Practical Decision-Making in Electricity Consumption Forecasting: Insights from Combined Models

Jéssica Catarina Cristina Pires

152022014

Dissertation written under the supervision of professor Ana Guedes,
with the collaboration of industry expert Ana Airports

Dissertation submitted in partial fulfilment of requirements for the MSc in Business Analytics, at the Universidade Católica Portuguesa,

January 2024.

Abstract

This study focuses on forecasting daily electricity consumption at Faro Airport, utilizing machine learning models such as Random Forest (RF) and AutoRegressive Integrated Moving Average with exogenous variables (ARIMAX). In addition to demonstrating the forecasting prowess of these models, this study evaluates their performance by comparing them with the ensembled model blending RF and ARIMAX. Performance metrics not only gauge the accuracy of predictions but also provide a nuanced understanding of the models' effectiveness. The hybrid model emerges as a standout performer, showcasing superior forecasting precision. Its ability to leverage the strengths of both RF and ARIMAX contributes to more robust predictions, especially in the context of daily electricity consumption at Faro Airport. Beyond numerical accuracy, the study incorporates Shapley values for interpretability, offering a transparent view of the factors influencing electricity consumption trends. This interpretability aids airport management in making informed decisions related to energy resource allocation, infrastructure planning, and operational efficiency. This approach not only optimizes resource utilization but also positions the airport to proactively address challenges and opportunities in its energy consumption patterns. The combination of accurate predictions and interpretable insights positions Faro Airport to not only meet current energy demands effectively but also to navigate future challenges with strategic foresight.

Title: Practical Decision-Making in Electricity Consumption Forecasting: Insights from Combined Models

Author: Jéssica Catarina Cristina Pires

Keywords: Consumption; Electricity; Airport; Machine Learning; Time Series; Accuracy; Performance; Faro; Portugal

Resumo

Este estudo concentra-se na previsão do consumo diário de eletricidade no Aeroporto de Faro, utilizando modelos de Machine Learning, Random Forest (RF) e AutoRegressive Integrated Moving Average com variáveis exógenas (ARIMAX). Para além de demonstrar a capacidade de previsão desses modelos, o estudo avalia o seu desempenho, comparando-os com o modelo híbrido que combina RF e ARIMAX. As métricas de desempenho não só medem a precisão das previsões, como também proporcionam uma melhor compreensão da eficácia dos modelos. O modelo híbrido destaca-se com um desempenho excepcional, exibindo uma boa precisão. A sua capacidade de aproveitar os pontos fortes de RF e ARIMAX contribui para previsões mais robustas, especialmente no contexto do consumo diário de eletricidade no Aeroporto de Faro. Além da precisão numérica, o estudo incorpora valores de Shapley para interpretabilidade, oferecendo uma visão transparente dos fatores que influenciam as tendências de consumo de eletricidade. Essa interpretabilidade auxilia a gestão do aeroporto na tomada de decisões informadas relacionadas à alocação de recursos energéticos, planeamento de infraestrutura e eficiência operacional. Esta abordagem, para além de otimizar a utilização de recursos, também posiciona o aeroporto para abordar proativamente desafios e oportunidades nos seus padrões de consumo de energia. A combinação de previsões precisas e insights interpretáveis permite ao Aeroporto de Faro responder eficazmente às necessidades energéticas atuais, mas também preparar futuros desafios com uma visão estratégica.

Título: Tomada de Decisões Práticas na Previsão do Consumo de Eletricidade: Insights de Modelos Combinados

Autor: Jéssica Catarina Cristina Pires

Palavras-chave: Consumo; Eletricidade; Aeroporto; Machine Learning; Séries Temporais; Precisão; Desempenho; Faro; Portugal

Acknowledgments

I would like to express my sincere gratitude to my family, my mother, Paula Pires, and my father, Paulo Pires, for their unwavering support throughout the entire thesis journey. Their encouragement and understanding have been invaluable, providing me with the strength and motivation needed to complete this endeavor.

A special appreciation goes to my partner, Bruno Pereira, whose constant encouragement and belief in my abilities have been a source of inspiration. His unwavering support has made this academic journey all the more meaningful.

I am also deeply thankful for the support of all my friends who stood by me during the highs and lows of this thesis. Their encouragement and fellowship have made this academic pursuit a truly enriching experience.

I extend my heartfelt thanks to Professor Ana Guedes for her guidance, expertise, and valuable insights throughout the research process. Her mentorship has played a pivotal role in shaping the outcome of this work.

A profound appreciation is extended to the leadership and staff of Faro Airport for their assistance and the provision of essential materials. I would like to express my gratitude, in particular, to Business Unit Controller Paulo Romão for his generous support and willingness to share his expertise, which greatly contributed to the success of this research.

Thank you to everyone who has played a role, no matter how big or small, in making this academic pursuit a reality. Your support has been instrumental, and I am truly grateful for the collaborative spirit that has enriched this academic journey.

May this project be more than the outcome of my effort and dedication, but also a constant reminder that I am capable of overcoming challenges and achieving great accomplishments. The journey does not end here; inspiration persists for new horizons and achievements.

Thank you.

Jéssica Pires

Contents

List of Figures

List of Tables

1 Introduction

Airports, serving as pivotal hubs for air travel, are intricate infrastructures characterized by high energy consumption. The incessant activity of aircraft arrivals, departures, and movement of passengers places a substantial demand on electricity. With operations spanning approximately 20 hours a day, airports incur significant costs associated with electricity consumption.

The fundamental components of an airport include runways, the passenger terminal, where boarding and disembarkation occur, the cargo terminal for handling goods, and auxiliary structures such as hangars, workshops, and maintenance facilities. The seamless operation of an airport relies on the harmonious interplay of these functionalities and conditions.

Efficient electricity consumption is imperative for optimizing operational costs. Forecasting daily electricity consumption emerges as a strategic imperative for airport managers seeking to enhance resource allocation and cost-effectiveness. As an intern at Faro Airport, my daily exposure to the intricacies of airport operations has heightened my awareness of the critical role electricity plays in maintaining seamless functionality. Recognizing that electricity costs contribute significantly to operational expenses, the motivation for this study stems from the desire to contribute actionable insights to enhance the efficiency of electricity consumption. By understanding the underlying patterns and factors influencing electricity usage, I aim to empower airport managers with informed decision-making tools for resource allocation and cost-effectiveness.

This study zeroes in on the electricity consumption patterns of Faro Airport, delving into the factors that exert notable influence. The primary objectives of the study are framed by key questions:

1. Does the integration of machine learning and time series models exhibit superior performance metrics when compared to each individual model?
2. Are the insights derived from these models useful in making practical decisions or formulating strategies?
3. Do the interpretability methods employed shed light on any instability or uncertainty inherent in the predictions of the models?

The study follows a structured outline. In the literature review, an exploration of existing literature on electricity consumption in airport operations occurs, with an examination of studies on the integration of machine learning and time series models, along with a review of global strategies employed by airports to optimize electricity usage and reduce costs. The methodology section provides details on the data collection process, including the types of data gathered from Faro Airport. It explains the application of machine learning and time series models for electricity consumption forecasting, outlining specific methods used for model integration and performance evaluation, along with the interpretability. Transitioning to the results, the study presents findings regarding electricity consumption patterns at Faro Airport. It compares performance metrics of integrated models with individual models and highlights practical insights derived from the models, discussing their implications for decision-making and strategy formulation, based on the global and local interpretability of the models. In the discussion, the analysis focuses on the significance of the study's findings in the broader context of airport operations. Practical implications of the insights gained from the models are discussed, along with an exploration of the interpretability methods employed and their role in identifying instability or uncertainty in predictions. This comprehensive approach aims not only to enhance the understanding of electricity consumption dynamics in airport operations but also to provide practical tools and insights for optimizing resource management and cost-efficiency in airport facilities. The study contributes meaningfully to the ongoing discourse on sustainable and efficient airport operations.

2 Literature Review

2.1 Electricity consumption of Airports

According to the sustainable report from ANA airports, it's evident that the aviation industry carries significant responsibility for global energy consumption and greenhouse gas emissions, accounting for approximately 2% of worldwide emissions. These statistics underscore the importance of managing electricity consumption in this sector. The report also provides specific insights into the electricity consumption at airports in Portugal for the year 2022. The total electricity consumption for all airports in the country amounted to a substantial 375944.36GJ. Notably, Faro Airport's consumption contributed significantly, making up 9.2% of this total with a consumption of 34493.45GJ. An interesting development in the report is the inauguration of a solar plant at Faro Airport in the same year. This solar plant has the capacity to provide about 30% of the airport's electricity consumption, marking a pioneering effort in Portugal and setting an example for other airports seeking to reduce their carbon footprint and operational costs through sustainable energy solutions (ANA, 2022).

Understanding and managing the electricity consumption characteristics of terminal devices are crucial for optimizing the efficiency of primary equipment systems and reducing operational expenses. To achieve this, various approaches have been explored, particularly in the context of predicting electricity consumption in public buildings like airports.

Electric power load is a stochastic process, which means that it is unpredictable and can vary significantly over time. So, probabilistic load forecasting is an approach that provides a probability distribution for future power load and is a valuable tool for power demand management (Hong & Fan, 2016).

According to Ortega Alba and Manana (2016) , the dynamics of energy consumption in airports are multifaceted and influenced by various factors, rendering them stochastic, nonlinear, and dynamic. Given the substantial impact of energy costs on airport operations, managers continually seek ways to reduce consumption. A detailed understanding of energy use and behavior is crucial, allowing for precise forecasts and identification of unnecessary expenditures or areas requiring improvements.

Airports worldwide have employed various prediction models to forecast their energy consumption. These models encompass a wide range of techniques, including sample data collection, statistical analysis, and machine learning methods like neural networks and decision trees.

2.1.1 Patterns and data features

Data mining is a powerful tool that can be used to extract valuable insights from large datasets and can be used to improve operational efficiency and decision-making in commercial buildings, including airports (Xiaowei, 2008).

In an attempt to understand the components with the highest energy expenditure in an airport, studies were sought that analyzed these consumptions and, consequently, the variables that would have the greatest impact on predicting electricity consumption. In China, Costa et al. (2012) concluded that around 28% of electricity consumption is due to lighting and 20% to Heating Ventilation and Air Conditioning (HVAC), with the remainder distributed among general airport activities. Ortega Alba and Manana (2017), in a study carried out to characterize energy demand patterns in airports, concluded that airport energy consumption is influenced by several factors, including climate, air traffic, passenger numbers, and time of day, and the main energy consumers in airports are HVAC systems, lighting, and auxiliary equipment, confirming what the previous study stated. With these findings, managers can do an effective energy management in airports to reduce costs, improve sustainability, and ensure the efficiency of operations. To do this, it is crucial that airports invest in energy efficiency measures, such as installing more efficient HVAC and lighting systems, using renewable energy, and optimizing operations.

Yang et al. (2023) characterized the behavior of energy consumption at Beijing Capital Airport and concluded, using a combination between regression and K-means cluster, and then Pearson's correlation analysis method, that electricity consumption can essentially be divided into three categories: consumption in the terminal, consumption related to passenger flow and consumption related to outside temperature. Detailed analyses have revealed that heating, ventilation, and air conditioning (HVAC) systems, along with lighting, are the primary contributors to airport energy consumption. Furthermore, temperature, daily light hours, and aircraft operations are identified as the most influential factors affecting energy consumption, providing valuable insights for energy management strategies.

At the 9th Airport Terminal Maintenance & Facilities Management Summit (2023), Cem Onater, electrical systems manager at TAV airports, namely at Izmir airport in Turkey presented a detailed daily forecast of their consumption for the following day in order to access lower energy prices. They analysed different performances for different factors in a multiple regression model and the factors that showed the best performance in their study was temperature and number of passengers.

