a training RMSE of 17.56 and a testing RMSE of 66.65. However, the GBM results obtained showed a significant decrease in performance, with a training RMSE of 83.39 and a testing RMSE of 83.92. The decrease in the training R² values from 0.97 in the original to 0.35 and the testing R2 from 0.57 to 0.29 in this work suggests that the algorithm was unable to capture the complexity of the data as effectively. This could be due to differences in hyperparameter optimization, like the number of boosting rounds, learning rate, or maximum tree depth. Additionally, it is possible that changes in feature engineering or data preprocessing steps in this work led to less informative features being used, thereby reducing the model's predictive accuracy.

For the Random Forest (RF) model, the original results showed a solid performance with a training RMSE of 29.61 and a testing RMSE of 68.48, with a training R² of 0.92 and a testing R² of 0.54. The results from this study, however, demonstrated a significantly improved performance with a lower training RMSE of 25.31 and a testing RMSE of 65.51, accompanied by a training R-Squared of 0.94 and testing R-Squared of 0.57. This improvement suggests that the adjustments made, possibly in the form of better hyperparameter tuning (like the number of trees or maximum features) or more effective feature selection, led to a more robust model that generalized better to the testing data. The reduced errors and higher R-Squared values indicate that the RF model in this work was able to recognize the patterns in the data more effectively than in the original study.

With regards to the XGBoost (XGB) model, the original testing RMSE of 63.86 and testing R² of 0.61 suggest strong performance. The results from this study also showed a high level of performance with a testing RMSE of 64.36, along with a testing R-Squared of 0.57. The slight increase in testing RMSE in these results compared to the original indicates that while the model was well-tuned during training, it may have experienced a slight overfitting, or the data split may have had slightly different characteristics. Nevertheless, the high R² values across both results demonstrate the effectiveness of XGBoost as a powerful predictive model in this context. The ET model, like XGBoost, demonstrated strong performance in both the original and this study. The original model had a testing RMSE of 59.61 and a testing R² of 0.66, while this study showed a comparable testing RMSE of 63.28 and a testing R² of 0.60. The close alignment between these results suggests that the Extra Trees model is highly consistent and robust across different implementations, likely due to its ability to minimize variance through ensemble learning. The slight differences could be attributed to variations in the number of trees or other hyperparameters, but overall, the ET model appears to be a reliable choice for this prediction task.

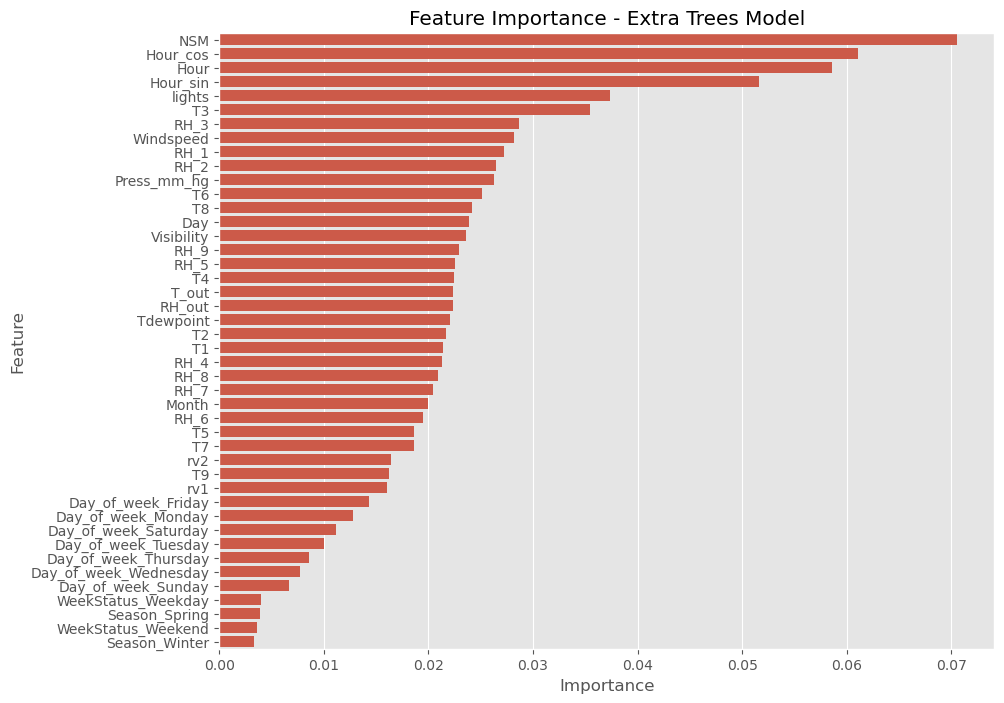
Finally, the Decision Tree (DT) and Neural Network (NN) models in this study did not perform as well as some of the other models. The DT model had a testing RMSE of 90.53 and a testing R² of 0.18, while the NN model showed a testing RMSE of 78.51 and a testing R² of 0.38. These results suggest that both models struggled to capture the difficult patterns of the data, likely due to overfitting in the case of DT (which is sensitive to tree depth) and insufficient tuning or architecture selection in the case of NN.

### Important Features of the Extra Trees model

The feature importance plot presented in Figure 3 reveals that temporal features play a dominant role in predicting appliance energy consumption in the Extra Trees model. The most important feature is NSM (Number of Seconds from Midnight), which suggests that the time of day is a critical factor in determining energy usage patterns. This is further supported by the high importance of the hour features which capture cyclical daily patterns, and the direct hour of the day. These features indicate that the model relies heavily on the time-specific characteristics of energy consumption.

In addition to temporal features, the energy usage of light fixtures (Lights) and room-specific conditions, such as temperature and humidity (e.g., T3, RH\_1, RH\_3), also play significant roles. These factors are likely linked to the operation of climate control systems and other appliances that respond to environmental conditions. The moderate importance of these features suggests that while the time of day drives the overall pattern of energy use, the specific conditions within the home also influence the intensity and timing of appliance usage. Environmental conditions, such as wind speed and atmospheric pressure, have a noticeable impact, indicating that external weather factors do play a role in energy consumption. Features like Season and Week Status are found to be the least important. This suggests that broader temporal aggregates are less informative than specific time-of-day data.

