Predicting Appliances' Energy Consumption

Ebad Akhter

Shiley-Marcos School of Engineering, University of San Diego

ADS 506: Applied Time Series Analysis

Professors Erin Cooke and Dave Hurst

November 30, 2023

Abstract:

The purpose of this paper is to present and discuss the data-driven predictive models for energy consumption by various appliances in a household. The data used throughout this project includes measurements of temperature and humidity sensors from a wireless network, weather reports from a nearby city-airport station, and recorded energy use of lighting fixtures within the household. The paper includes data filtering, to remove non-predictive features and feature ranking techniques. Four statistical models were coached and valued on a validation dataset. Multiple Linear Regression (MLR), Auto-Regressive Integrated Moving Average (ARIMA) Model, Naïve Model, and various Exponential Smoothing (ETS) Algorithms. The best model, MLR, was able to explain the dataset with the highest percentage of Bins within Tolerance, being 72.41%. The second best on the list was ARIMA with External Predictors, with Bin Classification Accuracy Percentage of 64.28%. Using the wireless network, data from Kitchen, Laundry, and Living room were considered as of highest importance, as these are the main rooms in a house, utilizing the most electricity. The prediction models used energy consumption data along with the weather data as the most relevant data variables in the prediction process. Atmospheric pressure may be important to include in energy prediction models and for building performance models for future use.

Table of Contents

1. Project Background

1.1. Introduction:

Research on understanding energy consumption in buildings has been extensively explored, with a particular emphasis on appliances, which constitute a substantial portion (ranging from 20% to 30%) of the overall electrical energy demand (refer to Figure 1) (Barbato et al., 2011). Notably, a UK-based study on domestic buildings attributed a 10.2% increase in electricity consumption to appliances like televisions and consumer electronics operating in standby mode (Firth et al., 2008). To solve the complexities of energy use, regression models play a pivotal role, aiding in comprehending the relationships between various variables and quantifying their impacts.

Prediction models for electrical energy consumption in buildings offer multi-layered applications, such as determining the ideal sizing of photovoltaics (SETO, 2023) and energy storage to reduce grid power flow, identifying abnormal energy use patterns, integrating into energy management systems for load control (Barbato et al., 2011), supporting model predictive control applications, facilitating demand side management (DSM) and demand side response (DSR) initiatives (Kristen S., 2016), and serving as input for building performance simulation analyses.

According to insights from (Firth et al., 2008), electricity consumption in domestic buildings is influenced by two primary factors: the type and quantity of electrical appliances and their usage by occupants. Naturally intertwined, these factors leave discernible signals in the indoor environment, including temperature, humidity, vibrations, light, and noise. The occupancy levels in different locations within the building further contribute to understanding appliance usage. This study employs diverse data sources and environmental parameters, encompassing data from a nearby airport weather station, temperature, and humidity readings from various rooms via a wireless sensor network, and the electrical energy consumption of specific appliances (lights) through sub-metering.

Four regression models undergo testing in this study: Multiple Linear Regression (MLR), Auto-Regressive Integrated Moving Average (ARIMA) Model, Naïve Model, and various Exponential Smoothing (ETS) Algorithms. each configured with different combinations of predictors. While the primary focus of this work is on predicting aggregate appliance energy usage, rather than delving into the specifics of modeling individual appliance energy loads, the literature review encompasses both aspects. The review not only addresses studies related to aggregate energy predictions but also includes insights from research exploring the modeling intricacies of individual appliance loads within buildings.

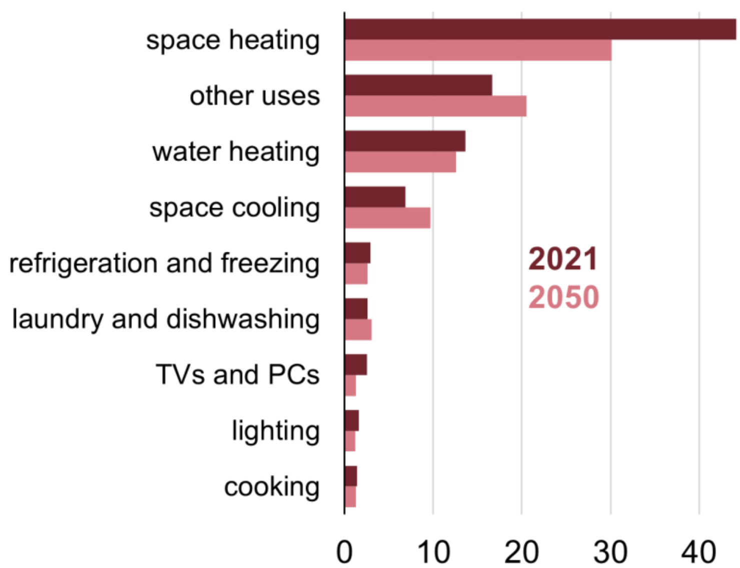

Figure 1: Residential energy intensity by end use (AEO2022 Reference Case)

1.2. Literature Review:

1.2.1. Appliances' Loads in Buildings and Numerical Modeling of Their Consumption:

This section provides an extensive overview of literature addressing the modeling of appliances and socio-economic factors, offering insights into diverse data sources and methodologies employed in comprehending appliances' energy consumption.

Hourly end-use data, collected from 454 houses and 140 commercial buildings in the Pacific Northwest, is a notable contribution (Pratt et al., 1993). The data acquisition systems monitored 12 to 16 channels for energy consumption. The study revealed significant temporal variations in end-use loads, with distinct evening peaks observed in cooking, dishwasher, lights, and small appliances, while refrigeration and freezer load exhibited relative stability.

A noteworthy method for generating occupancy data for UK households using the Markov Chain Monte Carlo technique is presented in (Richardson et al., 2008). This model is recommended for estimating the energy demand of appliances, lighting, and heating. Additionally, a methodology for estimating building energy consumption from Energy Plus benchmark models is discussed in (Fumo et al., 2010), proposing the use of predetermined coefficients to estimate hourly energy consumption from utility bills, thus alleviating the need for dynamic simulation.

A recent study focusing on major household appliances, including refrigerators, clothes washers, clothes dryers, and dishwashers, delineates daily energy use profiles, emphasizing the user-dependency of certain appliances. Furthermore, research on demand response potential ranks clothes dryers highest due to their substantial power demand.

The development of a model to detect and estimate individual home appliance loads using explicit duration Hidden Markov models is a notable contribution. A literature review scrutinizes socioeconomic factors, dwelling characteristics, and appliances influencing electricity consumption in domestic buildings (Jones et al., 2015). Appliances such as desktops, laptops, TVs, video players/recorders, video consoles, electric ovens, range hoods, refrigerators, freezers, dishwashers, washing machines, and tumble dryers are highlighted as having a significant positive effect on domestic electricity consumption.

While these models are primarily employed in energy building simulation studies to evaluate diverse building designs and predict their impact on energy balances, the subsequent section will delve into the challenge of predicting energy use during the operational phase.

1.2.2. Electricity Load Prediction

Traditionally, various models such as multiple regression, neural networks, engineering methods, support vector machines, and time series techniques have been employed to predict electricity demand. These models consider diverse parameters, including time of day, outdoor temperature, month, weekend, holidays, yesterday’s consumption, rainfall index, global solar radiation, wind speed, and occupancy.

A comprehensive study investigated the impact of weather variables on monthly electricity demand in England and Wales from 1983 to 2003 (Hor et al., 2005). Utilizing a parameter multiple regression model, the study incorporated heating degree days, cooling degree days, gross domestic product, and humidity, explaining the monthly demand variability between 91% and 95%.

Exploring electrical energy use patterns in buildings, a study presented one-minute interval power measurements from 12 houses, unveiling significant daily variations in temporal distribution for individual devices. Another study analyzed electricity data from 1628 households, identifying crucial variables such as weather, location, floor area, number of refrigerators, and entertainment devices. Similarly, research found that being at home during the day correlated with lower appliance efficiency, attributed to increased appliance usage when the house is occupied more frequently.

A predictive system for individual appliance consumption was proposed, leveraging past consumption, hour, day, season, and month data [8]. Notably, the last 24 hours were deemed the most relevant for accurate prediction.

Recent investigations monitored 23 houses in Ottawa, Canada, at one-minute intervals, highlighting the number of occupants as the strongest predictor for non-HVAC energy use [17]. In the UK, a study examined appliance ownership and use for 183 dwellings, revealing a correlation between owning more than thirty appliances and high electrical energy demand [5]. However, this study's limitation was reliance on survey data rather than direct measurements.

The literature review underscores key points:

- Appliances contribute significantly to residential sector electricity demand (up to 30%) and are crucial for grid power management [9].
- Identifying primary contributors to energy consumption becomes vital with an increasing number of owned appliances.
- Energy use patterns vary widely among appliances, with weather parameters proving relevant for predicting electricity consumption.
- In highly insulated buildings, the thermal impact of appliances gains importance in building energy performance [2].

This paper addresses several questions, including the representativeness of weather data from a nearby station for improving appliance energy consumption prediction, the utility of temperature and humidity measurements from a wireless network in energy prediction, the identification of crucial parameters in energy prediction models, and the impact of including sub-metered energy measurements related to occupancy (light) on prediction accuracy.

1.2.3. Predictive Techniques for Large-Scale Building Energy Applications

Building energy prediction techniques are pivotal tools in the pursuit of sustainable built environments, playing crucial roles in shaping energy policies and guiding the development of the building sector. This literature review offers a comprehensive examination of prevalent prediction techniques employed in large-scale building energy applications, considering various scopes and archetypes. The methods covered encompass black-box, white-box, and grey-box approaches.

The review delves into the advantages and disadvantages of these prediction techniques, providing a nuanced understanding of their applications within the context of large-scale buildings. The analysis reveals that prediction techniques have been extensively applied in addressing diverse aspects of large-scale building energy, including energy consumption forecasting and prediction, energy consumption profiling, as well as energy mapping and benchmarking of buildings.