2.1.2 Modelling approaches

The choice of prediction model largely depends on the specific requirements of the forecasting task, including the timeframe of the prediction. For operational and short term predictions, methods like fuzzy logic, neural networks, time series decomposition, and probabilistic techniques are commonly used. For medium and long-term predictions, more complex machine learning algorithms, such as linear regression or moving averages, may be preferred, balancing the need for accuracy with model interpretability (Klyuev et al., 2022).

In this realm, it’s worth noting that there is an abundance of methods available for predicting electricity consumption, with more than 200 distinct techniques reported. However, many of these methods are variations or combinations of existing models, highlighting the constant innovation and adaptation in the field of energy consumption forecasting (Klyuev et al., 2022).

When dealing with datasets exhibiting cyclical patterns and temporal dependencies, it is imperative to incorporate this information accurately to enhance the precision and robustness of predictions. Simply creating numeric variables to represent a temporal component, for identifying days, weeks, months, or quarters, can lead to misinterpretations, as the numeric value associated with the day does not necessarily reflect its position relative to the month. Instead of adopting this conventional approach, the application of the “Cyclic Encoding” technique stands out for transforming and preserving the cyclical nature of these variables. This technique employs trigonometric functions such as sine and cosine to represent these variables, preserving the continuity and intrinsic periodicity of temporal features. This allows the model to understand the nuances of seasonal patterns more accurately (Bescond, 2021).

When modeling, two different approaches can be used: using the time series with only the consumption target variable or using a dataset with the target variable and several explanatory variables. The choice between these two lies in the cost/precision trade-off, since the second approach requires a lot of historical data and much more laborious data preparation.

Using the first approach, a study showed that the Unbiased Grey Markov Predictive Model (UGMPM) is an effective method for predicting airport's monthly energy consumption. The UGMPM is able to capture the nonlinear and stochastic dynamics of airport energy consumption and it seems more accurate comparing to the traditional GM (1,1) prediction model (TGPM) and the Unbiased GM (1,1) prediction model (UGPM). UGMPM, for one of the five airports in the study, had an RMSE of $2323.4kWh$, while the RMSE of TGPM and UGPM was $17697.9kWh$ and $18289.9kWh$, respectively (J. Chen & Xie, 2013).

In this respect, although not in an airport environment, it is also possible to use an Auto Regressive Integrated Moving-Average (ARIMA) model, which takes into account the lags of the variable, the auto correlation function and the partial correlation functions of residuals. An extensive study, employing 12 different methods on the time series of consumption, was conducted using simple ARIMA models, Holt Winters (HW) decompositions and Exponential Smoothing State Space (ETS) models, with different parameters and bagging them, across three countries (Brazil, Mexico, and Turkey). The Remainder Sieve Bootstrap (RSB) ARIMA stood out in two out of three countries, achieving a MAPE of 4.359% in Brazil and 3.041% in Mexico (De Oliveira & Cyrino Oliveira, 2018). However, since energy consumption is very volatile due to a number of factors, it is possible to apply the AutoRegressive Integrated Moving Average with eXogenous variables (ARIMAX) model to combine in the model information on the historical values of the consumption and variables that prove to be significant in determining it. ARIMAX extends the traditional ARIMA model by incorporating additional external variables or predictors, allowing for a more comprehensive analysis and better forecasting accuracy. For example, in a study carried out in the city of Tallahassee, 3 cases were created using these techniques: ARIMA model, ARIMAX model with weather and ARIMAX model with weather and traffic. The results of the study showed that the latter was able to obtain much more accurate forecasts, with a MAPE of 3.3%, compared to 9.22% and 6.281% for ARIMA and ARIMAX with weather, respectively (Madhavi et al., 2017). Another study collected occupancy and electricity hourly used data from a three-story office building in eastern Canada. The data was used to train an ARIMAX model and the results suggest that building-level occupancy data can be used to improve the accuracy of ARIMA-based electricity use forecasts, since the MAE of the ARIMAX model was $2.889kW$, the MSE was $17.032kW^2$, and the MAPE was 1.217% (Iftikhar et al., 2023).

Using machine learning models, Kaytez et al. (2015) used three methods to predict annual electricity consumption of buildings in Turkey: Multiple Linear Regression (MLR), Artificial Neural Network (ANN) and Least Squares Support Vector Machines (LS-SVM), using as independent variables the installed capacity (IC), gross electricity generation (GEG), population (P), and total subscribership (TS). Looking at the accuracy metrics of the models, the LS-SVM method resulted in lower test MAPE of 1.004%, MSE of $2.06TWh^2$ and a RMSE of $1.435TWh$, comparing to the other two models (Kaytez et al., 2015).

Several other studies use tree based methods. The decision tree method is an effective method for modeling building energy demand, since it is able to capture the nonlinear and stochastic dynamics of building energy demand. This method was trained with energy demand data from residential buildings in China and, using the entropy calculated at each node of the tree, a selection of attributes was made. The model used data on indoor environment parameters, household characteristics and other issues such as occupant behavior and energy saving measures. The results showed that the variable with the greatest impact was temperature, followed by others related to the variation of this same parameter, such as the heating method and heat loss, and was able to predict building energy demand with an accuracy of up to 92% (Yu et al., 2010).

A study carried out at a university in Korea used various methods to predict electricity consumption in buildings, dividing them into three clusters according to their function (academic, scientific research and residential) and concluded that a hybrid Random Forest (RF) model with Multilayer perceptron (MLP) model is very effective at prediction. The study used variables with temporal, climatological and operational information (week, holiday, academic year, temperature, week-ahead load, year-ahead load, and Long Short-Term Memory network). In addition, a time series cross validation technique was used, which makes it possible to focus on a single prediction horizon for each test set. With this analysis, the RF model obtained a MAPE of 5.641% and a RMSE of $4675.762kWh$. RF+MLP model was the one that obtained the best metrics, with a MAPE of

4.674% and a RMSE of $3894.495kWh$, in the second cluster (cluster with the highest consumption) (Moon et al., 2018).

In a comparative analysis between time series and machine learning approaches, a study focused on load data from Covenant University. Various forecasting methods were applied, including the different points Moving Average, exponential smoothing (a time series method), and Artificial Neural Network (ANN) models. Among the different options explored, the model using a 3-Point Moving Average achieved a MAPE of 10.3%. The model incorporating exponential smoothing with a desmoothing constant of 0.3 demonstrated a MAPE of 8.56%, while the Artificial Neural Network produced the lowest MAPE at 8.25% (Samuel et al., 2016).

2.2 Interpretability and Explainability

Doshi-Velez and Kim (2017) define interpretability, in the context of machine learning, as the ability to explain a model output to a human in understandable terms. However, not all machine learning systems need to be interpretable, as there may be no consequences for invalid results or the problem in question may be so common and used in real situations that the results provided by the system become valid. Therefore, it must be understood when interpretability is necessary. In addition, the level of interpretability needed must be realized. Global interpretability requires knowing the existing patterns in general, while local interpretability requires knowing the reasons for each specific decision (Doshi-Velez & Kim, 2017).

A data analytics-based process for benchmarking energy performance of buildings can be made more interpretable by using explainable artificial intelligence (XAI) techniques. XAI techniques can reveal the contributions of each input variable to the benchmarking results, which can help users to understand the factors that influence the energy performance of buildings. In a recent study by Galli et al. (2022), an explainable AI-based benchmarking framework was developed to estimate the membership to specific energy performance classes of a large set of Energy Performance Certificates (EPCs) of flats in Italy. The framework used a Shapley additive explanations (SHAP) method to explain the model results and causal effects between the predictors and target variable to better understand the model behavior, and the motivations behind correct and wrong performed classifications (Galli et al., 2022).

In recent study by Gao and Ruan (2021), to predict and interpret the energy consumption of a building, was showed that the attention mechanism can be used to improve the interpretability of deep learning models for building energy consumption prediction. The attention weights can be used to identify the features that are most important for predicting building energy consumption. For example, the attention weights for the model trained on the commercial office building data showed that the most important features were the outdoor temperature, the indoor temperature, and the occupancy level. The attention weights can also be used to explain why the model made a particular prediction. For example, if the model predicted that the building's energy consumption would be high on a particular day, the attention weights could be used to show that the model was influenced by the high outdoor temperature and the high occupancy level (Gao & Ruan, 2021).

Studies of Eletric consumption in the metallurgical context have also been using machine learning approaches. Carlsson et al. (2020) have analysed eletrical energy consumption of an Eletric Arc Furnace (EAF) using data from a commercial steel plant. The model was then interpreted using SHAP to reveal the contributions from each input variable on the electrical energy consumption for every single heat in the prediction domain. The SHAP values for each input variable were evaluated based on process metallurgical experience to identify the most important factors influencing the electrical energy consumption of an EAF. This can help process engineers to better understand the factors influencing these processes and to identify ways to improve them (Carlsson et al., 2020).

Interpretable machine learning (IML) models have the potential to significantly improve building energy management by identifying the factors that influence building energy consumption and developing strategies for reducing building energy consumption. However, there are some challenges that need to be addressed for the widespread adoption of IML models in building energy management, such as lack of standardization, the need for high-quality data, and the difficulty of explaining IML model predictions to non expert users. IML models can achieve mean absolute errors (MAEs) of up to 5% for forecasting residential building energy consumption, and MAEs of up to 10% for forecasting commercial building energy consumption (Z. Chen et al., 2023).

Another technique for explaining outputs is LIME, which is a model-agnostic post-hoc method for approximating any ML model locally. In a study for Solar Photovoltaic Power Generation Forecasting, the authors employed LIME to explain the predictions of a deep learning model trained on historical photovoltaic and weather data. By highlighting the most influential features for each forecast, LIME provides valuable information about the model's decision-making process, allowing users to identify the main factors affecting photovoltaic generation. Here, surface solar radiation, time of day and upper net solar radiation are the most important features, while the total net water column, the 10-meter U wind component and total precipitation are the lowest in terms of numerical contribution (Kuzlu et al., 2020).

3 Methodolody

3.1 Data Collection and description

The analysis in this thesis was conducted within the context of Faro Airport from ANA – Aeroportos (VINCI Group), which experiences, in 2022, 8 millions passengers over the year, reaching its peak during the summer months due to its location in a region with high demand for tourism and leisure at that time. This company, which is responsible for managing airport movements throughout the region, incurs huge costs at all levels, particularly in terms of electricity consumption.

For the analysis and implementation of this study, Faro Airport provided historical data on daily electricity consumption in kWh, covering the period from January 1st, 2013, to September 30th, 2023.

The dataset also contained information on airport operations during the same period, including the number of flights, passenger counts, airport opening and closing times, and the respective operational duration (activity period).

Another set of crucial factors from the outset were related to climatic conditions. Faro Airport experiences a significant influx of tourist movements, with a higher percentage attributed to leisure travel rather than business purposes. It serves as a gateway to the South of Portugal and Southern Spain, popular vacation destinations throughout the year. These climatic data were downloaded from the website Visual Crossing (Crossing, 2023).

Thus, there were four datasets: weather data, period data, traffic data, and consumption data.

Table A1 in the appendix A contains a data dictionary with the original variables used.

3.2 Data preprocessing

3.2.1 Data Cleaning

Were identified some missing dates for operational variables. After analyzing the reason for these missing points, with Faro Airport, it was found that they coincided with periods of lockdown or airport restrictions due to the global Covid-19 pandemic. Therefore, as there was no arrival or departure movement on those days, and the airport's operating hours were from $06AM$ to $00AM$, these values were replaced as follows:

- 0 for the variables $Movements$ and $Passengers$, since there is no flights;
- 6 for the variable $Opening$, since it is the opening hour ($6AM$);
- 0 for the variable $Closing$, since it is the closing hour ($0AM$);
- 18 for the variable $Duration$, since the working period is from $6AM$ to $00AM$ (18 hours).