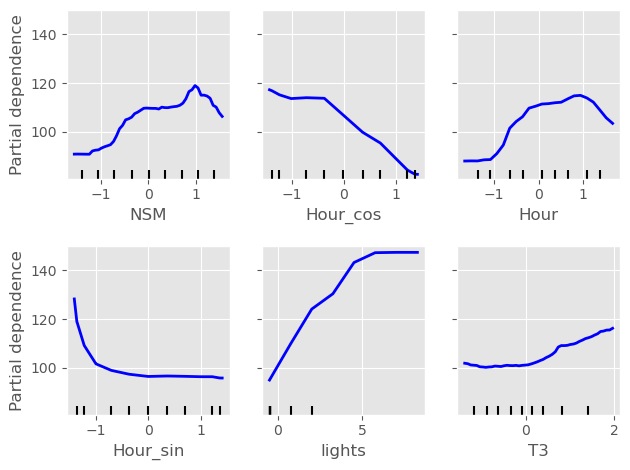
1. Feature Importance from the Extra Trees Model



### Partial Dependence

The partial dependence plots presented in Figure 4 provide a detailed look into how each of the most significant features influences appliance energy consumption while holding other features constant. This allows us to find the unique effect of any of the variables on appliance energy use. The plot for the Number of Seconds from Midnight (NSM) shows a generally upward trend, indicating that energy consumption tends to increase as the day progresses. This aligns with typical daily activities where energy use rises from morning through the day, likely peaking during evening hours when households are most active. The plot captures this daily trend, suggesting that time-of-day is a critical determinant of energy consumption patterns.

1. Partial dependence of top 6 features

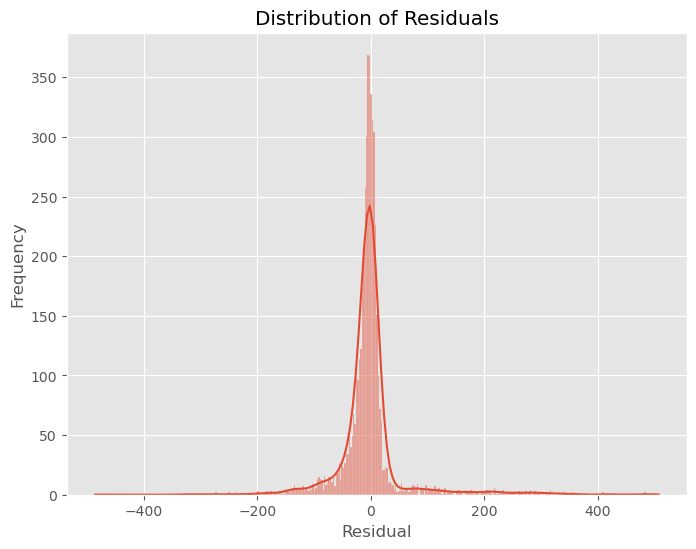


The Hour plot shows different peaks in energy consumption during certain hours, particularly in the morning and evening, which corresponds with common patterns of household energy use. The “Lights” feature, on the contrary, shows a strong direct relationship with overall energy consumption. As the use of lighting increases, so does the total energy consumption, which is expected given that lighting is a major contributor to household energy use, especially during the evening when lights are most frequently used. The temperature in the laundry room (T3) also shows a slightly upward trend, indicating that as the temperature rises, so does energy usage. These plots provide a comprehensive view of how time-related features and specific appliance usage drive energy consumption, offering valuable insights for managing energy use more efficiently.

### Residual Distribution

The residual plot in Figure 5 gives a visual representation of how frequently different levels of error occur across all predictions. A well-calibrated model will produce a residual distribution that is centred around zero, with most residuals clustered close to zero, indicating that the model's predictions are generally accurate. In the histogram provided, the residuals are indeed centred around zero, forming a sharp peak, which suggests that most of the predicted values of the algorithm are very close to the real values. This sharp peak is a positive indicator, showing that the model has low bias. However, the presence of long tails on either side of the distribution indicates that

1. Residuals Distribution

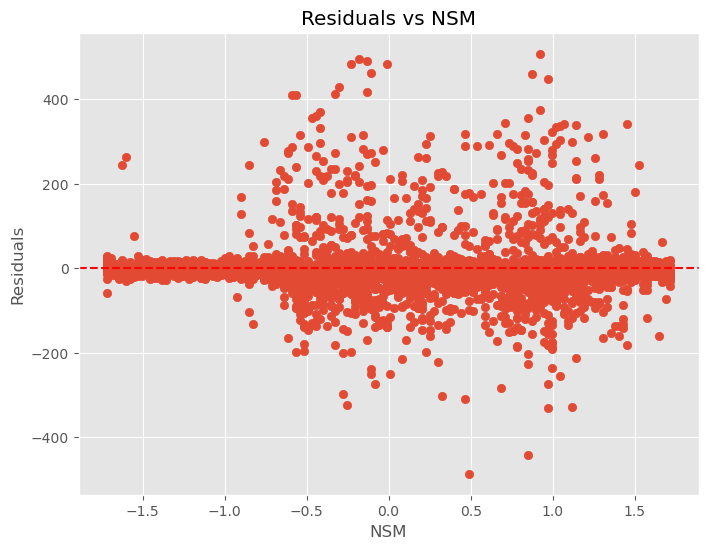


there are some predictions with larger errors. These tails are areas of potential concern as they suggest that while the model performs well for most predictions, there are specific instances where it does not perform as accurately. Investigating these instances could uncover specific conditions or data characteristics that challenge the model, offering opportunities for further refinement.

### Residual Plots

While Figure 5 shows the overall distribution of the residuals, the residual plots in Figure 6 display the distribution of the residuals for given values of NSM. In this context, the residual plot against the “NSM” feature (Number of Seconds from Midnight) is particularly important because NSM is the dominant contributing factor to appliance early prediction as indicated earlier. Ideally, residuals should be randomly scattered around the horizontal axis (at zero), indicating that the model has no systematic errors and that its predictions are unbiased across all ranges of the feature. This is the case in the residual plot in Figure 6. The points are spread out horizontally across different values of “NSM”, with no clear pattern or trend. This randomness suggests that the model does not deliberately overpredict or underpredict energy consumption based on the time of day. However, there are some larger residuals (both positive and negative), which indicate instances where the model's predictions are off by a significant margin. These outliers could be due to specific conditions or anomalies in the data that the model fails to capture.