Despite the progress made, the review identifies research gaps that warrant further attention. Notably, the need to incorporate occupant behavior in white-box models and the explicit representation of end-uses in black-box models emerge as key areas requiring exploration. The review underscores the importance of addressing these gaps to enhance the accuracy and applicability of prediction techniques in large-scale building energy applications.

This paper outlined essential tasks for modifying the current prediction approach framework. These modifications are envisioned to contribute to forecasting future energy use changes during the retrofit process, incorporating renewable energy technology, and ultimately aiding in the development of sustainable strategies for the built environment.

2. Data Preparation and Exploratory Data Analysis:

2.1. Data Ingestion and Pre-Processing:

The dataset used in this project originates from a residence in Stambruges, approximately 24 km from the City of Mons in Belgium. Constructed in December 2015, the building incorporates new mechanical systems, aligning with passive house certification standards (Feist, W., 2007). The design targets an annual heating and cooling load below 15 kWh/m2 per year, following the Passive House Planning Package (PHPP) guidelines. Notably, a wood chimney serves as the primary heat source, with detailed monthly records of wood type and quantity manually documented.

Covering around 4.5 months, the dataset records data at 10-minute intervals. House temperature and humidity conditions were monitored through a ZigBee wireless sensor network, with each node transmitting information every 3.3 minutes. The wireless data was then aggregated into 10-minute intervals. Energy consumption data, recorded every 10 minutes for appliances, was collected using m-bus energy meters.

To enhance the dataset's comprehensiveness, weather data from Chievres Airport in Belgium were integrated, aligning entries based on date and time columns (Candanedo, L., 2017). This analysis specifically focuses on energy consumption data recorded at 10-minute intervals for household appliances and includes another sub-metered load—lights—known for its predictive value in room occupancy when combined with relative humidity measurements.

Temperature and humidity readings from the wireless sensor network were averaged for corresponding 10-minute periods and integrated into the energy dataset based on date and time. All data analysis was conducted using the R programming language. The dataset spans 137 days (4.5 months), as illustrated in Figure 2, showcasing the energy consumption profile for the specified duration, revealing considerable variability.

2.2. Data Transformation:

For better understanding of the dataset, there were a few transformation procedures performed on the dataset. Initially, the dataset contained Features for each room in the house, with their respective Temperature Readings and Humidity values. These Features were renamed to reflect a better readability for these measures. Figure 2 shows the relation between each feature variable, with respect to original dataset. (Candanedo, L. et al, 2017)

Data variables and description.

Data variables	Units	Number of features
Appliances energy consumption	Wh	1
Light energy consumption	Wh	2
T1, Temperature in kitchen area	°C	3
RH1, Humidity in kitchen area	%	4
T2, Temperature in living room area	°C	5
RH2, Humidity in living room area	%	6
T3, Temperature in laundry room area	°C	7
RH3, Humidity in laundry room area	%	8
T4, Temperature in office room	°C	9
RH4, Humidity in office room	%	10
T5, Temperature in bathroom	°C	11
RH5, Humidity in bathroom	%	12
T6, Temperature outside the building (north side)	°C	13
RH6, Humidity outside the building (north side)	%	14
T7, Temperature in ironing room	°C	15
RH7, Humidity in ironing room	%	16
T8, Temperature in teenager room 2	°C	17
RH8, Humidity in teenager room 2	%	18
T9, Temperature in parents room	°C	19
RH9, Humidity in parents room	%	20
To, Temperature outside (from Chièvres weather station)	°C	21
Pressure (from Chièvres weather station)	mm Hg	22
RHo, Humidity outside (from Chièvres weather station)	%	23
Windspeed (from Chièvres weather station)	m/s	24
Visibility (from Chièvres weather station)	km	25
Tdewpoint (from Chièvres weather station)	°C	26
Random Variable 1 (RV_1)	Non dimensional	27
Random Variable 2 (RV_2)	Non dimensional	28
Number of seconds from midnight (NSM)	s	29
Week status (weekend (0) or a weekday (1))	Factor/categorical	30
Day of week (Monday, Tuesday... Sunday)	Factor/categorical	31
Date time stamp	year-month-day hour:min:s	–

Figure 2 Data Variable Descriptions

The dataset did not contain any null values within the records; therefore, no transformation or imputation was required for their replacement. Then the DATE column was reformatted to reflect date and time in readable format of Year, Month, Date, Hour, minute, and seconds.

After the initial transformation, I visualized the daily Energy Consumption for the entire four and a half months of data, by using a simple line graph (Figure 3).

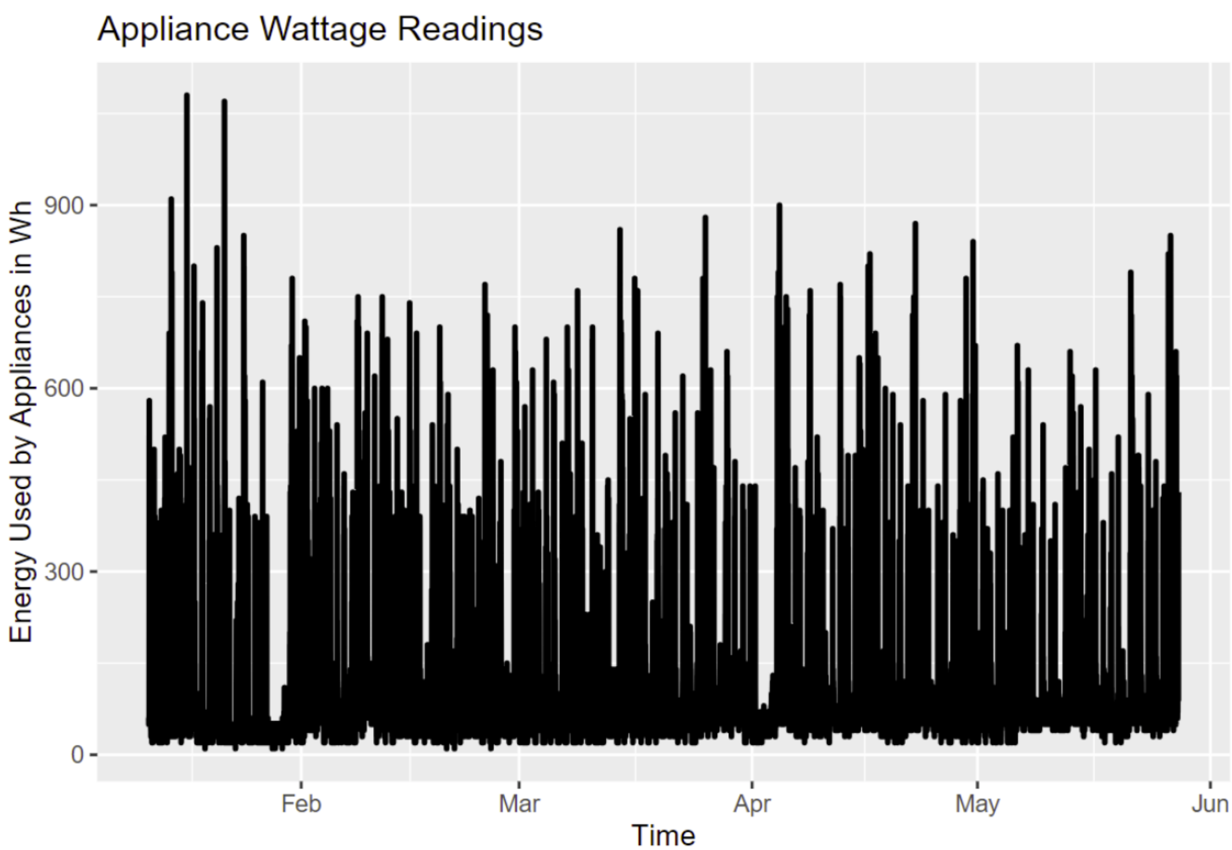

Figure 3 Daily Appliance Energy Consumption in Watts

The new plotted the energy consumption for the first week, using the 10-minute interval data points and getting a view of sample dataset. Here, the visualization shows that there are a few redundant values. The goal was to find the frequency within the dataset to see where the pattern for data values to repeat. (Figure 4).