Then, a missing value was detected in the variable $Sea\_level\_pressure$ on February 12th, 2021. Since there was no apparent reason for this gap, the value was replaced by the mean between the previous and subsequent days. The analysis of the graph below indicates that this was the most appropriate measure given the evolution of the values.

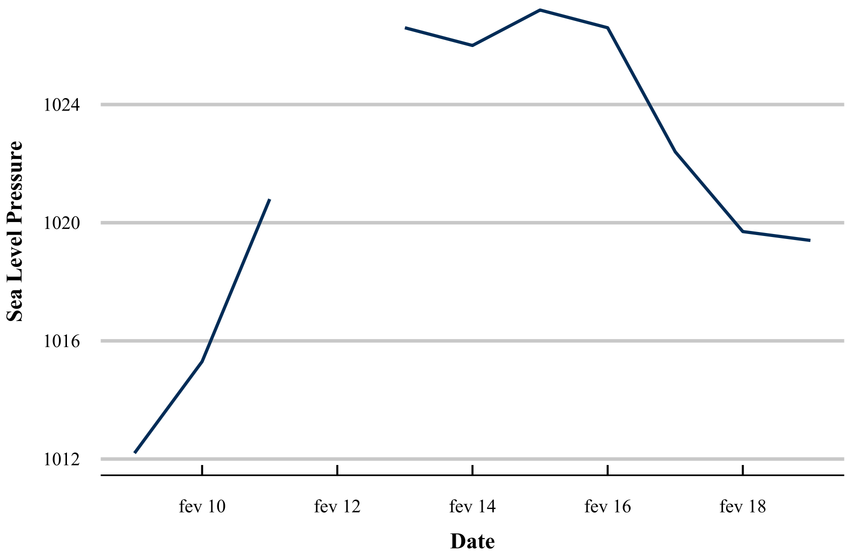


Figure 1: Sea Level Pressure Variation from 9th to 17th February

Since this analysis focuses on Faro Airport, it was observed that some climatic variables did not make sense due to the location or the specific study. Thus, these variables were eliminated. Table A2 in the appendix A contains a data dictionary with the deleted variables.

Due to reasons unrelated to this study, it was not possible to obtain data for the variables $Solar\_radiation$, $Solar\_energy$, and Uv_index in the period $[07-01-2023; 09-30-2023]$, so only the period $[01-01-2013; 06-30-2023]$ was considered for the study.

3.2.2 Feature engineering

With the complete data, the presence of outliers was assessed. The outliers do not seem to have an apparent reason for their existence. To avoid causing deviations in the moving averages (possible features to create) and, considering that the distribution is very robust on a weekly basis, the value from the previous week was used.

Finally, with a clean, consistent, and representative dataset, proceed to feature creation. Several ideas for new variables that could introduce power into the modeling emerged.

Since it is a time series, when using machine learning methods, it was necessary to create variables that captured temporal information: day, day of the week, week number, month, quarter, and year. However, as these variables are numeric, the model could capture incorrect information from this data, such as giving more importance to Friday (6) than Monday (2) just because its numeric value is higher. Thus, cyclic encoding was used. Cyclic encoding is a type of data encoding that uses a cycle to represent them, dividing them into blocks and then encoding each block using a standard encoding algorithm. The encoded blocks are then connected in a cycle to form the encoded data. In this case, trigonometric periodic functions were used to capture this effect, i.e., a sine and a cosine variable for each of these temporal components.

Additionally, the following variables were created:

- $Hours\_sun$ - Computes the daily number of hours, i.e., $Sunset$ - $Sunrise$;
- $Temp\_range$ - Calculates the daily temperature range, i.e., $Temp\_max$ - $Temp\_min$;
- $SeasonSummer$, $SeasonSpring$, and $SeasonWinter$ - Three dummy variables to capture the season of the year;
- $Holiday$ - A dummy variable to identify whether it is a holiday (with the assistance of a previously created table with the main national holidays);
- $Hours\_without\_sun$ - Records the number of hours the airport was in operation but still without sunlight (indicating a greater need for artificial light);

- $Grow\_Rate\_Percent$ - Calculates the quarterly growth rate in percentage.

Time series are characterized by having a temporal order, where each observation is influenced by previous observations. For this reason, it is crucial to consider temporal dependence when making predictions.

To improve the analysis of the time series, the rolling means technique was applied to the data. This approach involves creating moving averages over time, replacing individual values with averages calculated over consecutive windows. This smoothing of the data provides a clearer view of the underlying trends, filtering out temporary fluctuations and highlighting long-term patterns in order to reduce the impact of noise or random variations, making it easier to identify more consistent patterns in the time series and, consequently, improving the accuracy of the associated analyses and forecasts.

For the application of rollmeans, it is necessary to choose its size. In this study, weekly and monthly rollmeans were used for most explanatory variables that seemed most suitable in the model: $Temp\_max$, $Temp\_min$, $Temp$, Dew, $Humidity$, $Precip$, $Wind\_speed$, $Sea\_level\_pressure$, $Cloud\_cover$, $Solar\_radiation$, Uv_index, $Duration$, $Passengers$, $Movements$, $Hours\_sun$, and $Temp\_range$. After this technique, it is necessary to eliminate rows where missing values were created due to the absence of historical data in the creation of rolling means. Thus, the empty rows were removed.

Table A3 in the appendix A contains a data dictionary with the new variables.

3.3 Modelling Approaches

Having defined the objective of this study, it was necessary to formulate a logical framework to guide the research. To predict the daily electricity consumption at Faro Airport based on various factors, a quantitative method was deemed most appropriate. Through the analysis of historical data, characteristic patterns that could be extrapolated into the future could be identified. In this research, Machine Learning and Time Series manipulation methods and models were used to forecast daily electricity consumption at Faro Airport.

The data was split into two sets, one for training and one for testing, preserving the temporal dependence between observations. The training set was used to train the model, while the testing set was used to evaluate the model's performance, using an 80% training and 20% testing proportion, namely:

- Training set: $01 - 01 - 2013$ to $05 - 31 - 2021$
- Testing set: $06 - 01 - 2021$ to $06 - 30 - 2023$

Figure 2 shows the splitting into training set and testing set.

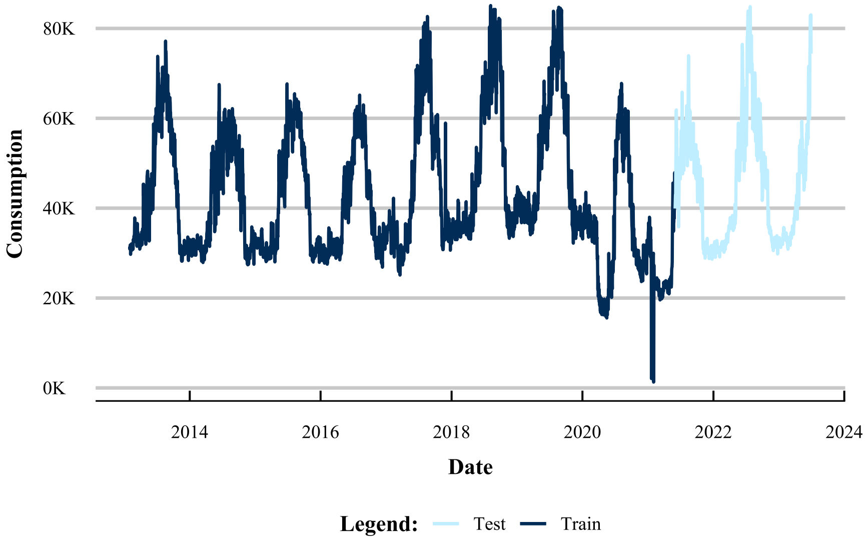


Figure 2: Training (dark blue line) and Testing (light blue line) Split

After dividing the data into training and testing, it is essential to estimate the model's accuracy on unseen data. A well-known technique for assessing the generalization ability of a model from a dataset and for avoiding overfitting is cross validation. Traditionally, this technique divides the dataset into various subsets (folds) and then trains the model on each fold, excluding the test fold. Finally, the model's performance is evaluated on the test fold. In the case of time series, this method is not suitable since random division can break the temporal dependence between observations. Therefore, time series cross validation can be used. This technique preserves the temporal dependence between observations, by carrying out sequential splits and ensuring that training takes place with data prior to the test data, effectively simulating the forecasting scenario in the real world. This method, executed in multiple splits, allows the model to be evaluated iteratively, adjusting its parameters only with information available up to the cut-off point. The initial window size was set to 30% of the training data, and the test window size was 30% of the testing data. The time window was fixed, meaning the same data window is used for training and testing in each iteration of cross-validation.

The approach described in this work focused on combining ML models (Random Forest, RF) with time series manipulation technique (AutoRegressive Integrated Moving Average with eXplanatory variables, ARIMAX).

3.3.1 Random Forest

This Supervised Machine Learning model uses a set of decision trees to predict the outcome of the explained variable. Decision trees are models that use a tree structure to represent the relationship between explanatory variables and the explained variable and can represent complex relationships between variables, making them suitable for a wide range of applications.

To improve this model, an adjustment of hyperparameters was carried out, which involves choosing the ideal values for these parameters in order to optimize the model's performance. In this study, various combinations of hyperparameters were tried in order to find the configuration that best fits the data. The minimum number of observations needed to create a node, since the training dataset is made up of a considerable number of observations, started out low, to allow the trees to grow deeper. The maximum depth of each tree was also adjusted, to try to learn more complex relationships, but always paying attention to the number placed, since if it is too high, the risk of overfitting increases. Other parameters adjusted were the minimum number of observations required for a leaf and the minimum number of observations needed to consider splitting an internal node.

Feature importance was calculated using the decision tree importance technique. Decision tree importance measures the contribution of each variable to reducing impurity in the decision tree. Analyzing these parameters is useful for improving the model or interpreting the model's results.

Final model evaluation was performed on test data.

3.3.2 Time Series based Models

The second technique was the study of time series for the application of an ARIMAX model.

First, a time series object was created, containing only the daily electricity consumption information.

ARIMA models consist of three components:

- Trend component: This component represents the long-term trend of the time series.
- Seasonality component: This component represents the seasonal variations of the time series.
- Noise component: This component represents the random variability of the time series.

ARIMAX models can estimate the parameters of these components, which can then be used to predict future values of the time series. Furthermore, it attempts to capture additional patterns and trends that can contribute to more accurate and robust modeling, taking into account the joint influence of the selected variables. Thus, the structure of an ARIMA model is represented by

$$y_t^{(d)} = \mu + \sum_{i=1}^{p} \phi_i y_{t-i}^{(d)} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \sum_{m=1}^{M} \beta_m X_{m,t} + \epsilon_t \tag{1}$$

where p is the number of autoregressive terms, d is the number of differences required to make the time series stationary, q is the number of moving average terms, and M is the number of variables to use in the model.

The exogenous variables were selected based on the feature importance obtained from applying the Random Forest (RF) model. The top 20 variables identified as most relevant by the RF were chosen as inputs for the ARIMAX model, except for any dummy variables (seasons and holidays). Recursively, the model was applied considering expanding sets of variables, starting with the top 20's first variable and progressively adding more variables. The choice of the optimal number of variables was determined based on performance measured by the Mean Absolute Error (MAE), and the final model was applied using this optimized set of variables.