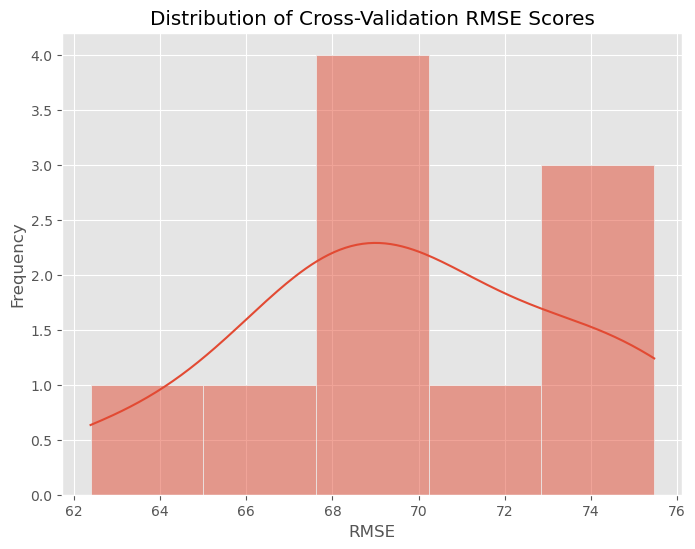
1. How the Residuals change with the NSM



### Cross-Validation RMSE Distribution

The cross-validation RMSE distribution in Figure 7 is vital for evaluating the stability and reliability of the estimated model across varying subsamples of the data. Cross-validation is a robust method to measure how efficiently the model generalizes to independent data sets. The approach splits the data into different folds and evaluates how the model performs on each fold. The histogram of cross-validation RMSE scores shows how these errors are distributed across different folds of the data. In the graph, the RMSE values appear to be concentrated around a central peak with values of 68 - 70, indicating that the efficiency of the algorithm is consistent across several other splits of the data. However, the spread in the RMSE values—reflected in the width of the histogram—indicates variability in performance, which could be due to the model's sensitivity to specific data characteristics.

1. Distribution of Cross Validation Scores

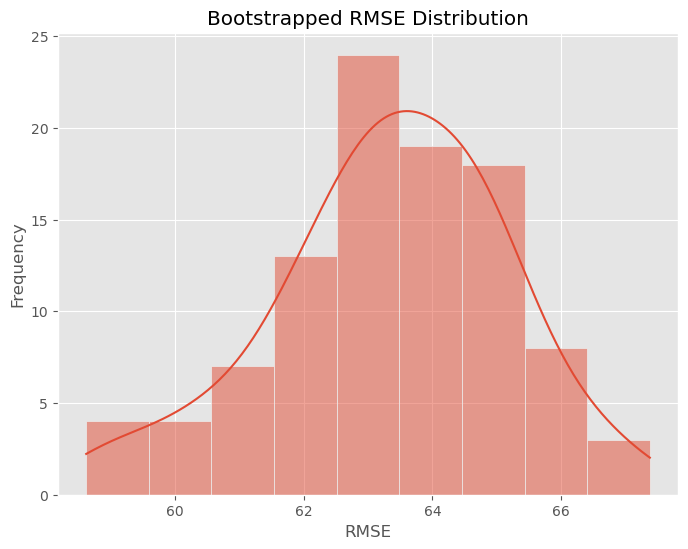


### Bootstrapped RMSE Distribution

A similar visualization to the cross validation RMSE errors in Figure 7 is the Bootstrapped RMSE distribution in Figure 8. Bootstrapping is a technique where multiple training sets are generated by repeatedly sampling with replacement, the dataset. The model is then trained and assessed using these varied sets. This process generates a distribution of RMSE estimates, allowing us to estimate how the performance of the model might vary under different data conditions.

In the bootstrapped RMSE distribution, the histogram shows the spread of RMSE values across these bootstrapped samples. The horizontal axis represents the RMSE estimates, ranging from approximately 58 to 68, while the vertical axis represents how many times each range of RMSE values occurred across the bootstrap samples. The histogram shows that the majority of RMSE values are clustered around the centre, particularly between 60 and 64. This clustering indicates that the model's prediction errors are relatively consistent across different bootstrap samples, with most errors falling within this central range. The peak of the histogram, around an RMSE of 62, suggests that this is the most common error rate across the different samples, reflecting the model's typical performance. Additionally, the distribution appears to be roughly normal but with a slight skew to the right, meaning that while most bootstrapped RMSE values are concentrated around the central value, there are some instances where the model yields higher errors, resulting in a longer tail on the right side of the distribution. The Kernel Density Estimate (KDE) curve, drawn on the histogram provides a smooth approximation of this distribution, further highlighting the central tendency and the spread of errors.

1. Bootstrapped RMSE distribution



# Conclusion

## Summary

In this project, I successfully applied a series of machine learning models to estimate how much of appliance energy will a household consume given various factors, including temporal data and environmental conditions. The analysis involved an exploratory data analysis (EDA) to get an overview of how the variables are related and identify key patterns. The eight machine learning models were subsequently trained and assessed using performance metrics like RMSE, MAE, MAPE, and R². The models trained and tested include Multiple Linear Regression (LM), Support Vector Regression (SVR), Random Forest (RF), Decision Tree (DT), Gradient Boosting (GBM), XGBoost (XGB), Extra Trees (ET), and Neural Networks (NN).

## Evaluation

When evaluating the objectives outlined in section 1.2, the project achieved its general objective of identifying the most effective model for predicting appliance energy consumption. The study introduced Decision Tree and Neural Network models aimed to see if these models could surpass the performance of the previously used models like Multiple Linear Regression (LM), SVR, RF, and GB. However, the results indicate that DT and NN did not outperform these models. Specifically, the DT model had a testing RMSE of 90.53 and a testing R² of 0.18, while the NN model had a testing RMSE of 78.51 and a testing R² of 0.38. These performance metrics are lower compared to models like RF, XGBoost, and Extra Trees, which exhibited superior performance in terms of both RMSE and R². The Extra Trees model showed superiority in performance with RMSE and R2 values of 63.28 and 0.6, respectively. The study therefore crowned the Extra Trees (ET) model as the best model for predicting appliance energy usage, just as [2] concluded.

The study also intended to assess the impact of additional feature engineering on the dataset. The overall results show mixed outcomes. For example, while models like RF and XGBoost showed improvements in testing RMSE and R², the performance of GB and SVM models deteriorated compared to the original studies. The DT and NN models, which also incorporated the additional features, did not perform as well as anticipated. This suggests that the new feature engineering techniques may not have universally enhanced model performance across all algorithms. Instead, their impact was variable, benefiting some models more than others.