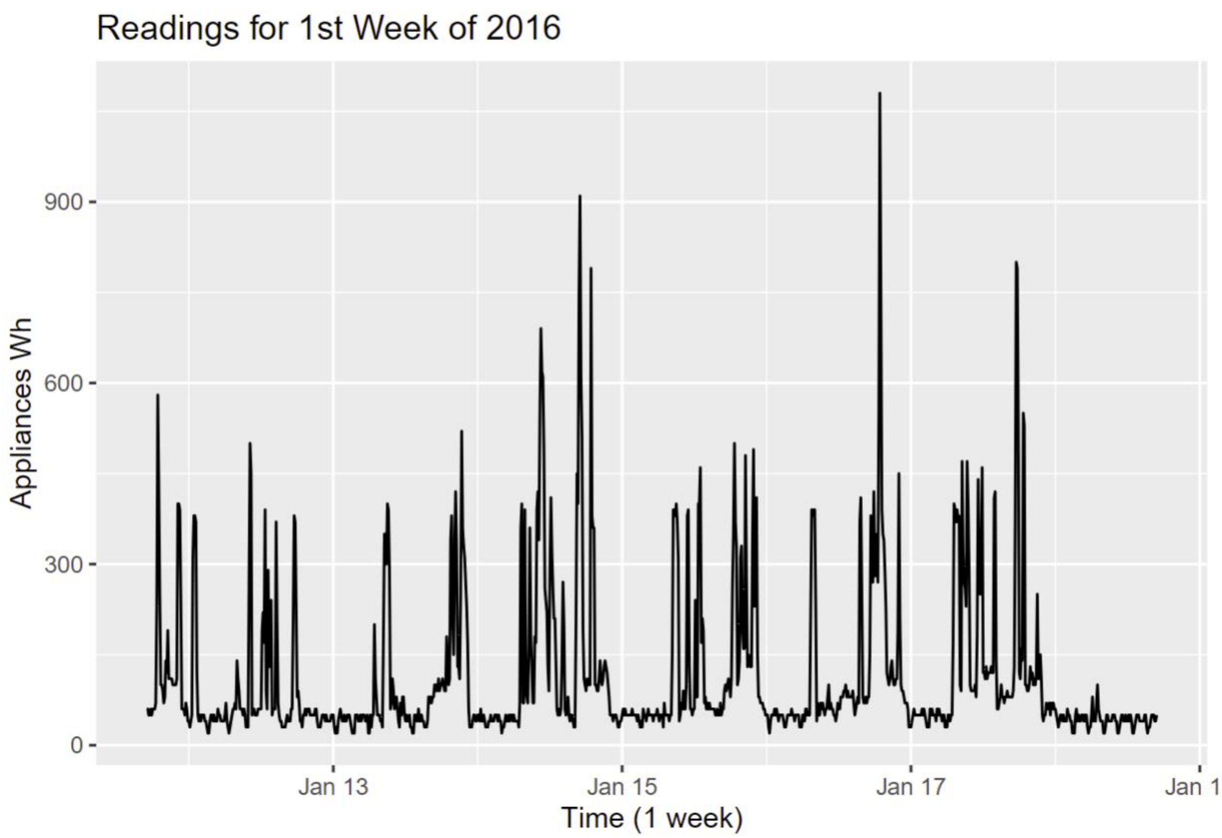

Figure 4 Daily Usage for the First Week of 2016

After transforming the DATE column, I created new variables separately for Date, Time, Hours, Weekday, Weekend, and Second Day measures to compare different measurements of Temperature and Humidity across these varying time values. I also incorporated Day of the week to see if that reflects a different pattern of frequency in our dataset. Figure 5 reflects on the energy consumption by appliances averaged across the days of the week.

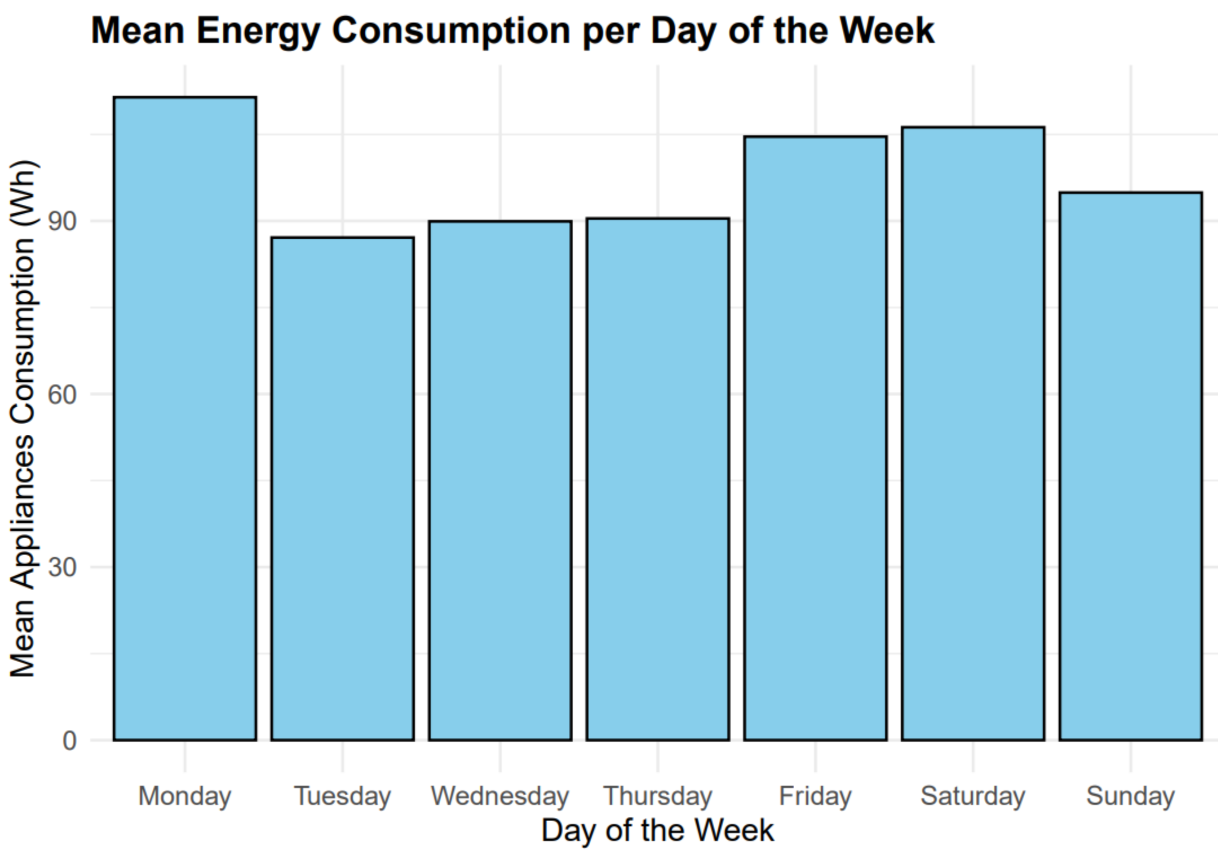

Figure 5 Consumption Average by Day of Week

The energy consumption profile shows a high variability. Figure 6 shows a boxplot of the data. As can be seen the data distribution has a long tail. In the boxplot, the median is represented with a thick black line inside the blue rectangle and has a value of 60Wh. The lower whisker has a value of 10Wh, and the upper whisker has a value of 170Wh. It also shows that the data above the median is more dispersed and that there are several outliers (marked with the round circles above the upper whisker).

The dataset was then framed into a time series data frame, to visualize a time series plot of daily energy consumption in the house across the entire 4.5 months of time span. (Figure 7).

In our analysis of the Appliance Energy Consumption dataset, we explored the application of both smoothing and differencing techniques to enhance the ease of future forecasting. Smoothing was conducted through Holt-Winters (ETS) smoothing. The resultant figures, illustrating the original and smoothed time series, reveal minimal disparities. Notably, the expected benefits of smoothing, such as the removal of drastic outliers, were not prominently observed. Consequently, the decision was made to retain the original Appliance Energy Consumption time series data, ensuring the preservation of inherent fluctuations.

The Smoothed Daily Energy Usage, as shown in Figure 8, did not result in positive impact on the dataset, as it smoothed the data to the mean value across different months. The smoothed Time series for Hourly Energy Usage shows that there of minimal disparities between the original and smoothed datasets (Figure 9).