The model parameters were automatically selected using the $auto.arima$ function after applying a Box Cox transformation with automatically calculated lambda and highlighting the seasonal component. To obtain accurate forecast values, the Box Cox transformation was reversed using the lambda calculated by the model.

3.4 Ensemble Models

Ensemble models were employed as part of modeling approach, specifically using the ensemble technique of averaging predictions generated by different models. The concept behind ensemble models is to combine predictions from multiple individual models to enhance the accuracy and robustness of forecasts. By averaging predictions from diverse models, the aim was to capture a more comprehensive and balanced view of the system's behavior while minimizing potential biases inherent in specific models. This ensemble approach is valuable as it integrates the collective wisdom of multiple models, mitigating individual limitations and yielding more reliable and generalizable results.

3.5 Evaluation Metrics

When evaluating the predictive performance of regression or forecasting, various accuracy measures are commonly employed. Although each measure offers unique insights, relying on a single metric can be insufficient. Therefore, a comprehensive evaluation should include multiple measures. In this study, the Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE) were used.

Mean Absolute Error (MAE): MAE represents the average of the absolute differences between predicted values and actual values. A lower MAE indicates a closer approximation to the actual data.

$$MAE = \frac{1}{N} \sum |y_i - \hat{y}_i|, \tag{2}$$

where N is the number of predictions, y_i is the actual value, and $\hat{y}_i$ is the predicted value.

Mean Squared Error (MSE): MSE calculates the average of the squared differences between predicted and actual values. A lower MSE implies a better fit.

$$MSE = \frac{1}{N} \sum (y_i - \hat{y}_i)^2, \tag{3}$$

where N is the number of predictions, y_i is the actual value, and $\hat{y}_i$ is the predicted value.

Root Mean Squared Error (RMSE): RMSE calculates the square root of the MSE, allowing a better analysis of the scale of the data, since the MSE uses the square of the units. A lower RMSE implies a better fit.

$$MSE = \sqrt{\frac{1}{N} \sum (y_i - \hat{y}_i)^2}, \tag{4}$$

where N is the number of predictions, y_i is the actual value, and $\hat{y}_i$ is the predicted value.

Mean Absolute Percentage Error (MAPE): MAPE measures the mean absolute percentage error. It is particularly useful for comparing models when the scale of the data varies.

$$MAPE = \frac{1}{N} \sum \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100, \tag{5}$$

where N is the number of predictions, y_i is the actual value, and $\hat{y}_i$ is the predicted value.

The choice of the precision measure depends on the specific context and data characteristics. The MAE is less sensitive to outliers than the MSE and the MAPE is scale-independent, making it suitable for comparing models with different data scales.

3.6 Interpretability with SHAP analysis

The Random Forest model is considered a black box model in machine learning. This means that its results are incomprehensible to humans. To interpret the results of the study, SHAP was used to address this need.

SHAP (SHapley Additive exPlanations) is a theoretical approach based on game theory, specifically Shapley values, introduced by Lloyd Shapley in 1953. In the context of interpreting machine learning models, SHAP provides a systematic and theoretically sound way to assign individual contributions to the prediction of a model for each feature in a dataset. The main goal of this method is to explain the model's response in a way that is fair, consistent, and globally coherent. It is possible to perform a local analysis for each observation or a global analysis with an average of SHAP values for each point.

SHAP values for a specific feature represent its average contribution to the difference between the model's prediction and the mean of predictions. Considering all possible combinations of features, the average contribution of each feature in all these combinations is calculated.

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(N - |S| - 1)!}{N!} [v(S \cup \{i\}) - v(S)], \tag{6}$$

where ϕ_i is the Shapley value for observation i, S is a coalition of observations, $v(S)$ is the payoff for this coalition, N is the total number of observations and $N \setminus \{i\}$ is all the possible coalitions not containing i.

For this purpose, bee swarm, waterfall and dependence plots were used, which are more complex and informative than a simple variable importance plot. They provide not only the importance of variables but also reveal the true relationship between the explanatory variables and the target variable.

3.7 Softwares and Tools

For the execution of this study, data were imported from files in *.csv* and *.xlsx* formats from Microsoft Excel. Data manipulation, exploration, and modeling were conducted in R version 4.2.2 (2022-10-31 ucrt), using the open-source integrated development environment for R, RStudio 2022.12.0+353. This software serves as a functional interface that employs a programming language for statistical calculations and graphics. Various specific packages were utilized for data manipulation, visualization, and modeling.

For data manipulation, the packages included were *openxlsx*, *tsibble*, *tidyverse*, *lubridate*, *timeDate*, *utils*, and *dplyr*. In terms of modeling, packages such as *randomForest*, *forecast*, *caret*, and *kernelshap* were employed. Additionally, the *knitr* package, along with *kableExtra*, was used for generating enhanced tables in LaTeX. Finally, for data visualization, the *ggplot*2 package, along with *shapviz*, *formatR*, *ggthemes*, and *gridExtra*, were utilized for formatting and refining the graphics.

4 Results

4.1 Exploratory Data Analysis

Daily energy consumption over time is shown in Figure 3 (a), highlighting a clear seasonal pattern that appears to be annual. In Figure 3 (b), this annual seasonality is confirmed. From May to September, it can be seen that, for all the years under analysis, consumption is significantly higher. There are several hypotheses about this phenomenon, as these are months with higher temperatures, which implies a greater need for air conditioning, and also a high number of air movements and passengers. Additionally, in these months, air conditioning is done with electrical devices, while in the remaining months, it is done using gas boilers, thus reducing electricity consumption. Considering the evaluation of seasonality on another time basis, Figure 3 (c) shows that there is no clear pattern of consumption over the days of the month, so there is no monthly consumption pattern.

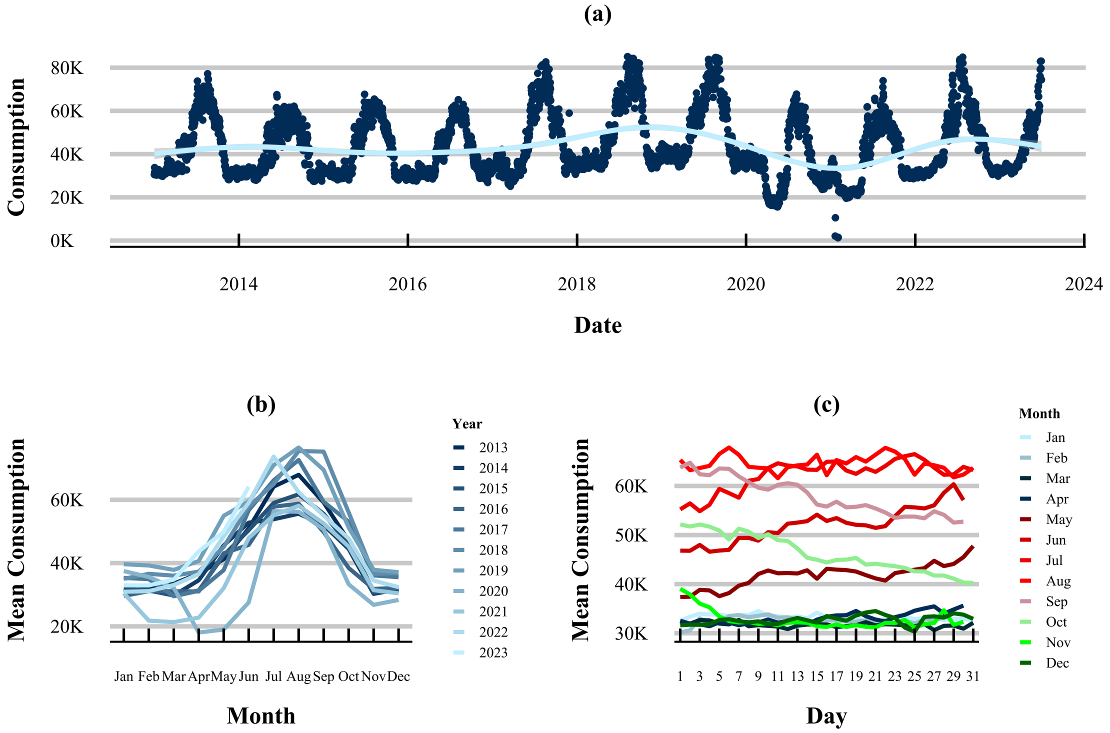


Figure 3: (a) Distribution of the time series (dark blue dots) over time accompanied by a trend line (light blue line); (b) Distribution of average monthly consumption for each year; (c) Distribution of average daily consumption for each month grouped for all years

Since most airlines apply lower fares on specific weekdays, one might expect a pattern in the distribution of consumption throughout the weekdays due to a higher number of passengers. In an attempt to assess this distribution, Figure 4 shows a heatmap with the average consumption on weekdays for each month. It is possible to highlight, once again, the annual seasonality. However, there is no clear pattern throughout the week, contrary to what was assumed.

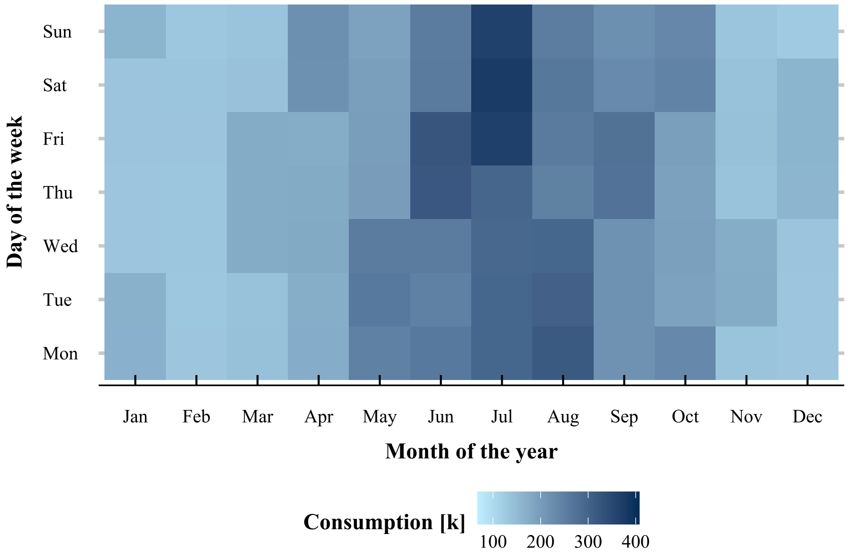


Figure 4: Consumption per Month and Day of the Week, where darker zones means more consumption

Another factor that could also affect airport activity and, consequently, electricity consumption is holidays. In Figure 5, it is shown that electricity consumption on holidays is, on average, mostly lower than on non-holiday days.