## Future Work

While this study successfully identified key factors influencing household energy consumption and developed predictive models, there are areas for further exploration. The results showed that while models like Extra Trees, Random Forest and XGBoost were effective, others like Support Vector Regression and Neural Networks did not perform as well. Future work could focus on refining these models, possibly by experimenting with different kernels, architectures, or feature scaling techniques. Additionally, further research could explore the inclusion of more complex features or the use of ensemble methods to improve predictive accuracy.

Furthermore, while this study focused on temporal and environmental variables, future studies could incorporate socio-demographic factors or real-time external data such as energy prices to enhance the models' applicability in real-world energy management systems. Finally, cross-validation with different datasets and further comparison with other existing models could help to generalize the findings and improve model robustness.

References

[1] L. Candanedo, V. Feldheim, and D. Deramaix, “Data driven prediction models of energy use of appliances in a low-energy house,” Energy Build, vol. 140, pp. 81–97, 2017, doi: 10.1016/J.ENBUILD.2017.01.083.

[2] C. F. Assadian and F. Assadian, “Data-Driven Modeling of Appliance Energy Usage,” Energies (Basel), p., 2023, doi: 10.3390/en16227536.

[3] Y. Xie and A. I. M. Noor, “Factors Affecting Residential End-Use Energy: Multiple Regression Analysis Based on Buildings, Households, Lifestyles, and Equipment,” Buildings, p., 2022, doi: 10.3390/buildings12050538.

[4] P. Rickwood, “Residential operational energy use,” Urban Policy and Research, vol. 27, no. 2, 2009, doi: 10.1080/08111140902950495.

[5] A. Kavousian, R. Rajagopal, and M. Fischer, “Determinants of residential electricity consumption: Using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants’ behavior,” Energy, vol. 55, pp. 184–194, 2013, doi: 10.1016/J.ENERGY.2013.03.086.

[6] E. Leahy and S. Lyons, “Energy use and appliance ownership in Ireland,” Energy Policy, vol. 38, pp. 4265–4279, 2010, doi: 10.1016/J.ENPOL.2010.03.056.

[7] K. S. Cetin, P. Tabares-Velasco, and A. Novoselac, “Appliance daily energy use in new residential buildings: Use profiles and variation in time-of-use,” Energy Build, vol. 84, pp. 716–726, 2014, doi: 10.1016/J.ENBUILD.2014.07.045.

[8] A. Kavousian, R. Rajagopal, and M. Fischer, “Ranking appliance energy efficiency in households: Utilizing smart meter data and energy efficiency frontiers to estimate and identify the determinants of appliance energy efficiency in residential buildings,” Energy Build, vol. 99, 2015, doi: 10.1016/j.enbuild.2015.03.052.

[9] J. Rouleau, L. Gosselin, and P. Blanchet, “Understanding energy consumption in high-performance social housing buildings: a case study from Canada,” Energy, vol. 145, pp. 677–690, 2018, doi: 10.1016/J.ENERGY.2017.12.107.

[10] A. Iwayemi, W. Wan, and C. Zhou, “Energy management for intelligent buildings,” Energy Manag. Syst, 2011.

[11] J. Reyna and M. Chester, “Energy efficiency to reduce residential electricity and natural gas use under climate change,” Nat Commun, vol. 8, p., 2017, doi: 10.1038/ncomms14916.

[12] S. Ullah, M. Nazeer, and N. Malik, “Machine Learning based Energy Consumption Prediction of Appliances in a Low Energy House,” International journal of Engineering Works, p., 2020, doi: 10.34259/IJEW.20.710326332.

[13] G. R. Duarte, L. G. da Fonseca, P. V. Z. C. Goliatt, and A. C. de C. Lemonge, “Comparison of machine learning techniques for predicting energy loads in buildings,” Ambiente Construído, vol. 17, no. 3, 2017.

[14] M. Rambabu, N. Ramakrishna, and P. K. Polamarasetty, “Prediction and Analysis of Household Energy Consumption by Machine Learning Algorithms in Energy Management,” E3S Web of Conferences, p., 2022, doi: 10.1051/e3sconf/202235002002.

[15] R. E. Edwards, J. New, and L. E. Parker, “Predicting future hourly residential electrical consumption: A machine learning case study,” Energy Build, vol. 49, pp. 591–603, 2012, doi: 10.1016/J.ENBUILD.2012.03.010.

[16] R. K. Jain, K. Smith, P. Culligan, and J. Taylor, “Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy,” Appl Energy, vol. 123, pp. 168–178, 2014, doi: 10.1016/J.APENERGY.2014.02.057.

[17] L. Xiang, T. Xie, and W. Xie, “Prediction model of household appliance energy consumption based on machine learning,” J Phys Conf Ser, vol. 1453, p., 2020, doi: 10.1088/1742-6596/1453/1/012064.

[18] E. Mocanu, H. Nguyen, M. Gibescu, and W. Kling, “Deep learning for estimating building energy consumption,” Sustainable Energy, Grids and Networks, vol. 6, pp. 91–99, 2016, doi: 10.1016/J.SEGAN.2016.02.005.

[19] J. Jiang, Q. Kong, M. D. Plumbley, and N. Gilbert, “Deep Learning Based Energy Disaggregation and On/Off Detection of Household Appliances,” ArXiv, vol. abs/1908.00941, p., 2019, doi: 10.1145/3441300.

[20] S. Bourhnane, M. Abid, R. Lghoul, K. Zine-dine, N. Elkamoun, and D. Benhaddou, “Machine learning for energy consumption prediction and scheduling in smart buildings,” SN Appl Sci, vol. 2, pp. 1–10, 2020, doi: 10.1007/s42452-020-2024-9.

[21] N.-T. Ngo, A. Pham, T. T. H. Truong, N.-S. Truong, N.-T. Huynh, and T. M. Pham, “An Ensemble Machine Learning Model for Enhancing the Prediction Accuracy of Energy Consumption in Buildings,” Arab J Sci Eng, vol. 47, pp. 4105–4117, 2021, doi: 10.1007/s13369-021-05927-7.

[22] M. Al-Rakhami, A. Gumaei, A. Alsanad, A. Alamri, and M. M. Hassan, “An Ensemble Learning Approach for Accurate Energy Load Prediction in Residential Buildings,” IEEE Access, vol. 7, 2019, doi: 10.1109/ACCESS.2019.2909470.