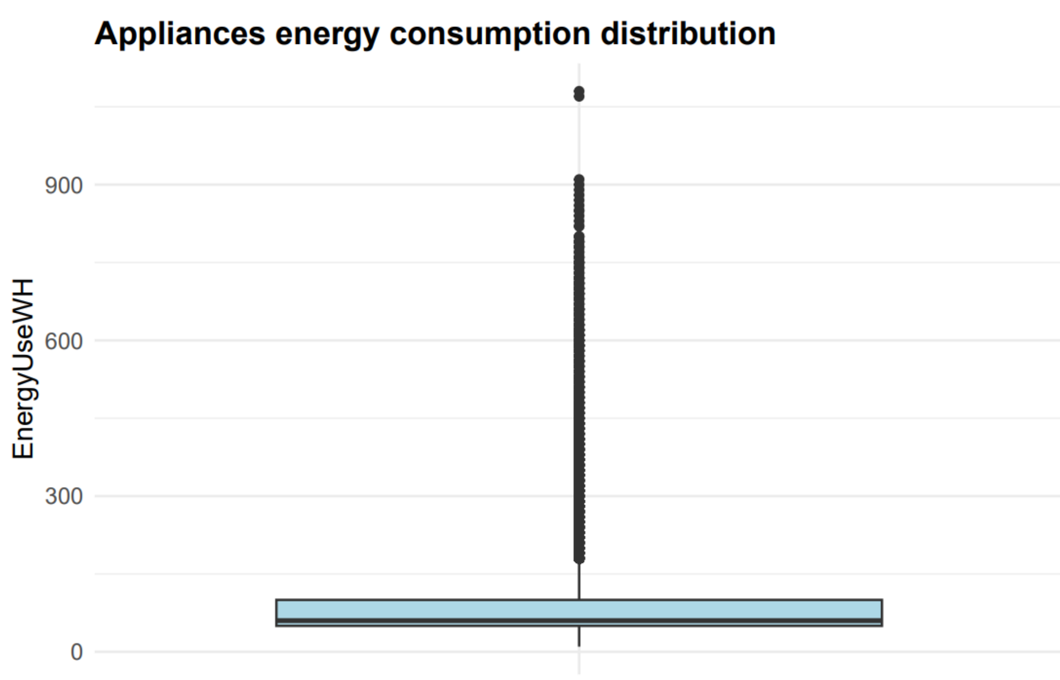

Figure 6 Boxplot for Daily Energy Consumption

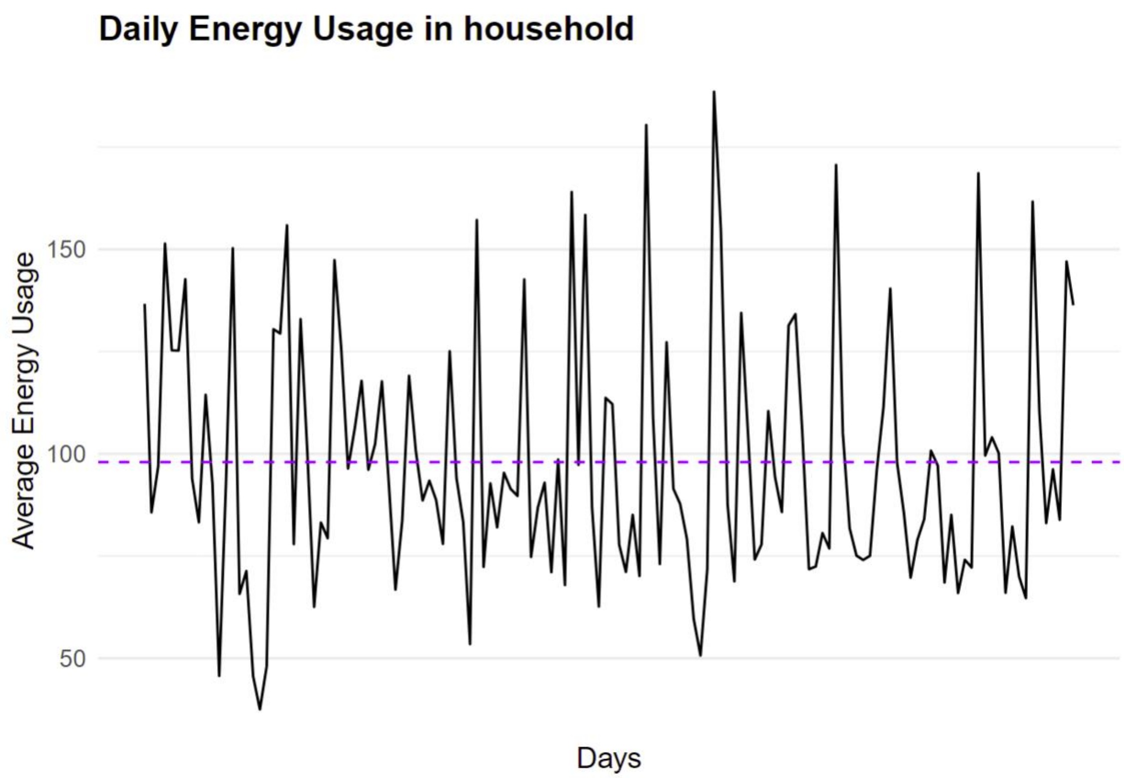

Figure 7 Time Series Plot for Daily Consumption

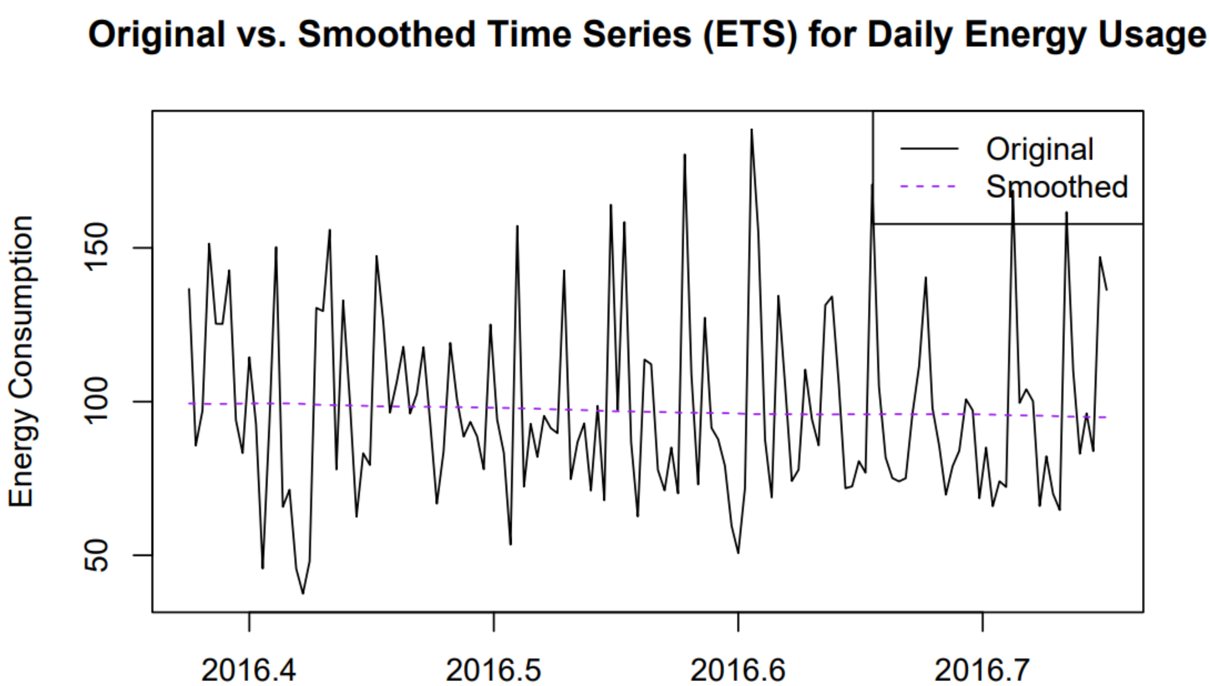

Figure 8 Original vs. Smoothed Time Series (ETS) for Daily Energy Usage

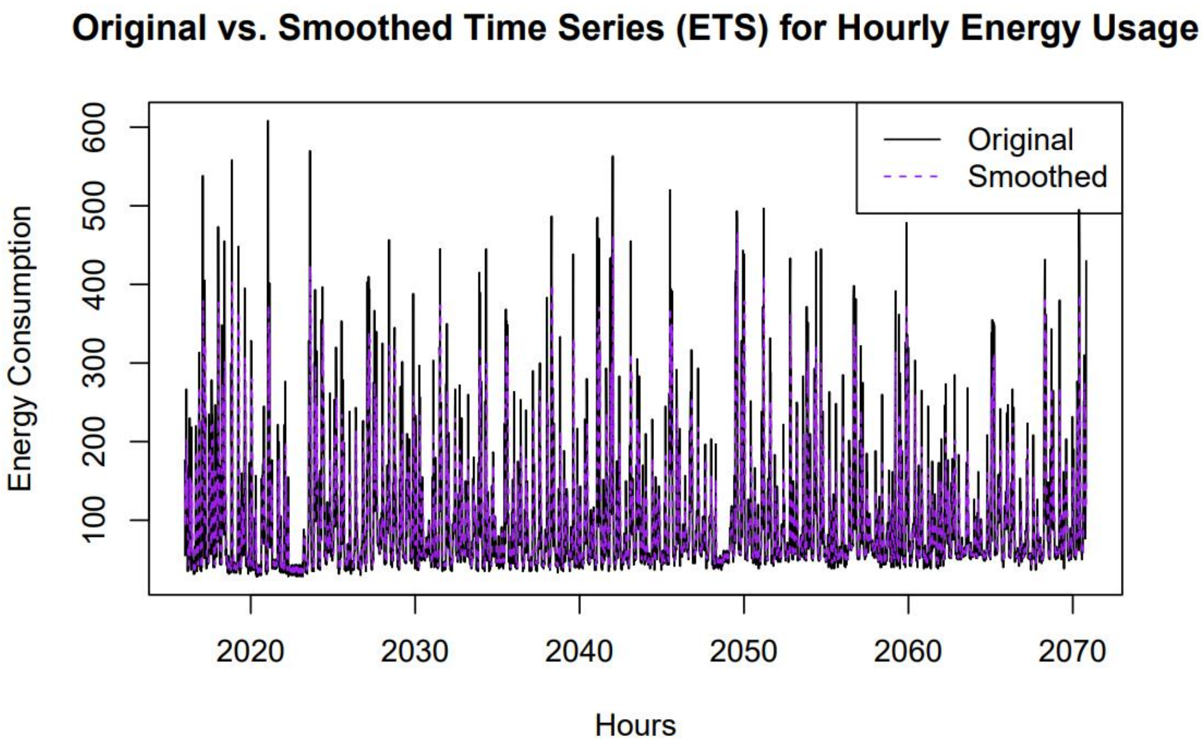

Figure 9 Original vs. Smoothed Time Series (ETS) for Hourly Energy Usage

Additionally, we investigated the impact of differencing on the time series. Figure 10 illustrates that differencing induced significant alterations in the time series, leading to the loss of crucial data related to seasonality. Subsequently, an Augmented Dickey-Fuller (ADF) test was employed to assess stationarity in the original data. While not excessively high, it did not necessitate differencing to enforce stationarity. Considering the substantial loss of important seasonal information associated with differencing, the decision was made to continue utilizing the original time series rather than the differenced series in our analysis of Appliance Energy Consumption.