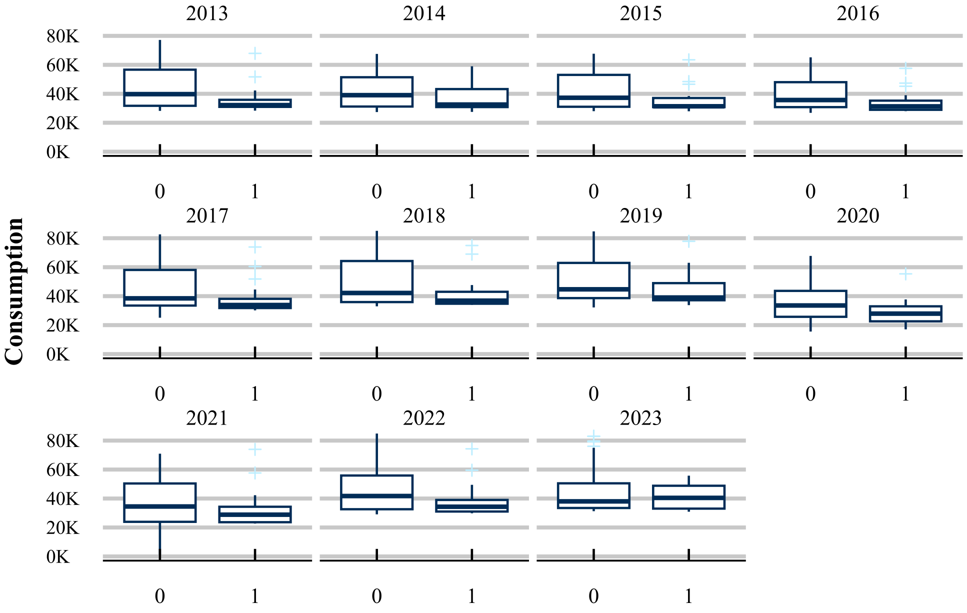


Figure 5: Distribution of consumption on holidays and non-holidays, where each panel represents a specific year, with distinct boxplots for holidays and non-holiday days, providing insights into the variability and centrality of consumption in different temporal contexts.

The distribution of consumption per level of UV index is shown in Figure 6, emphasizing that consumption is higher and more variable for higher UV index values.

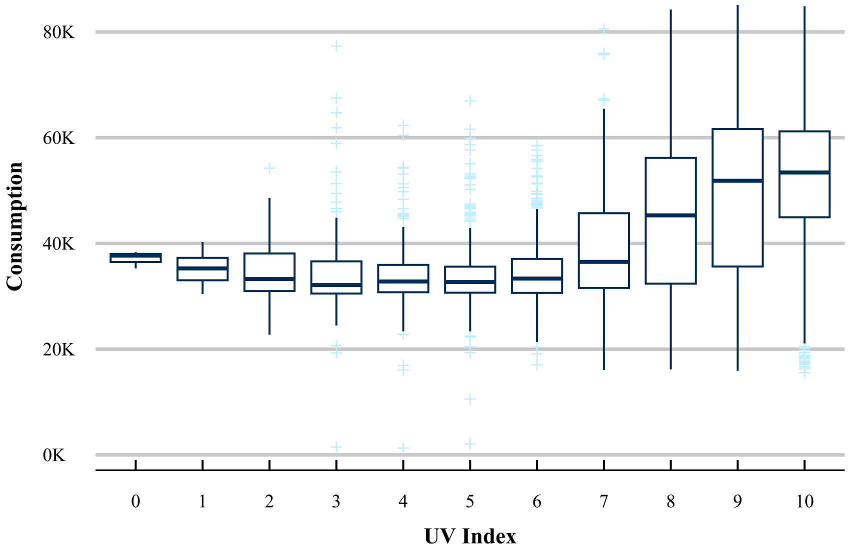


Figure 6: Boxplots relating energy consumption to different UV index levels.

The distribution of electricity consumption was also analyzed based on some numerical variables of the model, which can be seen in figure B1 of the appendix B. In general, it was observed that variables such as $Temp\_max$, $Temp$, $Temp\_min$, $Passengers$, and $Movements$ exhibit a positive relationship with electricity consumption. On the other hand, variables such as $Cloud\_cover$ and $Humidity$ negatively impact electricity consumption. There are also some variables where there seems to be no significant relationship, such as the $Temp\_range$ and $Visibility$ variables.

For time series manipulation, the correlation of the series with past periods was examined using an autocorrelation plot, shown in Figure 7, at a significance level of 95% (dark blue lines), once again confirming the annual seasonality in the data, since the positive peaks have a difference of 365 lags.

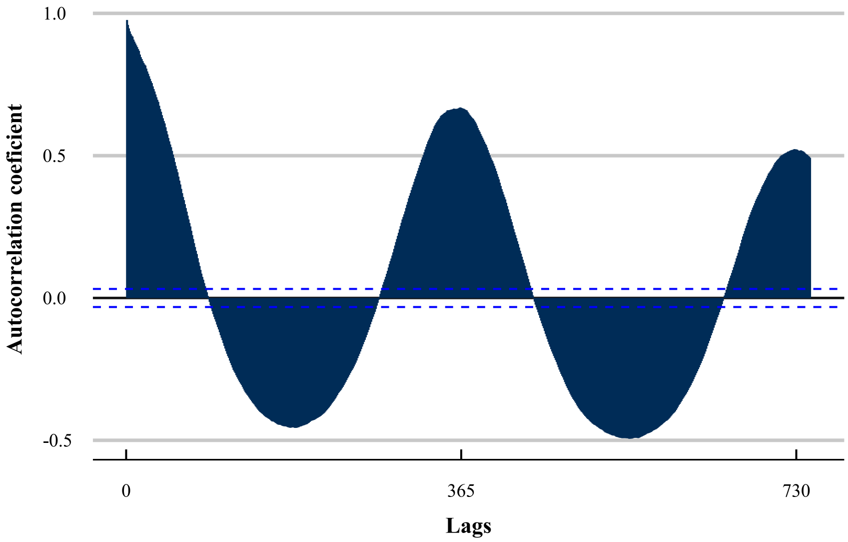


Figure 7: Autocorrelation Plot of Consumption Series

Additionally, the time series was decomposed into the three aforementioned components to understand the temporal evolution of the data, with an additive decomposition method. By isolating these elements, it becomes possible to identify long-term patterns, seasonal variations and irregularities in the data. This in-depth analysis makes it possible to extract valuable insights into underlying trends, recurring seasonal behaviors and temporal anomalies, providing a solid basis for informed decision-making and more accurate forecasting strategies. In this case, it helped in the choice of parameters for the construction of ARIMAX, in order to integrate a parameter that requires working with seasonality. Figure 8 shows the decomposition of the time series.

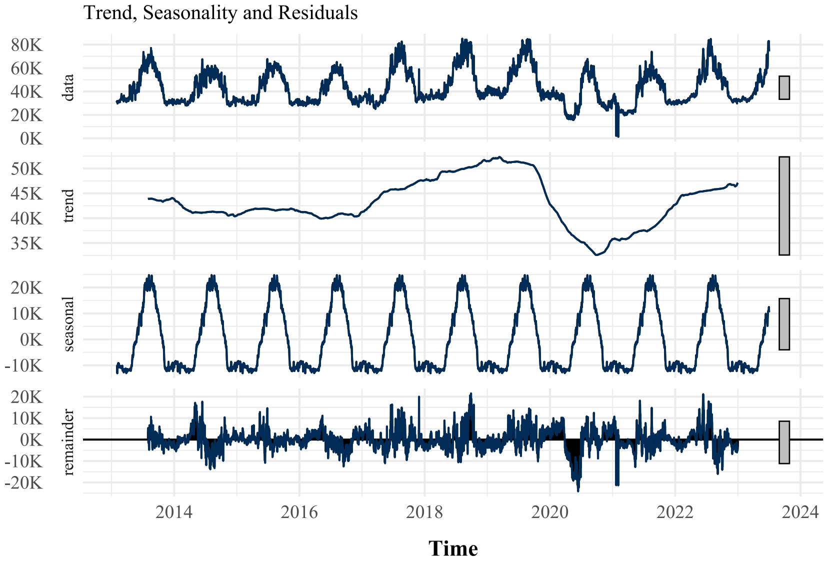


Figure 8: Additive Decomposition of Consumption time series

If the time spectrum was wider, it might be possible to see that the pattern visible in the trend was actually a cycle, another property of time series to consider when manipulating them.

4.2 Modeling findings

Training the Random forest model resulted in a model with a minimum of 10 observations needed to create a node, a maximum depth of 30 for each tree, a minimum of 15 observations required for a leaf, a minimum of 15 observations needed to consider splitting an internal node, 500 trees in each split and an optimum number of variables to use of 36, resulting in a model with explanability of 96.8% of the target variable.

Random Forest top 20 most important features is illustrated at Figure 9.

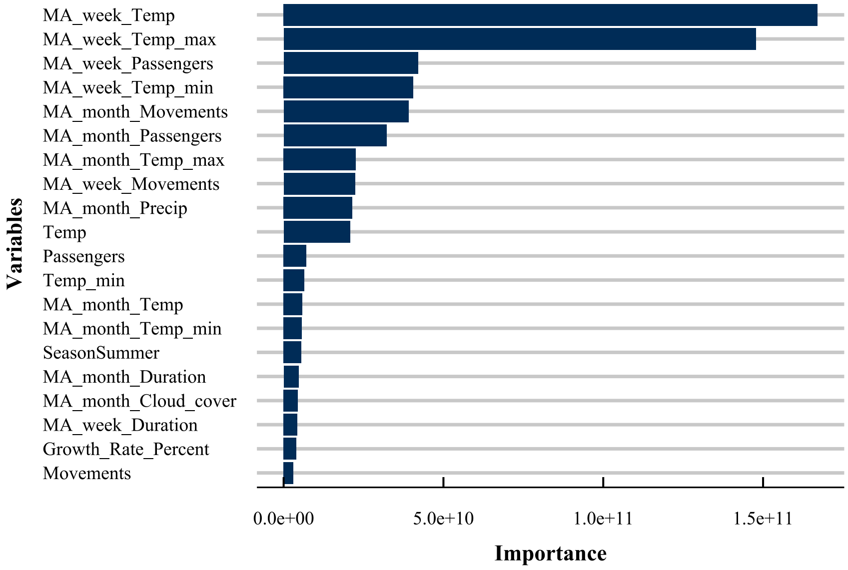


Figure 9: Random Forest Variable Importance

It can be observed that the moving averages of certain variables stand out the most, both on a weekly and monthly basis, indicating that temperature, movements, and passengers' variables hold higher importance values.

In Figure 10, a comparison between the actual data and predictions provided by the Random Forest model is shown.

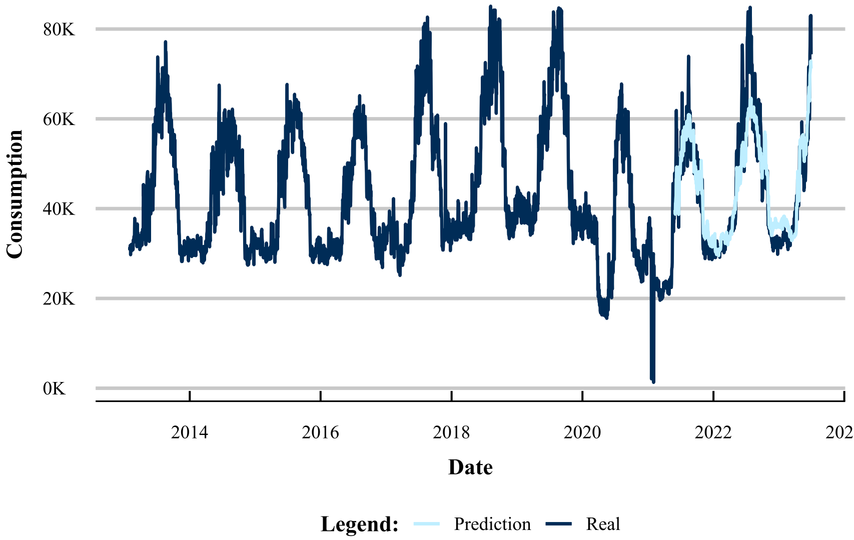


Figure 10: Real Data (dark blue line) vs. Random Forest Predictions (light blue line)

ARIMAX was trained and used to recursively predict on the previous set of top 20 variables. Figure 11 shows a graph of the model's performance as the number of variables from the top 20 used increases. The number of variables used corresponds to the minimum point on the graph.