[23] L. N. M, K. C. K. S, A. Mohan, and V. Gopal, “Appliance Prediction from Total Energy Data — A Demand Response Method Using Simple and Complex Networks,” 2019 IEEE 2nd International Conference on Power and Energy Applications (ICPEA), pp. 222–226, 2019, doi: 10.1109/ICPEA.2019.8818489.

[24] M. Sajjad et al., “A Novel CNN-GRU-Based Hybrid Approach for Short-Term Residential Load Forecasting,” IEEE Access, vol. 8, pp. 143759–143768, 2020, doi: 10.1109/access.2020.3009537.

[25] D. M. Ibrahim, A. Almhafdy, A. A. Al-Shargabi, M. Alghieth, A. Elragi, and F. Chiclana, “The use of statistical and machine learning tools to accurately quantify the energy performance of residential buildings,” PeerJ Comput Sci, vol. 8, p., 2022, doi: 10.7717/peerj-cs.856.

[26] A. Moradzadeh, B. Mohammadi-ivatloo, M. Abapour, A. Anvari‐Moghaddam, and S. S. Roy, “Heating and Cooling Loads Forecasting for Residential Buildings Based on Hybrid Machine Learning Applications: A Comprehensive Review and Comparative Analysis,” IEEE Access, vol. PP, p. 1, 2021, doi: 10.1109/access.2021.3136091.

[27] Z. Wu and H. He, “Traditional Machine Learning Models for Building Energy Performance Prediction: A Comparative Research,” Machine Learning Research, vol. 8, no. 1, pp. 1–8, 2023.

[28] N. S. I. M. Husin, S. Mostafa, M. M. Jaber, S. Gunasekaran, A. Al-shakarchi, and N. F. Abdulsattar, “Machine Learning Regression Approach for Estimating Energy Consumption of Appliances in Smart Home,” 2023 Al-Sadiq International Conference on Communication and Information Technology (AICCIT), pp. 229–233, 2023, doi: 10.1109/AICCIT57614.2023.10217991.

[29] D.-H. Tran, D.-L. Luong, and J.-S. Chou, “Nature-inspired metaheuristic ensemble model for forecasting energy consumption in residential buildings,” Energy, vol. 191, p. 116552, 2020, doi: 10.1016/j.energy.2019.116552.

[30] E. U. Haq, X. Lyu, Y. Jia, M. Hua, and F. Ahmad, “Forecasting household electric appliances consumption and peak demand based on hybrid machine learning approach,” Energy Reports, vol. 6, pp. 1099–1105, 2020, doi: 10.1016/j.egyr.2020.11.071.

[31] H. G. Zaini, “Forecasting of Appliances House in a Low-Energy Depend on Grey Wolf Optimizer,” Computers, Materials & Continua, p., 2022, doi: 10.32604/cmc.2022.021998.

[32] A. Al-Adaileh and S. Khaddaj, “Machine Learning Prediction Based Integrated Smart Energy Management System to Improve Home Energy Efficiency,” 2022 21st International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES), pp. 1–4, 2022, doi: 10.1109/DCABES57229.2022.00042.

[33] S. Iram et al., “An Innovative Machine Learning Technique for the Prediction of Weather Based Smart Home Energy Consumption,” IEEE Access, vol. 11, pp. 76300–76320, 2023, doi: 10.1109/ACCESS.2023.3287145.

[34] and C.-J. L. Chih-Wei Hsu, Chih-Chung Chang, “A Practical Guide to Support Vector Classification,” BJU Int, vol. 101, no. 1, 2008.

[35] S. Seyedzadeh, F. Rahimian, P. Rastogi, and I. Glesk, “Tuning machine learning models for prediction of building energy loads,” Sustain Cities Soc, p., 2019, doi: 10.1016/J.SCS.2019.101484.

[36] J. Bergstra and Y. Bengio, “Random search for hyper-parameter optimization,” Journal of Machine Learning Research, vol. 13, 2012.

[37] M. Feurer, A. Klein, K. Eggensperger, J. T. Springenberg, M. Blum, and F. Hutter, “Efficient and robust automated machine learning,” in Advances in Neural Information Processing Systems, 2015.

[38] T. A. Ryttov and R. Shrock, “Infrared fixed-point physics in so (Nc) and Sp (Nc) gauge theories,” Physical Review D, vol. 96, no. 10, 2017, doi: 10.1103/PhysRevD.96.105015.

[39] S. R. Young, D. C. Rose, T. P. Karnowski, S. H. Lim, and R. M. Patton, “Optimizing deep learning hyper-parameters through an evolutionary algorithm,” in Proceedings of MLHPC 2015: Machine Learning in High-Performance Computing Environments - Held in conjunction with SC 2015: The International Conference for High Performance Computing, Networking, Storage and Analysis, 2015. doi: 10.1145/2834892.2834896.

[40] D. Moldovan and A. Słowik, “Energy consumption prediction of appliances using machine learning and multi-objective binary grey wolf optimization for feature selection,” Appl. Soft Comput., vol. 111, p. 107745, 2021, doi: 10.1016/J.ASOC.2021.107745.

[41] Y.-W. Chen and C.-J. Lin, “Combining SVMs with Various Feature Selection Strategies,” in Feature Extraction: Foundations and Applications, I. Guyon, M. Nikravesh, S. Gunn, and L. A. Zadeh, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 315–324. doi: 10.1007/978-3-540-35488-8\_13.

[42] I. T. Jolliffe, Principal component analysis for special types of data. Springer, 2002.