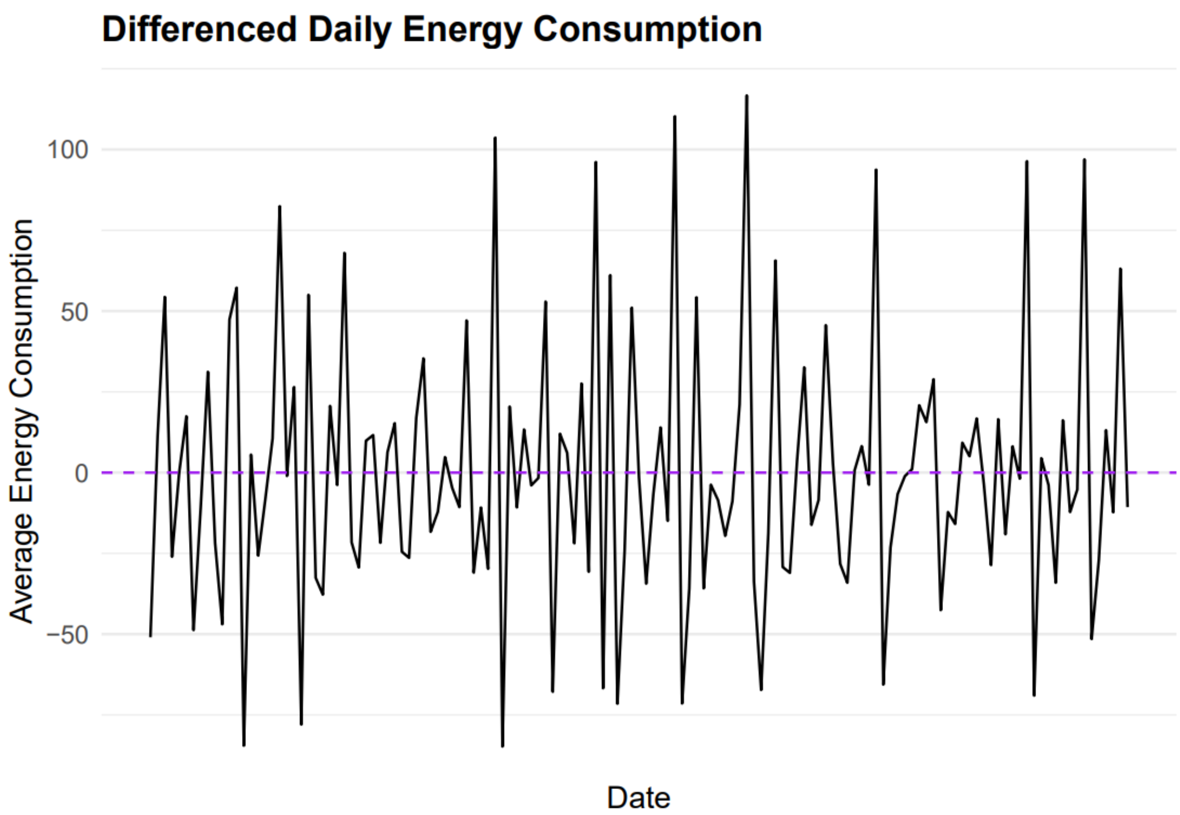

Figure 10 Differenced Daily Energy Consumption

2.3.Exploratory Data Analysis:

Figure 11 illustrates a correlation between various rooms' temperature and their humidity, with the second-largest correlation observed between Teen Room Temperature and Parents Room Temperature. High correlations among indoor temperatures, driven by the HRV unit for ventilation, are evident in the plots, such as the positive correlation between Kitchen and Laundry Room. Positive correlations have been found between Ironing, Teens', and Parents' room. Additionally, a positive correlation is observed between appliance consumption and outdoor temperature (Outside Temp C), indicating higher temperatures coincide with increased energy usage. Negative correlations were found with Outside

Humidity and pressure. Notably, a negative correlation is identified between pressure and wind speed, indicating that lower pressure is associated with higher wind speeds.

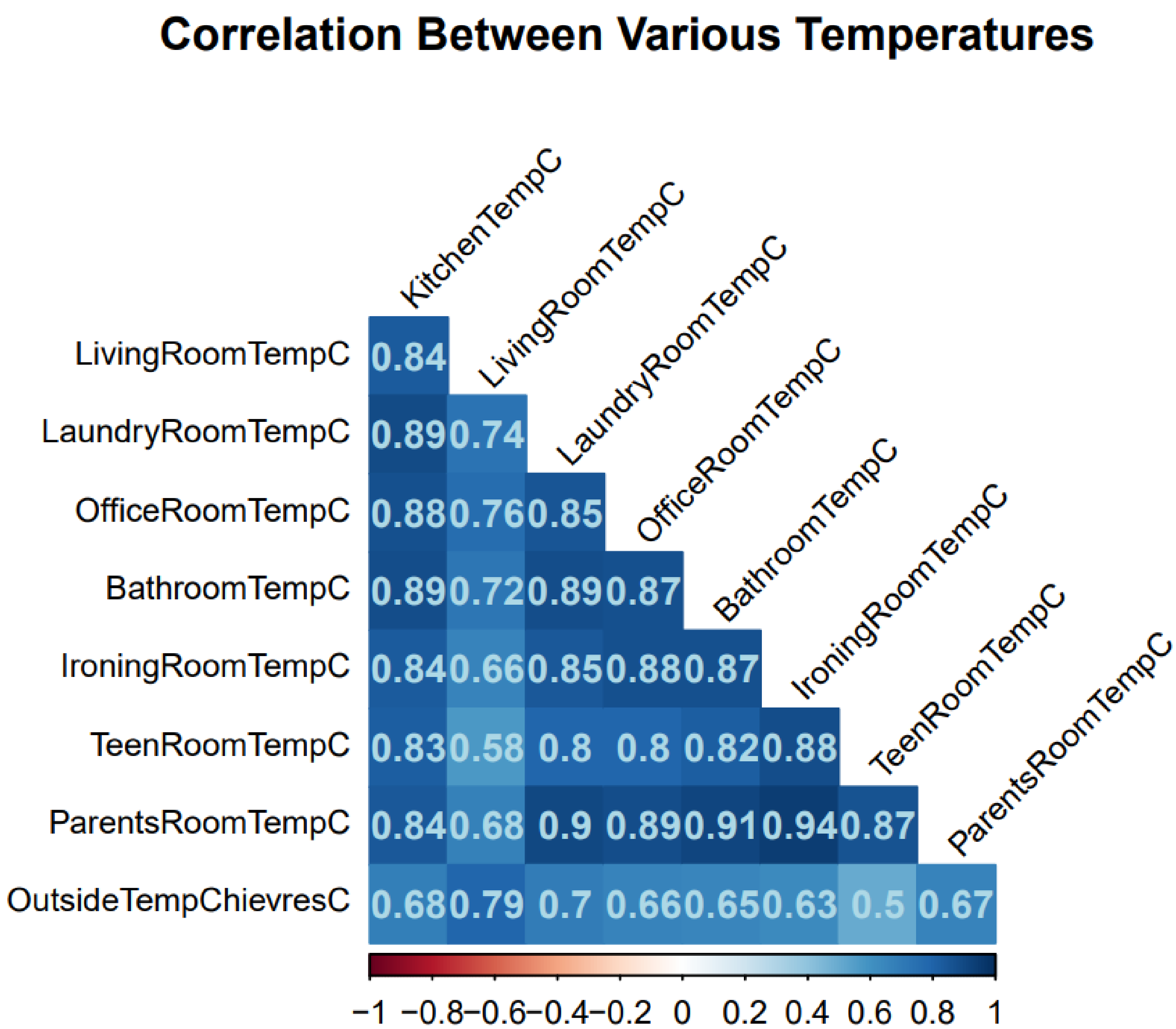

Figure 11 Correlation between Temperatures

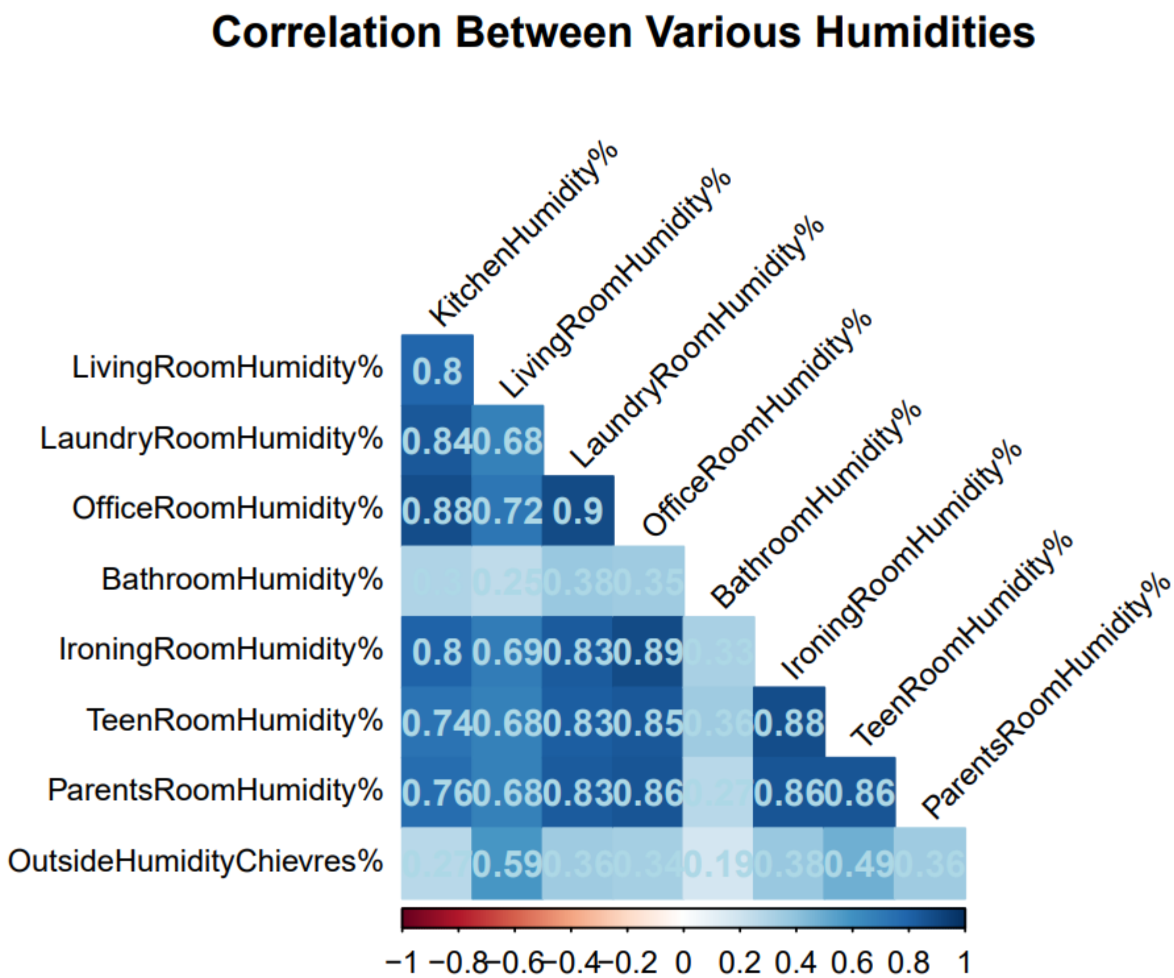

Figure 12 Correlation between Humidities

To explore time trends, a daily heatmap covering four consecutive weeks of data was generated (Figure 13). The heatmap highlights distinct time-dependent patterns in energy consumption, with increases around 6 in the morning, surges around noon, and heightened demand around 6 pm. While no clear pattern is discernible concerning the day of the week, the visualization underscores the pronounced temporal aspect in the energy consumption pattern. (Candanedo, L. et. al., 2017)