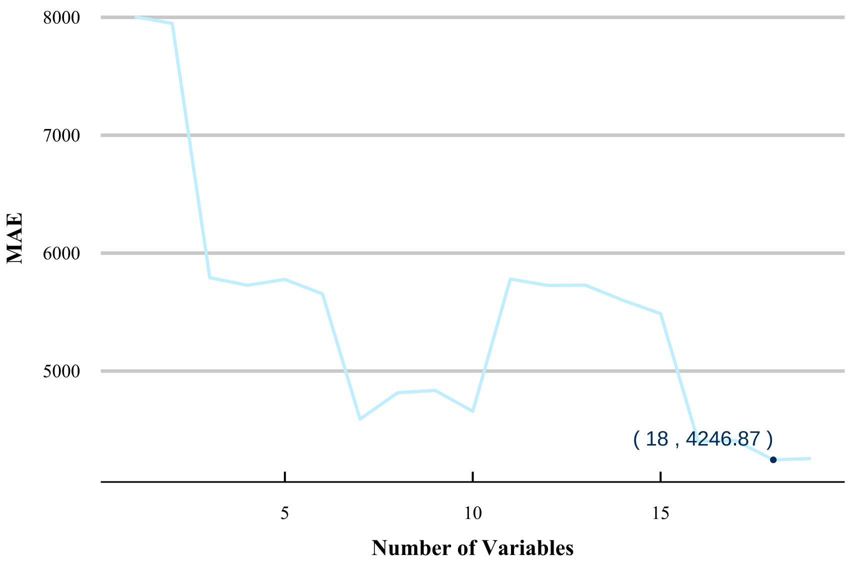


Figure 11: ARIMAX performance variation for different number of explanatory variables.

After applying the ARIMAX model, the following adjusted model was obtained: ARIMA(3,1,1) with Box Cox transformation lambda of 0.54, which means it uses 3 autoregressive lags, first-order differencing to make the series stationary, and 1 moving average lags.

In Figure 12, a comparison between the actual data and predictions provided by the ARIMAX model is presented.

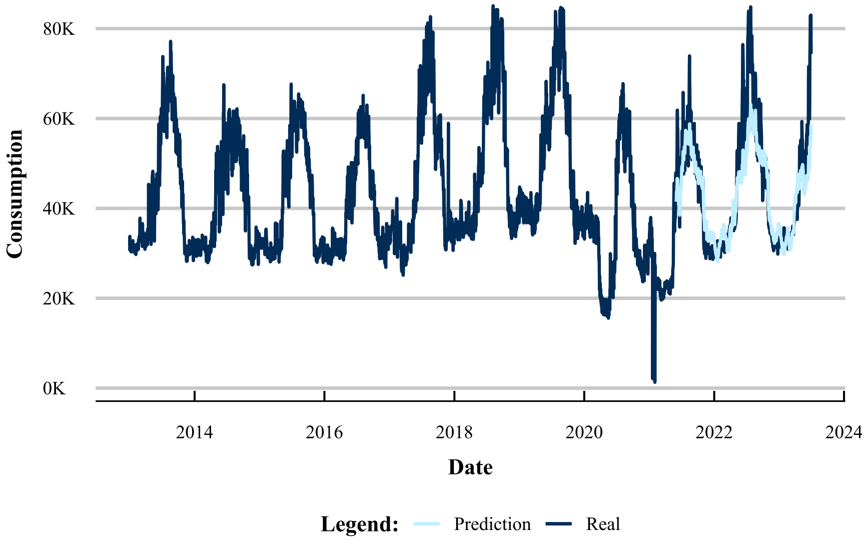


Figure 12: Real Data (dark blue line) vs. ARIMAX Predictions (light blue line)

To obtain the final forecasts, the two models were averaged, with a result on the test data, shown in figure 13.

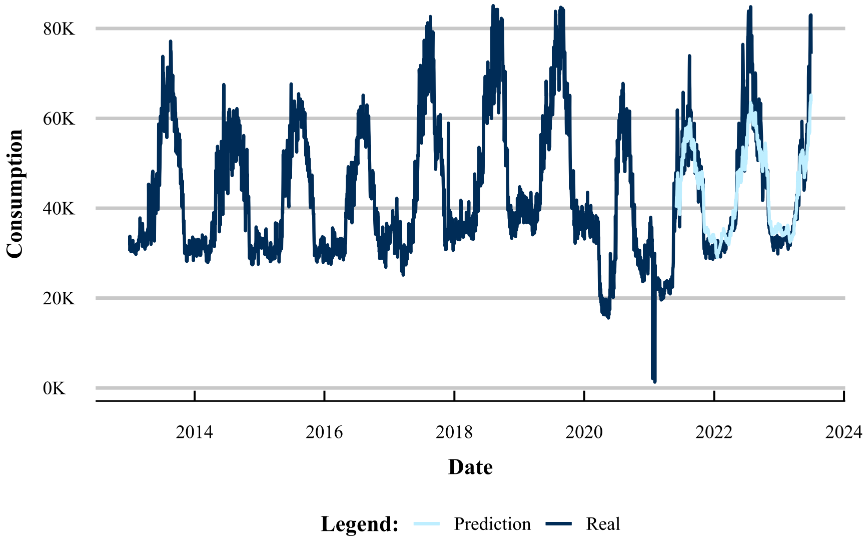


Figure 13: Real Data (dark blue line) vs. Final Predictions (light blue line)

4.3 Model Evaluation

To assess the performance of the models, as mentioned earlier, the following prediction error metrics were utilized: MAE, MSE, RMSE, and MAPE.

Table 1 displays the results of each of these metrics, calculated from the average using the predictions of the test data, for each model and their respective combination.

Table 1: Models Evaluation

Model	MAE	MSE	RMSE	MAPE
Random Forest	3743.217	27315008	5226.376	7.998852
ARIMA	4246.868	40724717	6381.592	8.499076
Combination	3593.716	29675541	5447.526	7.319280

4.4 SHAP Analysis

As mentioned in section 3.5, SHAP allows for the actual interpretation of the forecast of an explanatory variable, calculating the impact of each of these variables on the forecast made by Random Forest, considering the local importance of the variable and its changes as the values increase or decrease. The beeswarm graph, in Figure 14, orders the variables by their mean effect on the forecast and shows how the lowest and highest values of these variables impact on the forecast consumption value, in addition to ordering them according to their greatest importance, as was done previously in Figure 8 of section 4.2..

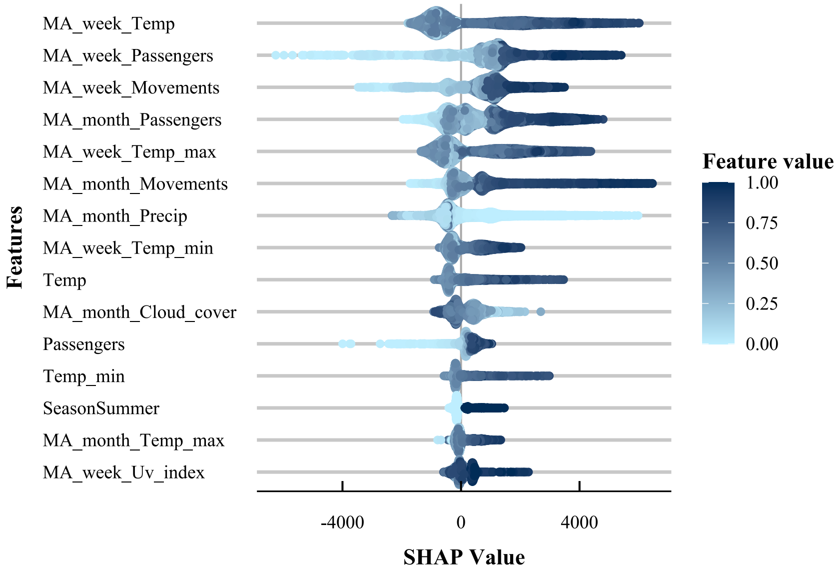


Figure 14: Shapley Values for some features, where dark blue represents higher feature values and light blue represents lower feature values.

Here it is possible to observe that, for example, higher values for the weekly moving averages of Movements or Passengers have positive SHAP values, while lower values have negative SHAP values. This implies that days with higher average Movements or Passengers have a higher predicted consumption. For lower weekly moving averages of cloud cover, positive SHAP values are associated, and vice-versa. The same analysis can be applied to the remaining variables. Additionally, the distribution of points is also informative. For instance, in the case of weekly moving averages of Passengers, a zone with a higher density of points with average values is located in an area of low SHAP values. On the other hand, low values for Passenger variable extend far to the left, suggesting that low values of Passengers have a stronger negative impact on daily consumption than the positive impact of high values of the same variable.

For a local interpretation, for example on 10-27-2015, a waterfall plot of SHAP values was reproduced. This graph can also be made for several observations and the average of the SHAP values is calculated.

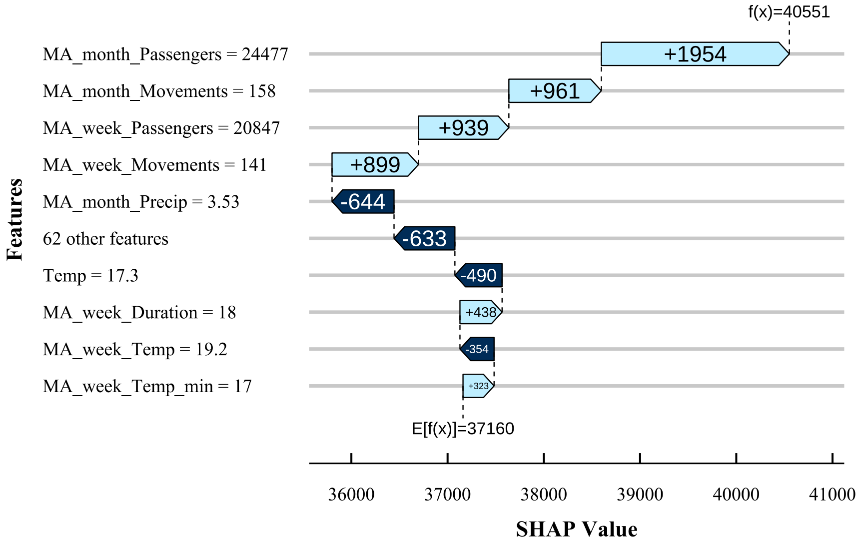


Figure 15: Shapley Values for some features, for the observation of 10-27-2015, where dark blue represents negative SHAP values and light blue represents positive SHAP values.

This graph shows that, on this day, the fact that the average rainfall over the last 30 days was $3.53mm$ reduced the electricity consumption forecast for that day by $644kWh$. Conversely, the fact that the average number of passengers over the last 30 days was 24477 increased the electricity consumption forecast for that day by $1954kWh$. This analysis can be done for any of the variables.

In addition to analyzing the effect of some features, it is also possible to examine the behavior of one variable in relation to another. Given that two of the variables that demonstrated greater importance in the study were $Temp\_max$ and $Movements$, a scatterplot was created, as shown in Figure 16, illustrating the impact of the interaction between these two features on the predicted value of electricity consumption.