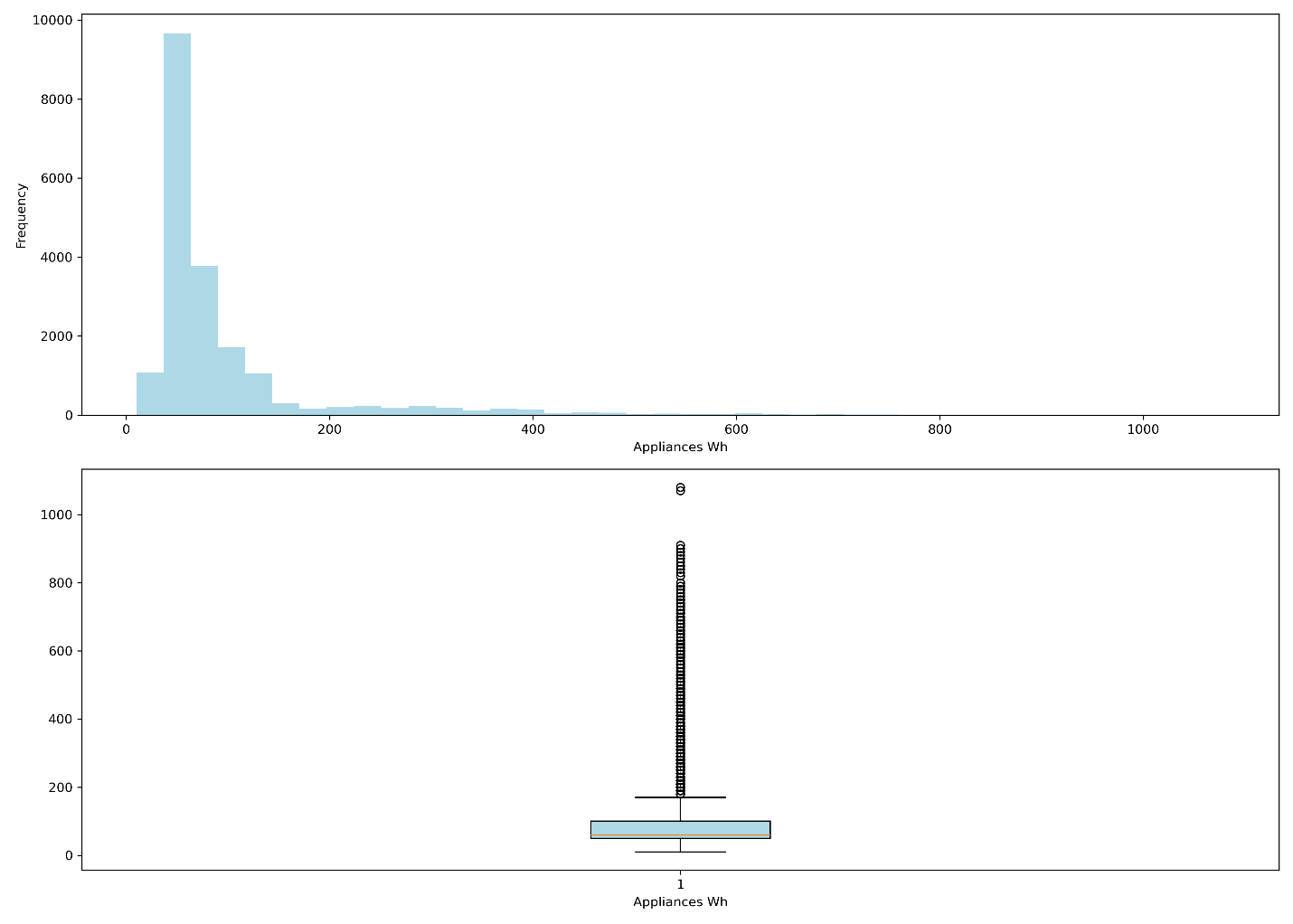
[43] H. Abdi and L. J. Williams, “Principal component analysis,” 2010. doi: 10.1002/wics.101.

[44] L. Candanedo, “Appliances Energy Prediction,” 2017. doi: https://doi.org/10.24432/C5VC8G.

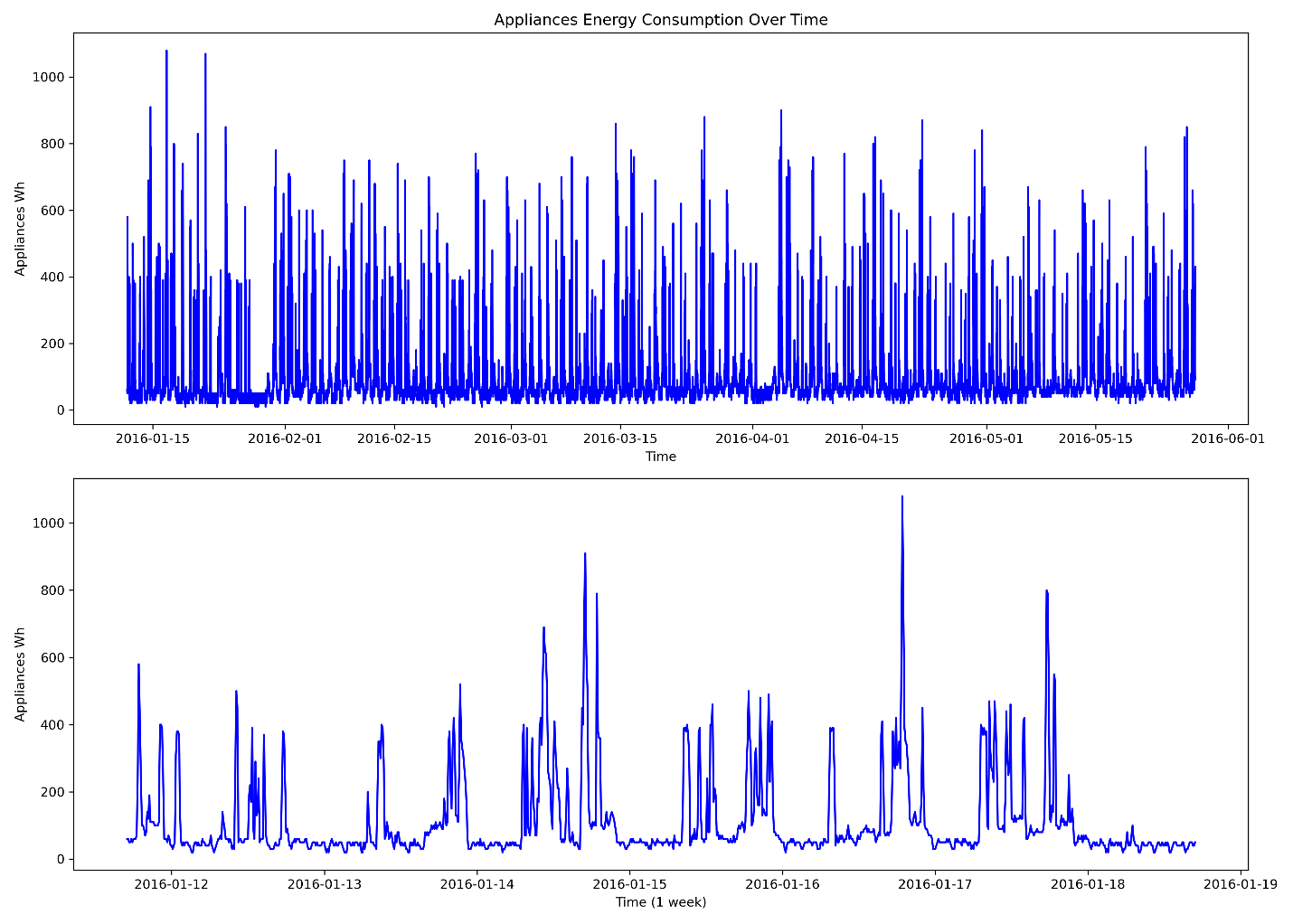
[45] S. Yadav and S. Shukla, “Analysis of k-Fold Cross-Validation over Hold-Out Validation on Colossal Datasets for Quality Classification,” in Proceedings - 6th International Advanced Computing Conference, IACC 2016, 2016. doi: 10.1109/IACC.2016.25.