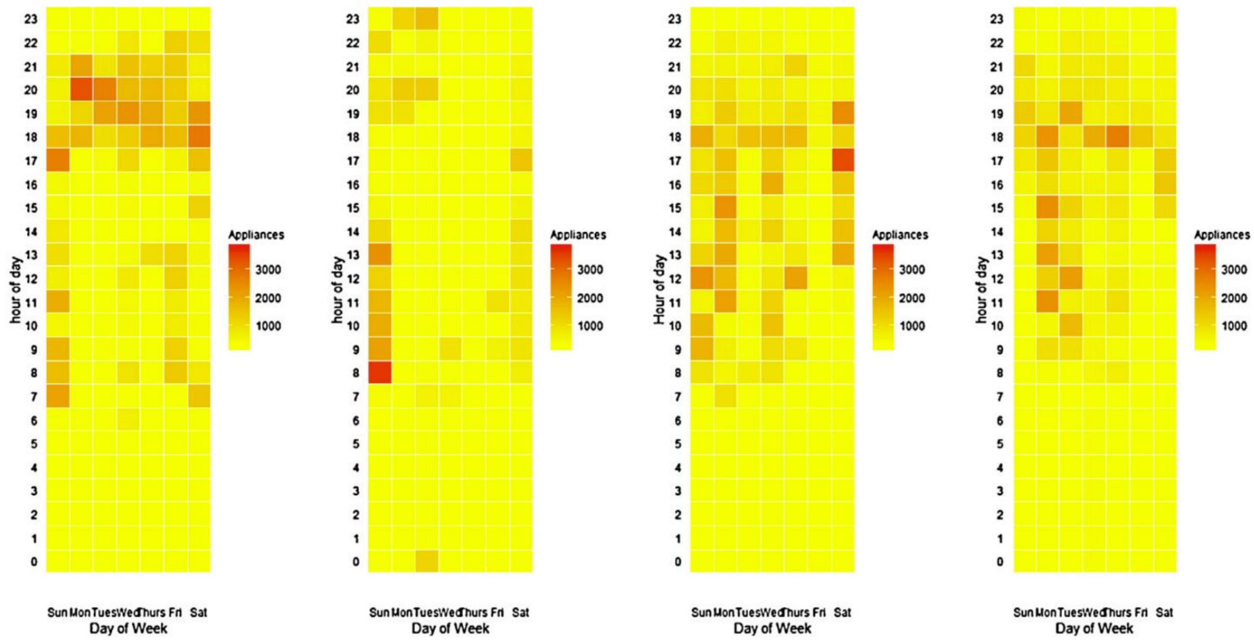

Figure 13 Hourly energy consumption of appliances heat map for four consecutive weeks

In Figure 14, the scatter plot depicting daily appliance energy consumption reveals a diverse distribution of data points. Positioned along the x-axis, each point corresponds to a specific day, while the y-axis represents the corresponding daily energy consumption values. Notably, there is a concentration of points below 150, denoting days with relatively low energy consumption. Simultaneously, the scatter plot portrays a scattering of points across the spectrum, illustrating the varied patterns of energy consumption on different days. The outliers are of significance in the plot, extending beyond 600, indicating instances of exceptionally high daily energy usage. This visualization highlights both typical patterns and extreme values, providing a comprehensive insight into the irregularities of energy consumption behavior over the observed period.

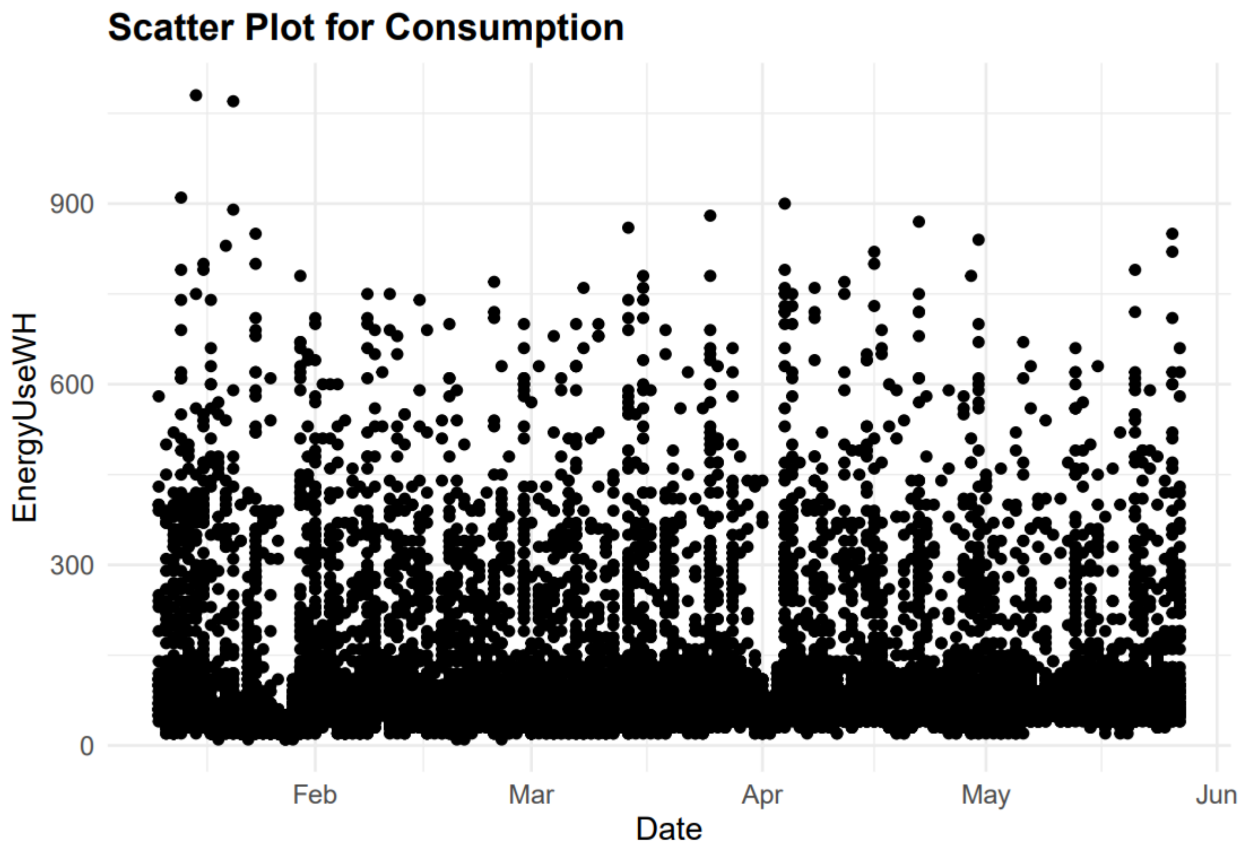

Figure 14 Scatter plot for Daily Energy Consumption

3. Modeling and Evaluation Metrics:

3.1.Energy Consumption Prediction:

This study aims to leverage historical data on appliance energy consumption to forecast future energy usage in households. In the realm of time series forecasting, the objective is twofold: to predict exact future values of energy consumption and to determine whether the values will surpass predefined thresholds. In pursuit of optimal accuracy, we explore both approaches in this paper. To establish baseline accuracy metrics before delving into more sophisticated models, we initiate the analysis with a straightforward naive forecast and a seasonal naive forecast. These initial models provide a foundational understanding of the dataset and help set the stage for the implementation of more intricate forecasting techniques.

An intricate decision in the model development process was the division of the dataset into training (80%) and validation (20%) sets. This division was guided by timestamp indices, ensuring that our models were adept at capturing temporal intricacies and were robust in their ability to generalize to unseen data. The consideration of temporal aspects in the data split aimed to enhance the models' capacity to understand and predict patterns that evolve over time.

Three widely employed metrics for assessing the accuracy of predictive models in the context of predicting appliance energy consumption are the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). RMSE quantifies the average magnitude of errors between predicted and actual values, emphasizing larger errors due to its squared nature. In contrast, MAE calculates the average absolute error magnitude without squaring, offering a more straightforward accuracy measure. MAPE expresses the average percentage difference between predicted and actual values, with a slight bias towards under-forecasts. While each metric provides valuable insights, a comprehensive evaluation of model performance requires considering a combination of these metrics during the validation phase.

In the exploration of time-related patterns in the Appliance Energy Consumption dataset, initial efforts involved the application of straightforward methods, specifically, the Naive and ETS (Seasonal Naive) models. These methods, while serving as fundamental starting points, revealed certain limitations in capturing the intricacies inherent in the dataset.

I initiated the modeling process by implementing a straightforward ETS (Error, Trend, and Seasonality) model after establishing our baseline models. This model, employing simple exponential smoothing, dynamically determines the optimal alpha, representing the smoothing constant for the learning rate. The optimization process involves maximizing likelihood over the training period, which equivalently minimizes the Root Mean Squared Error (RMSE). Once applied to the training time series, the ETS model generates forecast values for the upcoming year and captures residuals by comparing them with actual values.

The Naive model, characterized by its simplicity, provided predictions based on the overall average of the historical data. While this approach offers a basic benchmark, it falls short in discerning the nuanced variations and specific patterns present in the dataset. By relying solely on the overall average, the Naive model fails to capture the dynamic nature of energy consumption over time, overlooking potential trends, seasonality, or recurring patterns.