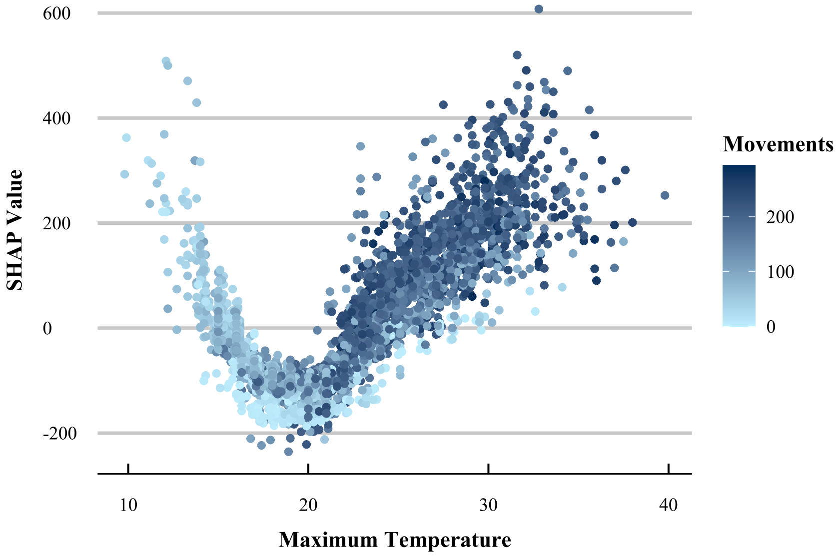


Figure 16: Shapley Values Maximum Temperature against Movements where darkblue represents higher values of Movements and light blue represents lower values of Movements.

It can be observed that, overall, high values of $Temp\_max$ paired with high values of $Movements$ exhibit higher SHAP values. In contrast, average/low values of $Temp\_max$ coupled with low values of $Movements$ result in negative SHAP values.

5 Discussion

The main objective of this study was to see whether the combination of two forecasting methods, combined with explainability and interpretability, would result in a good form of forecasting, with significant insights in this context. To assess the validity of the proposed research questions, a Random Forest model and an ARIMAX model were developed, as well as a combination of the two. In addition, a KernelSHAP explainer was developed to interpret the model. In this section, the results of the study will be discussed to verify whether there is evidence to support the research questions.

First, the Random Forest Machine Learning model was trained with different parameters and validated using the time series cross validation method, which resulted in the predictions seen in Figure 9 and the error measures in the first row of Table 1. Compared to Moon et al. (2018), who applied the RF model and then a combination of RF and MLP, the RMSE performance metric for RF model was slightly higher ($5226.376kWh$ vs. $4675.762kWh$), as well as MAPE (7.999% vs. 5.641%).

With regard to the second model, ARIMAX, the modeling carried out with the 18 most important variables in the Random Forest, which came up with the predictions seen in Figure 12 and the error measures in the second row of Table 1, resulted in a model capable of capturing the seasonal patterns and trend of the data, based on the automatically determined autoregressive lags, differentiations and moving average lags. This model, applied to the test data, achieved a reasonable performance (MAPE = 8.499% and MAE = $4246.868kWh$), compared to the MAPE of the ARIMAX with wheather and traffic of 3.3%, the MAPE of ARIMAX with weather, of 6.281%, by Madhavi et al. (2017), and the ARIMAX with building-level occupancy data, with MAE of $2.889kWh$ and MAPE of 1.217%, by Iftikhar et al. (2010).

The combination of these two models resulted in a model that outperformed each of the models individually, with a MAE of $3593.716kWh$ and a MAPE of 7.319%, thus confirming the study's first hypothesis, which asked whether the combination of Machine Learning and Time Series models collectively outperformed each of these models individually in terms of performance metrics.

From the explainability analyses provided by the Shapley Values visualizations, it is possible to draw concrete information on what each variable causes in each of the observations. This complexity in interpretability takes forecasts deeper, allowing managers, such as in Faro Airport, to get advance information not only on future consumption, but also on possible changes or improvements to be made to some of the characteristics that most affect electricity consumption at the airport, thus validating the second hypothesis, which questioned whether the insights derived from these models would be useful in making practical decisions or formulating strategies.

With regard to the last research question, which asked whether the insights derived from explainability methods highlight any instability or uncertainty in the models' predictions, it can be seen in the graphs produced from the Shapley Values that the residuals of the predictions can be explained in detail by each variable, in order to obtain the real value of consumption.

It should be noted that the data has suffered an imbalance in the usual seasonal trend and pattern, due to the global Covid-19 pandemic, which from March 2020, closed most establishments in Portugal and, obviously, canceled air travel for some time. Although the restrictions eased in 2021, there was still a big impact on air travel. Thus, these two atypical years clearly produced significant deviations in the model's forecasts, but they were kept in the dataset, since these low electricity consumptions are accompanied by variables that also accompanied this decrease, namely $Passengers$ and $Movements$, which at times remained 0.

6 Conclusion and Developments

Integrating advanced Machine Learning techniques with Time Series models has unequivocally proven to be an exceptionally efficacious strategy in substantially enhancing the predictive prowess of individual models. This, exemplified by the fusion of a sophisticated Random Forest model and an adept ARIMAX model, strategically leveraging the most salient variables identified by the Random Forest, not only culminated in forecasts of heightened accuracy but also fortified robustness. This strategic synergy becomes particularly conspicuous when grappling with the inherent volatility and intricacies characterizing time series data; the Random Forest adeptly captures intricate and nuanced patterns, whereas the ARIMAX adeptly navigates through temporal trends and seasonal variations.

Furthermore, the intrinsic elucidative capabilities of these models bear immense value in the realm of pragmatic decision-making and the formulation of strategic initiatives. The interpretability, facilitated by the Shapley Values, engenders a profound and comprehensive comprehension of the myriad factors that underpin forecasts, both at a global scale and within localized contexts. This unprecedented level of transparency stands as an imperative linchpin for transforming the intricate results generated by these models into a comprehensible and human-friendly format. This, in turn, expedites the seamless implementation of informed strategies within real-world scenarios.

However, as the study unfolded, a singular constraint emerged in the form of data oscillations observed during the anomalous COVID-19 period. Despite the potential remedial measure of excluding the 2020 data, it was decided to retain it owing to the influence of accompanying variables. Remarkably, removing this data did not result in a substantial shift in model performance. Additionally, the imputation of moving averages in certain variables encountered notable challenges, underscoring the intricacies involved in data manipulation.

The envisaged augmentation of the modeling framework involves the judicious inclusion of novel datasets, such as detailed information pertaining to the types of aircraft operational at the airport and comprehensive insights into their maintenance and ground operations. This thoughtful expansion aims to provide a more panoramic view of energy consumption patterns, as diverse aircraft types inherently exhibit disparate energy requirements during their operational lifecycles. Moreover, the contributory facets of maintenance activities and airfield operations offer invaluable nuances to the overarching consumption dynamics. By integrating these nuanced variables into the predictive models, an unprecedented level of granularity and context-specific insights can be harnessed, thereby leading to markedly more accurate and refined forecasts. This analytical refinement is grounded in an exhaustive comprehension of airport operations, fortified by the consideration of factors uniquely germane to the aviation industry, which, collectively, enriches the model's predictive acumen and furnishes a more holistic and accurate depiction of energy consumption within the airport milieu.

Another enhancement that could complement the study, resulting in improved performance, would be to include more historical data. In other words, extending the temporal scope could allow the visualization of more evident seasonal patterns and cycles.

A noticeable detail in Figure 13 is the model's inability to closely approximate the high consumption levels during the summer months. A suggested approach for future investigation would be to train a model exclusively on these summer months to accurately capture this specific information and generate more precise forecasts. A less sophisticated approach could involve incorporating a percentage of the previously made predictions for these months.

As the trajectory unfolds towards future projects, a thoughtful consideration revolves around the prospective imputation of values for explanatory variables. Given their inherent unknown nature for future periods - excluding the exceptions of passenger and movement figures, reliably projected nearly a year in advance - contemplating the calculation of extended 30-year averages for these unknown variables emerges as a sagacious approach. This meticulous approach ensures a robust foundation for future forecasting endeavors, duly factoring in the uncertainties inherent in variables yet to be known.

References

ANA. (2022). Environmental performance report. https://www.ana.pt/sites/default/files/2023-09/ANA_RS_ING03_25.09.2023.pdf

Bescond, P.-L. (2021). *Cyclical features encoding, it's about time!* [Medium]. https://towardsdatascience.com/cyclical-features-encoding-its-about-time-ce23581845ca

Carlsson, L. S., Samuelsson, P. B., & Jönsson, P. G. (2020). Interpretable machine learning—tools to interpret the predictions of a machine learning model predicting the electrical energy consumption of an electric arc furnace. *steel research international*, *91*(11), 2000053. https://doi.org/10.1002/srin.202000053

Chen, J., & Xie, K. (2013). A prediction model based on unbiased grey markov for airport energy consumption prediction. *2013 Chinese Automation Congress*, 291–294. https://doi.org/10.1109/CAC.2013.6775745

Chen, Z., Xiao, F., Guo, F., & Yan, J. (2023). Interpretable machine learning for building energy management: A state-of-the-art review. *Advances in Applied Energy*, *9*, 100123. https://doi.org/10.1016/j.adapen.2023.100123

Costa, A., Blanes Restoy, L., Donnelly, C., & Keane, M. (2012). Review of eu airport energy interests and priorities with repect to ict, energy efficiency and enhanced building operation.

Crossing, V. (2023). Weather data services - visual crossing. https://www.visualcrossing.com/weather/weather-data-services

De Oliveira, E. M., & Cyrino Oliveira, F. L. (2018). Forecasting mid-long term electric energy consumption through bagging ARIMA and exponential smoothing methods. *Energy*, *144*, 776–788. https://doi.org/10.1016/j.energy.2017.12.049

Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. https://doi.org/10.48550/arXiv.1702.08608

Galli, A., Piscitelli, M. S., Moscato, V., & Capozzoli, A. (2022). Bridging the gap between complexity and interpretability of a data analytics-based process for benchmarking energy performance of buildings. *Expert Systems with Applications*, *206*, 117649. https://doi.org/10.1016/j.eswa.2022.117649

Gao, Y., & Ruan, Y. (2021). Interpretable deep learning model for building energy consumption prediction based on attention mechanism. *Energy and Buildings*, *252*, 111379. https://doi.org/10.1016/j.enbuild.2021.111379

Hong, T., & Fan, S. (2016). Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting*, *32*(3), 914–938. https://doi.org/10.1016/j.ijforecast.2015.11.011

Iftikhar, H., Bibi, N., Canas Rodrigues, P., & López-Gonzales, J. L. (2023). Multiple novel decomposition techniques for time series forecasting: Application to monthly forecasting of electricity consumption in pakistan [Number: 6 Publisher: Multidisciplinary Digital Publishing Institute]. *Energies*, *16*(6), 2579. https://doi.org/10.3390/en16062579

Kaytez, F., Taplamacioglu, M. C., Cam, E., & Hardalac, F. (2015). Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines. *International Journal of Electrical Power & Energy Systems*, *67*, 431–438. https://doi.org/10.1016/j.ijepes.2014.12.036

Klyuev, R. V., Morgoev, I. D., Morgoeva, A. D., Gavrina, O. A., Martyushev, N. V., Efremenkov, E. A., & Mengxu, Q. (2022). Methods of forecasting electric energy consumption: A literature review [Number: 23 Publisher: Multidisciplinary Digital Publishing Institute]. *Energies*, *15*(23), 8919. https://doi.org/10.3390/en15238919

Kuzlu, M., Cali, U., Sharma, V., & Guler, O. (2020). Gaining insight into solar photovoltaic power generation forecasting utilizing explainable artificial intelligence tools. *IEEE Access*, *8*, 187814–187823. https://doi.org/10.1109/ACCESS.2020.3031477

Madhavi, K. S. L., Cordova, J., Ulak, M. B., Ohlsen, M., Ozguven, E. E., Arghandeh, R., & Kocatepe, A. (2017). Advanced electricity load forecasting combining electricity and transportation network. *2017 North American Power Symposium (NAPS)*, 1–6. https://doi.org/10.1109/NAPS.2017.8107312

Moon, J., Kim, Y., Son, M., & Hwang, E. (2018). Hybrid short-term load forecasting scheme using random forest and multilayer perceptron [Number: 12 Publisher: Multidisciplinary Digital Publishing Institute]. *Energies*, *11*(12), 3283. https://doi.org/10.3390/en11123283

Ortega Alba, S., & Manana, M. (2016). Energy research in airports: A review [Number: 5 Publisher: Multidisciplinary Digital Publishing Institute]. *Energies*, *9*(5), 349. https://doi.org/10.3390/en9050349

Ortega Alba, S., & Manana, M. (2017). Characterization and analysis of energy demand patterns in airports [Number: 1 Publisher: Multidisciplinary Digital Publishing Institute]. *Energies*, *10*(1), 119. https://doi.org/10.3390/en10010119

Samuel, I., Ojewola, T., Awelewa, A., & Amaize, P. (2016). Short-term load forecasting using the time series and artificial neural network methods.