Appendix 1

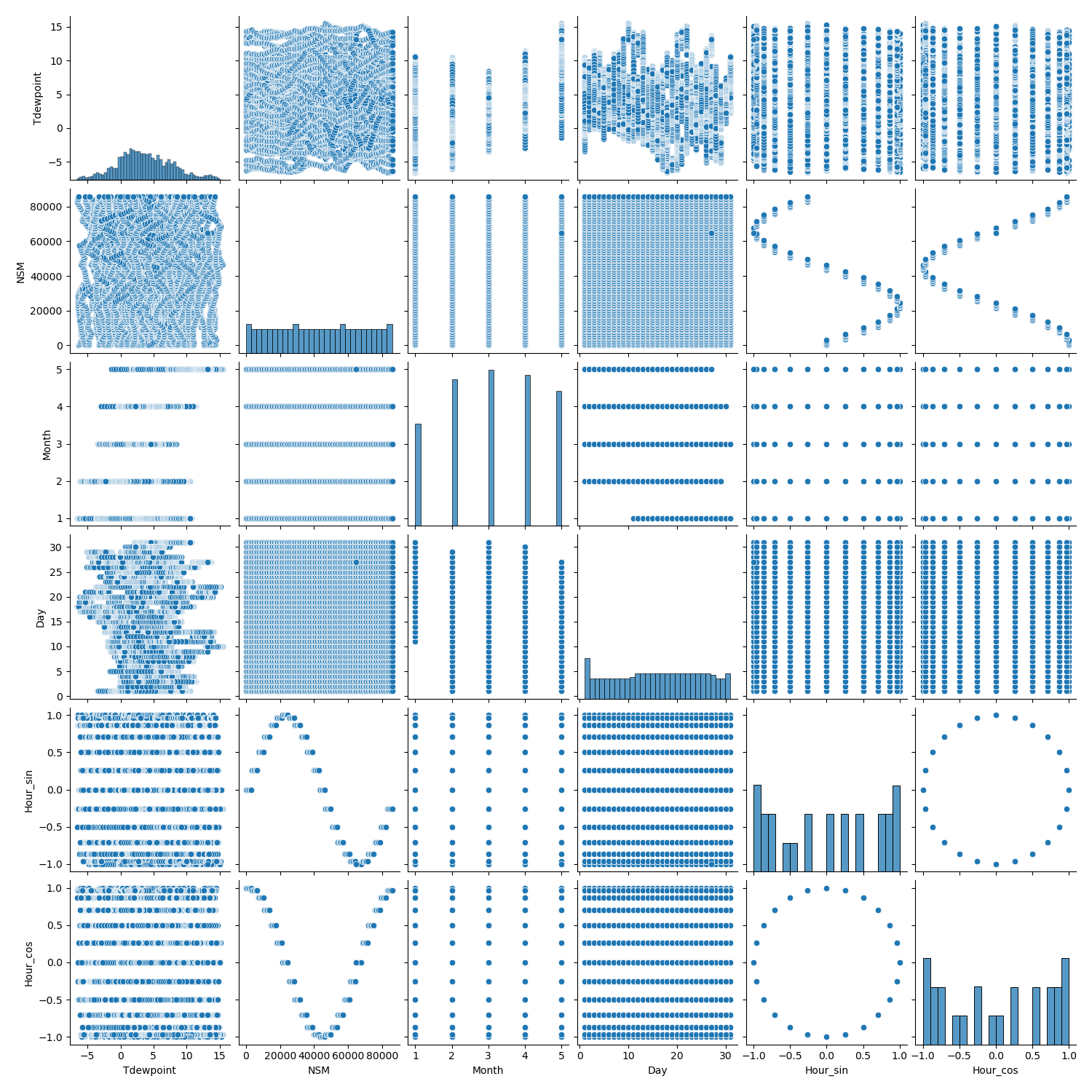
1. Distribution of Appliance Energy Consumption