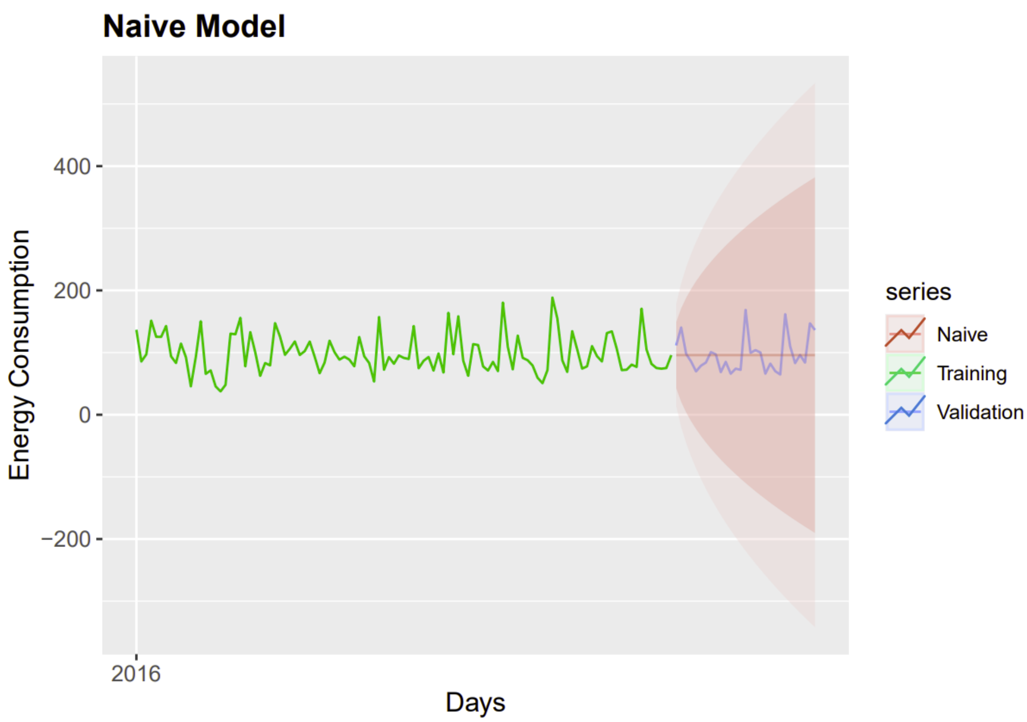

Figure 15 Naive Model Forecasts on Energy Consumption

Similarly, the ETS (Seasonal Naive) model, which introduces seasonality considerations, still operates on a relatively basic level. It relies on historical seasonal patterns to make predictions yet may overlook more complex temporal dependencies and variations that are crucial for accurate forecasting. This limitation underscores the need for more sophisticated models that can delve deeper into the dataset, capturing the subtleties and temporal dynamics that simplistic methods may overlook.

The recognition of these limitations pushed the exploration into more advanced modeling approaches to uncover and leverage the intricate temporal patterns within the Appliance Energy Consumption dataset. The subsequent deployment of more sophisticated models aimed to address these shortcomings and provide a more accurate representation of the underlying dynamics influencing energy consumption over time.

Acknowledging the inherent limitations of simpler methods such as the Naive and ETS (Seasonal Naive) models in capturing the intricate temporal patterns of Appliance Energy Consumption, a strategic pivot was made towards more advanced forecasting techniques. This shift aimed to harness the latent complexities within the time series data and enhance the accuracy and depth of predictions.

The versatility of the ETS model allows for manual adjustments to the smoothing constant and the fitting of error, trend, and seasonality components, tailoring the model to specific instances. The model's performance, assessed through key metrics like RMSE, Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE), demonstrates consistently low values during the testing period. Notably, it outperforms both baseline models – the naive and seasonal naive forecasts – positioning it as a recommended choice. In Figure 16, the time series plot of Appliance Energy Consumption for both the training and testing periods showcases the ETS model predictions against actual values. The precision evident in the predicted values, as depicted in the plot, surpasses that of the baseline models, affirming the model's effectiveness, as corroborated by the low error metrics.

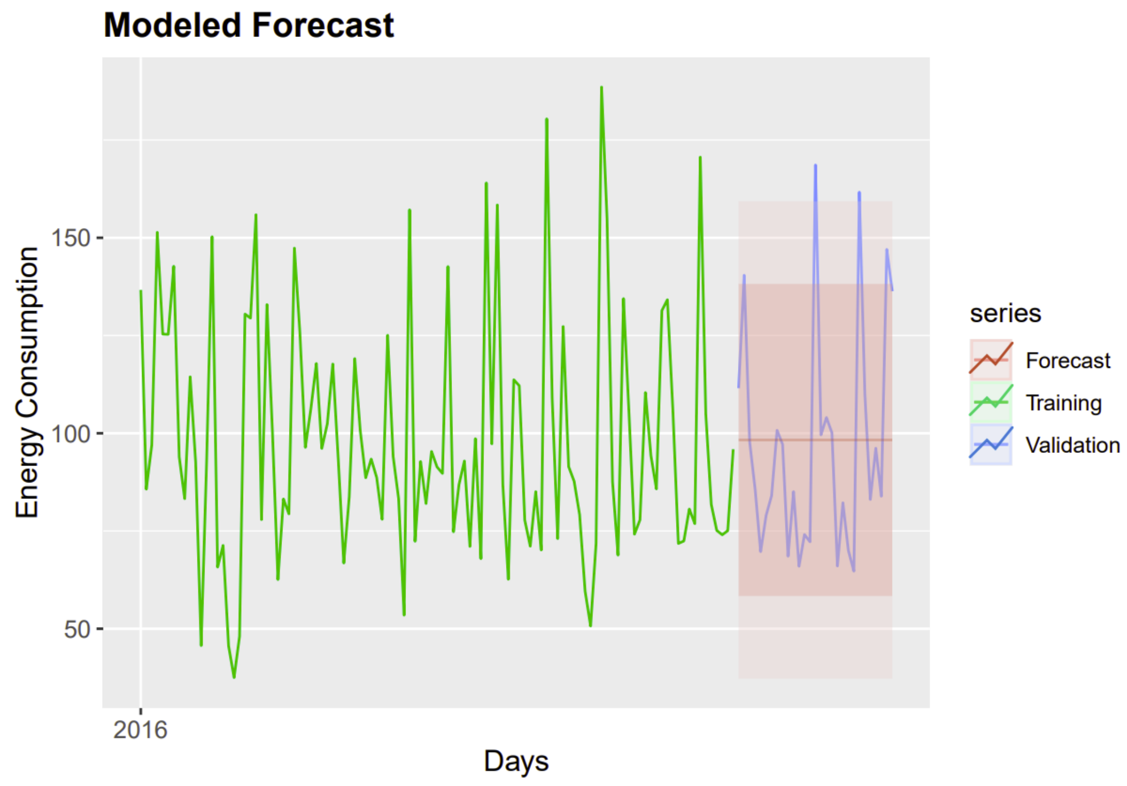

Figure 16 Modeled Forecasts with External Predictors

In addition to the ETS model, our research incorporates two distinct Autoregressive Integrated Moving Average (ARIMA) models for Appliance Energy Consumption prediction. ARIMA models encompass both the historical relationship between the target variable and its past values (autoregressive component) and the integration of forecast errors from recent predictions to refine subsequent predictions (moving average component). These models can utilize only the target variable's time series data or include external time series as well.

The initial ARIMA model focused solely on the past Appliance Energy Consumption data. By testing various autoregressive, differencing, moving average, and seasonality parameters (p, d, q) (P, D, Q), we identified the optimal model, designated as ARIMA (2,1,1)

(0,1,0). Subsequently, the scope of the second ARIMA model was expanded by incorporating external variables of temperature, to forecast future Appliance Energy Consumption values. Through thorough parameter testing, the optimal model for this scenario emerged as ARIMA (1,0,1).