Xiaowei, Z. (2008). Application of data mining technology in shanghai commercial building information database.

Yang, Y., Chen, C., Li, Z., Guan, D., & Su, F. (2023). Electricity consumption characteristics and prediction model of a large airport terminal in beijing. *City and Built Environment*, *1*(1), 11. https://doi.org/10.1007/s44213-023-00012-1

Yu, Z., Haghighat, F., Fung, B. C., & Yoshino, H. (2010). A decision tree method for building energy demand modeling. *Energy and Buildings*, *42*(10), 1637–1646. https://doi.org/10.1016/j.enbuild.2010.04.006

Appendix

A Data Dictionary

Table A1: Original Variables - Data Dictionary

Variables	Description
Date	Date, yyyy/mm/dd
Consumption	Electricity consumption, Kwh
Temp_max	Maximum Temperature, ºC
Temp_min	Minimum Temperature, ºC
Temp	Mean Temperature, ºC
Dew	Dew Point, ºC
Humidity	Relative Humidity, %
Precip	Precipitation, mm
Precip_prob	Precipitayion Probability, %
Wind_speed	Wind Speed, Kph
Wind_dir	Wind Direction, º
Sea_level_pressure	Sea Level Pressure, mb
Cloud_cover	Cloud Cover, %
Visibility	Visibility, Km
Solar_radiation	Solar Radiation, W/m^2
Solar_energy	Solar Energy, MJ^2
Uv_index	UV Index
Sunrise	Sunrise time, h
Sunset	Sunset time, h
Opening	Opening time, h
Closing	Closing time, h
Duration	Duration, h
Passengers	Passengers
Movements	Movements

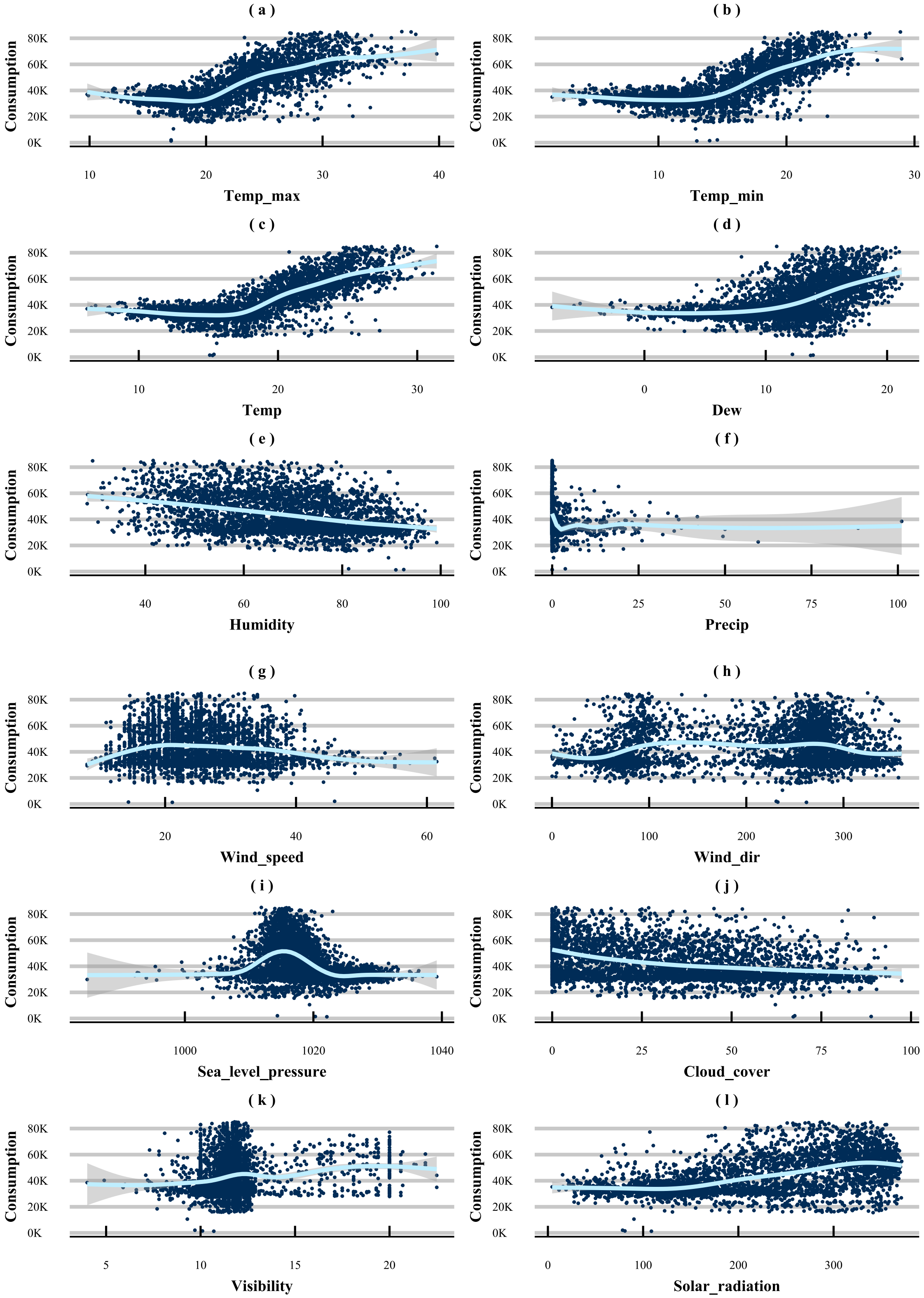
Table A2: Deleted Variables - Data Dictionary

Variables	Description
Name	City Name
Precip_type	Precipitation Type
Severe_risk	Severe Risk
Conditions	Text about the weather
Description	Description of the weather
Icon	Weather Icon
Stations	Stations source
Wind_gust	Wind Gust
Feelslike	Feels Like, ºC
Feelslike_max	Maximum Feels Like, ºC
Feelslike_min	Minimum Feels Like, ºC
Precip_cover	Precipitation Cover, %
Snow	Snow, cm
Snow_depth	Snow Depth, cm
Moon_phase	Moonphase

Table A3: New Variables - Data Dictionary

Variables	Description
Day_sin	Trigonometric encoding of the cycle in sine wave format for Day
Day_cos	Trigonometric encoding of the cycle in cosine wave format for Day
Weekday_sin	Trigonometric encoding of the cycle in sine wave format for Weekday
Weekday_cos	Trigonometric encoding of the cycle in cosine wave format for Weekday
Week_sin	Trigonometric encoding of the cycle in sine wave format for week
Week_cos	Trigonometric encoding of the cycle in cosine wave format for Week
Month_sin	Trigonometric encoding of the cycle in sine wave format for Month
Month_cos	Trigonometric encoding of the cycle in cosine wave format for Month
Quarter_sin	Trigonometric encoding of the cycle in sine wave format for Quarter
Quarter_cos	Trigonometric encoding of the cycle in cosine wave format for Quarter
Hours_sun	Number of Hours with Sun
Temp_range	Temperature Range
Holiday	1 if Holiday, 0 o/w
Hours_without_sun	Working Hours Without Sun
Growth_Rate_Percent	Quarterly Consumption Growth Rate Percent
SeasonSpring	1 if Spring, 0 o/w
SeasonSummer	1 if summer, 0 o/w
SeasonWinter	1 if Winter, 0 o/w
MA_week_...	Week Moving Average
MA_month_...	Month Moving Average

B Descriptive Analysis



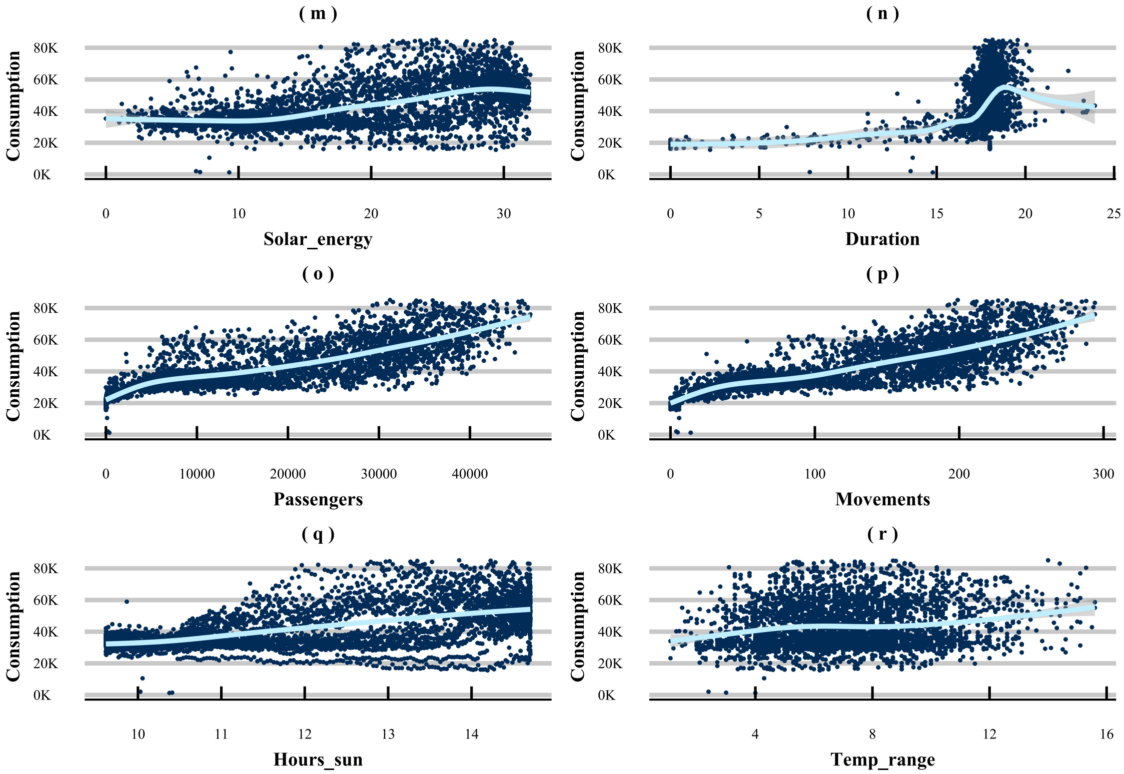


Figure B1: Distribution of Consumption over (a) Maximum Temperature; (b) Minimum Temperature; (c) Mean Temperature; (d) Dew Point; (e) Humidity; (f) Precipitation; (g) Wind Speed; (h) Wind Direction; (i) Sea Level Pressure; (j) Cloud Cover; (k) Visibility; (l) Solar Radiation; (m) Solar Energy; (n) Duration; (o) Passengers; (p) Movements; (q) Hours of Sun; (r) Temperature Range