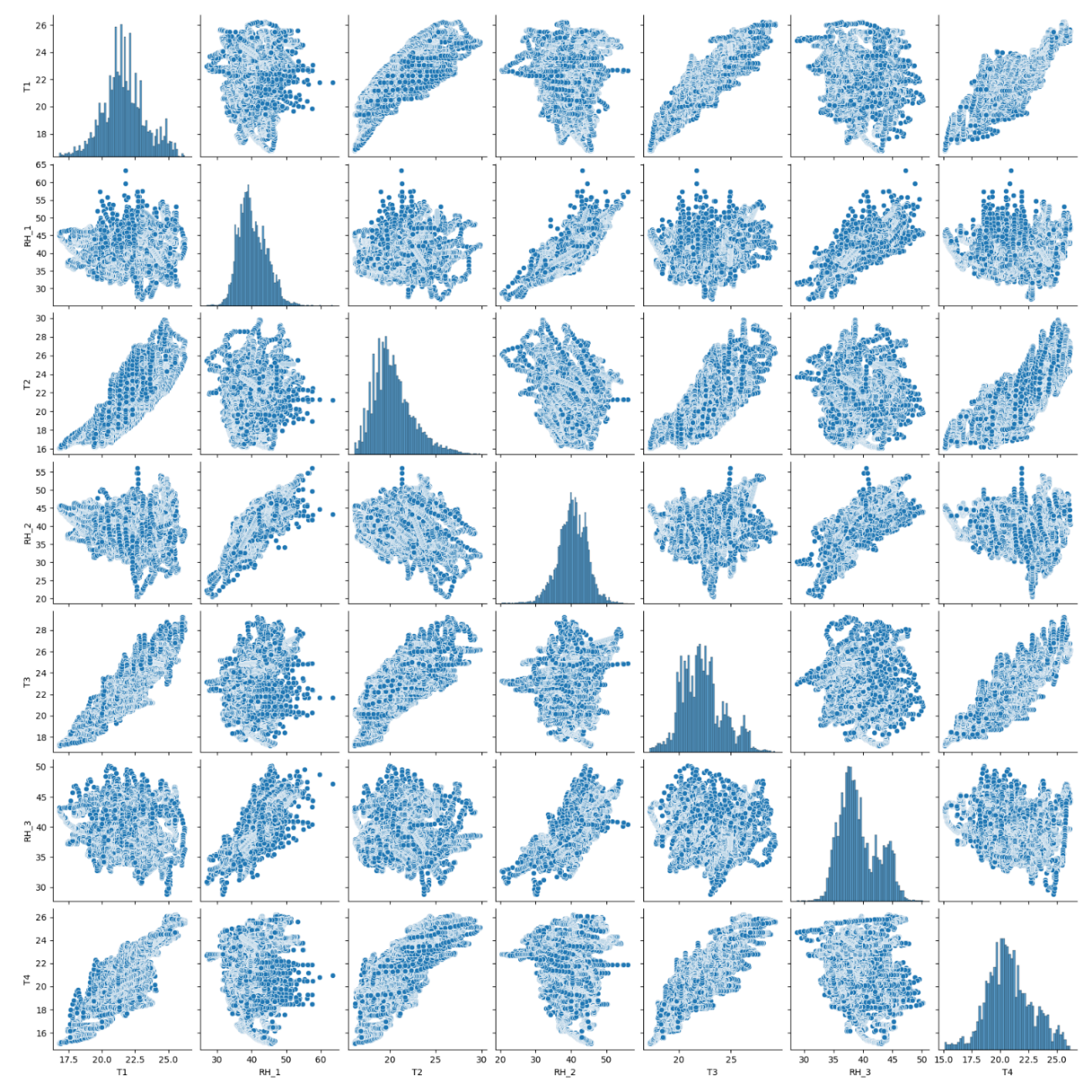
1. Appliance Energy Consumption: All time and Weekly plot



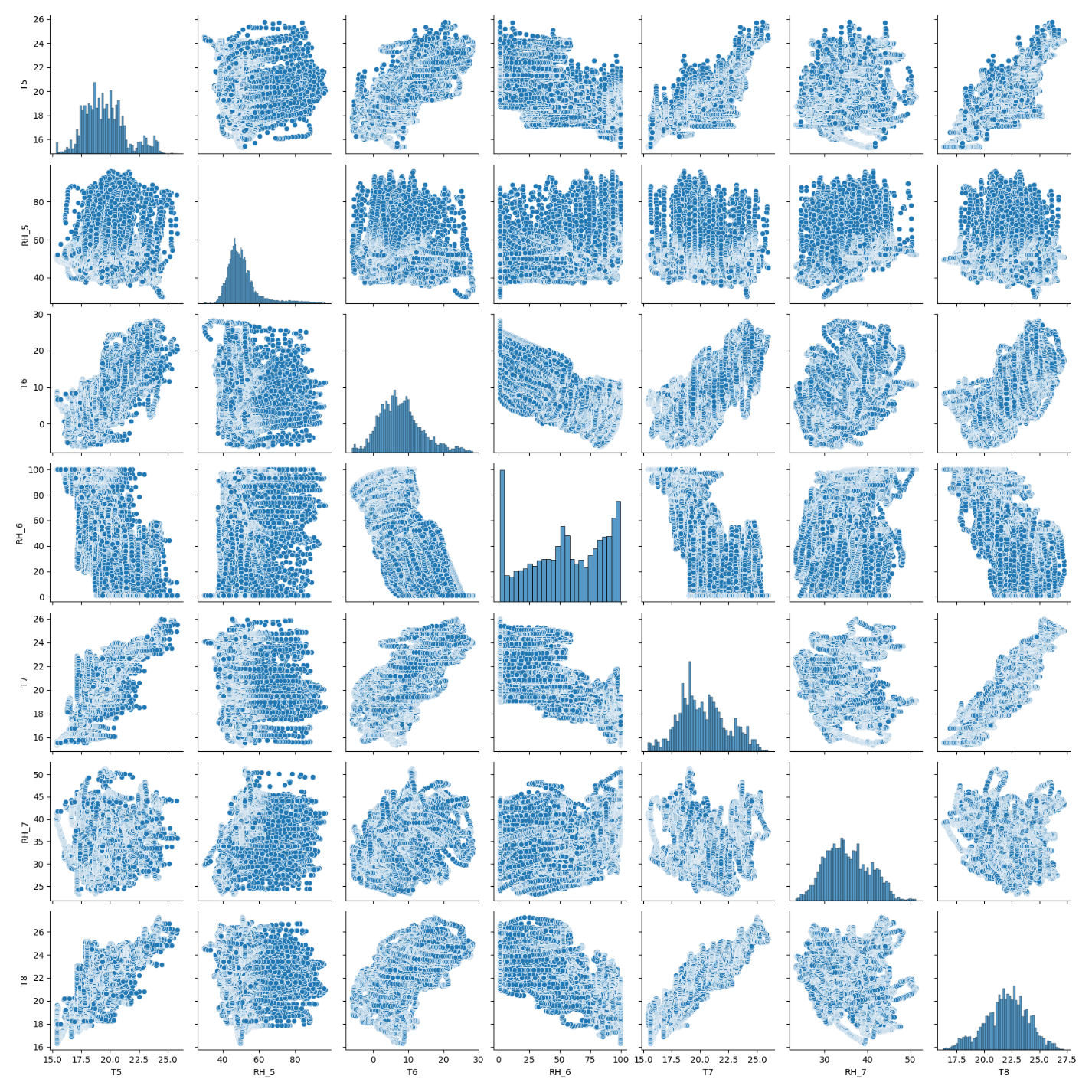
1. Pair Plot Set 1



1. Pair Plot Set 2



1. Pair Plot Set 3



1. Pair Plot Set 4