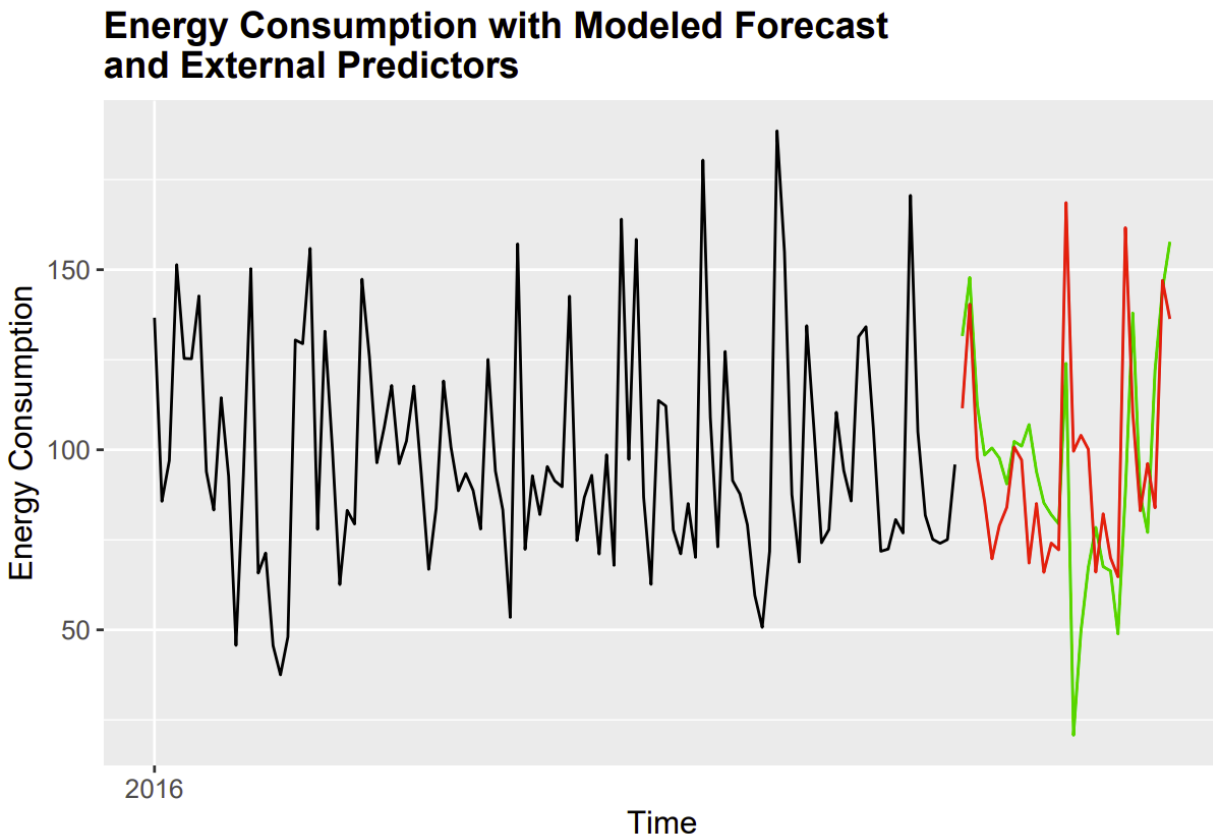

Figure 17 ARIMA with External Predictors

In addition to traditional time series models, the exploration expanded to incorporate Logistic Regression into the analytical toolkit. This alternative approach introduced a different paradigm, focusing on the probability of an event occurrence. The integration of Logistic Regression broadened the spectrum of tools available, enabling a more comprehensive exploration of trends and patterns within the Appliance Energy Consumption time series.

Each model brought its unique perspective and strengths to the analytical forefront. The collective insights derived from ARIMA, ETS, and Logistic Regression contributed to a richer understanding of the complex temporal dynamics governing energy consumption patterns. The diverse approaches employed in this exploration ensured a more holistic evaluation of the dataset, paving the way for informed decision-making and more accurate predictions. This multifaceted modeling strategy underscored the importance of adapting to the intrinsic complexities of time series data, ultimately enhancing the predictive capabilities of the analysis.

Figure 18 offers a comprehensive representation of the RMSE, MAE, and MAPE metrics for the two ARIMA models discussed. Remarkably, when gauging these metrics, the initial ARIMA model demonstrated notably inferior performance compared to the ETS model. Conversely, the ARIMA model, enriched by the inclusion of external predictors, exhibited a slightly enhanced performance in comparison to the ETS model. The graphic depiction of actual versus predicted forecast values for the ARIMA model with external predictors, showcased below, provides valuable insights into the predictive accuracy of this model.

Drawing upon established methodologies, I carefully considered an optimal model to forecast appliance wattage with precision. The combination of ARIMA and ETS, boosted by exogenous variables, served as the foundation for achieving strong predictive capabilities. At the heart of this model was the Exogenous ARIMA framework. This approach eliminated the limitations of historical data by incorporating temperature readings from various rooms within the house. Furthermore, our predictive scope expanded to encompass temperature insights from the external environment, sourced from a nearby airport weather station.

Table 1: Model Evaluation

Model	RMSE	MAE	MAPE	PCT_BIN
Naive	28.44332	21.83227	22.52375	48.27586
ETS	28.47278	22.16847	23.43231	48.27586
ARIMA	28.47277	22.16842	23.43221	48.27586
ARIMA with External Predictors	29.71261	22.15417	22.99229	64.28571
Logistic Regression with External Predictors	NA	NA	NA	72.41379

Figure 18 Model Evaluations

4. Discussion:

This study encounters several limitations that should be acknowledged. Firstly, the analysis is confined to a single house, which restricts the generalizability of findings. A more comprehensive understanding of appliances' energy consumption patterns could be gleaned from examining multiple houses, considering additional factors such as occupant demographics, pet ownership, and building geometry. Moreover, the study is limited by the duration of continuous analyzed data, potentially overlooking seasonal variations in energy use patterns. Exploring different seasons could uncover valuable insights into temporal dynamics.

Another constraint pertains to the proximity of the weather station. Improved predictions of appliances' energy use could be achieved with a weather station closer to the house. The optimal placement of wireless indoor sensors for enhanced energy prediction remains unexplored in this research, posing a potential area for future investigation. Additionally, the study does not delve into the potential benefits of employing more sensors or enhancing sensor accuracy, which could contribute to refining the precision of energy predictions.

These limitations underscore the need for future research endeavors to address these aspects and broaden the scope of insights into appliance energy prediction.

5. Conclusion:

In this paper, the examination of various predictive models, including the Naive, ETS, ARIMA, ARIMA with External Predictors, and Logistic Regression with External Predictors, has yielded valuable insights into forecasting appliance energy consumption. Pairwise comparisons and performance metrics reveal distinct characteristics of each model.

The Naive model, providing a simplistic baseline, serves as a straightforward benchmark. However, its reliance on overall averages may overlook nuanced patterns within the data. Moving beyond basic approaches, the ETS model leverages exponential smoothing, showcasing improved accuracy with precise value predictions. It outperforms the Naive model, offering a more nuanced understanding of the data's intricacies.

ARIMA models, considering both past relationships and forecast errors, contribute to the predictive landscape. The basic ARIMA model, focusing solely on past energy consumption values, demonstrates the challenges of predicting complex patterns. Incorporating external predictors enhances the second ARIMA model, showcasing improved metrics, although not surpassing the ETS model.

The Logistic Regression model with External Predictors emerges as a robust contender, achieving the highest accuracy when determining energy consumption levels above or below a specified threshold. Despite its success, the choice between models requires a balanced consideration of accuracy and granularity. While the Logistic Regression model excels in accuracy, it sacrifices the precision of exact value predictions evident in the ETS model.

The analysis and modeling revealed significant relationships among parameters, with ARIMA and ETS models outperforming others. Weather data significantly enhanced predictions. The wireless sensor network, particularly in specific rooms, contributed significantly. Future work may explore additional weather data, occupancy details, and expanded wireless sensor capabilities for more comprehensive predictions and occupant behavior tracking. In conclusion, the selection of the optimal model hinges on the specific goals and trade-offs in accuracy and granularity. The ETS model stands out for its nuanced predictions, balancing accuracy, and the granularity necessary for a comprehensive understanding of the complex temporal dynamics inherent in predicting appliance energy consumption.

References:

Candanedo, Luis. (2017). Appliances energy prediction. UCI Machine Learning Repository. https://doi.org/10.24432/C5VC8G

A. Barbato, A. Capone, M. Rodolfi and D. Tagliaferri (2011). *Forecasting the usage of household appliances through power meter sensors for demand management in the smart grid*, IEEE International Conference on Smart Grid Communications (SmartGridComm), Brussels, Belgium, 2011, pp. 404-409, doi: 10.1109/SmartGridComm.2011.6102356.

S. Firth, K. Lomas, A. Wright, R. Wall (2008). *Identifying trends in the use of domestic appliances from household electricity consumption measurements, Energy and Buildings*, Volume 40, Issue 5, 2008, Pages 926-936, ISSN 0378-7788, https://doi.org/10.1016/j.enbuild.2007.07.005. https://www.sciencedirect.com/science/article/pii/S0378778807002022

Solar Energy Technologies Office (2023). *Photovoltaics*. https://www.energy.gov/eere/solar/photovoltaics

Kristen S. Cetin (2016) *Characterizing large residential appliance peak load reduction potential.*

utilizing a probabilistic approach, Science and Technology for the Built Environment, 22:6, 720-732, DOI: 10.1080/23744731.2016.1195660

Abdo A. A. Gassar, & Seung H. Cha. (2020). *Energy prediction techniques for large-scale* *buildings towards a sustainable built environment: A review*. *Energy and Buildings, 224*, 110238. https://doi.org/10.1016/j.enbuild.2020.110238

U.S Energy Information Administration (2022). *Annual Energy Outlook 2022*. Retrieved from: https://www.eia.gov/outlooks/archive/aeo22//

Pratt, R. G., Conner, C. C., Cooke, B. A., & Richman, E. E. (1993). *Metered end-use consumption and load shapes from the ELCAP residential sample of existing homes in the Pacific Northwest. Energy and Buildings*, 19(3), 179-193. https://doi.org/10.1016/0378-7788(93)90026-Q

Richardson, I., Thomson, M., & Infield, D. (2008). A high-resolution domestic building occupancy model for energy demand simulations. Energy and Buildings, 40(8), 1560-1566. https://doi.org/10.1016/j.enbuild.2008.02.006.

Fumo, N., Mago, P., & Luck, R. (2010). Methodology to estimate building energy consumption using EnergyPlus Benchmark Models. Energy and Buildings, 42(12), 2331-2337. https://doi.org/10.1016/j.enbuild.2010.07.027.

Jones, R. V., Fuertes, A., & Lomas, K. J. (2015). The socio-economic, dwelling, and appliance-related factors affecting electricity consumption in domestic buildings. Renewable and Sustainable Energy Reviews, 43, 901-917. https://doi.org/10.1016/j.rser.2014.11.084.

Hor, Ching-Lai, Watson, S. J., & Majithia, S. (2005). Analyzing the impact of weather variables on monthly electricity demand. IEEE Transactions on Power Systems, 20(4), 2078-2085. doi: 10.1109/TPWRS.2005.857397.

Feist, W. (2007). Passive House Planning Package. Passive House Institute. Retrieved from https://www.eerg.it/passive-on.org/CD/4.%20PHPP%202007%20Demo/Contents_Manual.pdf

Candanedo, L. M., Feldheim, V., & Deramaix, D. (2017). Data-driven prediction models of energy use of appliances in a low-energy house. *Energy and Buildings*, *140*, 81-97. https://doi.org/10.1016/j.enbuild.2017.01.083.