A Comparison Between Seasonal and Non-Seasonal Forecasting Techniques for Energy Demand Time Series in Smart Grids

Sabereh Taghdisi Rastkar, Danial Zendehdel, Enrico De Santis\textsuperscript{a} and Antonello Rizzi\textsuperscript{b}

Department of Information Engineering, Electronics and Telecommunications, Sapienza University of Rome, 00185 Roma, Italy

Keywords: Time Series Energy Forecasting, Forecasting Algorithm, Seasonality Effect.

Abstract: Accurate energy consumption forecasting is essential for optimizing resource allocation and ensuring a reliable energy supply. This paper conducts a thorough analysis of energy consumption forecasting using XGBoost, SARIMA, LSTM, and Seasonal-LSTM algorithms. It utilizes two years of hourly electricity demand data from Italy and the PJM region (USA), categorizing algorithms into seasonality and non-seasonality groups. Performance metrics like RMSE, MAE, $R^2$, and MSPE are employed. The study underscores the importance of considering seasonality, with SARIMA and Seasonal-LSTM achieving high accuracy in the seasonality group. In the non-seasonality group, XGBoost and LSTM perform competitively. In summary, this research aids in choosing suitable forecasting algorithms for building an Energy Management System for smart energy management in microgrids, considering seasonality and data attributes. These insights can also benefit energy companies in efficient resource management, promoting sustainable energy practices and urban development.

1 INTRODUCTION

Consider an urban area that relies solely on renewable energy sources. In this setting, accurately predicting energy consumption and demand during different seasons and public holidays becomes critical. However, forecasting the utility consumption aids in balancing the generation and demand of energy.

Energy consumption problems have become a practical research topic in recent years. Energy problems are crucial for the security and well-being of societies (Ghalekhondabi et al., 2017). However, unlike many other energy sources, electricity must be consumed immediately after generation. Thus, forecasting future electricity demand is vital for power companies to allocate resources and guarantee sufficient supply effectively. This information about consumption and demand aids firms in implementing unique energy conservation strategies, as storing electricity is often prohibitively expensive, inefficient, or unfeasible. Consequently, balancing electricity consumption and generation becomes critical (Nguyen and Hansen, 2017). Thus, forecasting future power demand is crucial for power companies in their energy management efforts (Hamzaçebi et al., 2019).

Moreover, the growing world population and increasing use of advanced technologies are expected to drive electricity demand. The emergence of smart grids has made load prediction systems indispensable for sustainable growth and intelligent urban development (Azeem et al., 2021). With the advent of intelligent networks, power demand forecasting will become increasingly important (Hamzaçebi et al., 2019).

Moreover, Renewable Energy Communities equipped with Intelligent Energy Management Systems (EMS) have emerged as potent catalysts for change. These communities, which leverage a combination of renewable energy sources like solar, wind, and hydro, are reshaping production and consumption. However, the variable and stochastic nature of renewable energy generation presents unique challenges that require sophisticated management strategies. Hence, a high-performance forecasting system plays a significant role in tackling this problem. Accurate and reliable energy consumption and production forecasts are the backbone of any effective EMS. A well-calibrated forecasting model can enable real-time adjustments to energy distribution, ensuring supply reliably meets demand while minimizing waste. In this way, advanced

\textsuperscript{a} https://orcid.org/0000-0003-4915-0723
\textsuperscript{b} https://orcid.org/0000-0001-8244-0015
forecasting is not merely an adjunct but a central component of an Intelligent EMS in enhancing both Renewable Energy Communities’ economic and environmental sustainability.

Load forecasting has been a long-standing technique used to predict future demand. It plays a critical role in the precise design and placement of electrical loads at various time intervals within the planning horizon. Therefore, the potentially significant cost savings of accurate load forecasts bring significant benefits to electrical utilities (Singh et al., 2012). Predicting load demand and managing electricity are essential for energy conservation (Nepal et al., 2020). Numerous forecasting algorithms have been devised to enhance forecast accuracy, each with unique strengths and weaknesses. Selecting the ideal algorithm for a given situation requires comparing various algorithms in different settings. Understanding data trends and their association with different types of seasonality is crucial for accurately predicting energy demand (Hong et al., 2016). For instance, air conditioning increases during warm months, leading to higher power consumption. Conversely, the demand for heating rises in colder seasons, affecting energy usage. Energy consumption also fluctuates on holidays and weekends.

In this study, we categorize forecasting techniques into two primary groups: those reliant on seasonality, such as SARIMA and seasonality LSTM, and those independent of it, represented by XGBoost and LSTM. We will then conduct a comprehensive analysis of these methods and their inherent characteristics. This inquiry underscores the significance of utilizing precise forecasting techniques and tools to ensure a dependable energy supply and enhance energy management. Lastly, we will assess the performance of these methods on diverse data sets using metrics such as mean absolute error, mean square percentage error, and root mean square error.

The remainder of this study is organized as follows. In Sec. 2 the technical literature is revised. In Sec. 3

2 RELATED WORK

The growing importance of energy forecasting in the business sector is undeniable, particularly in the field of renewable energies. Accurate demand forecasting is vital for energy planning, efficient resource allocation, and cost savings for energy suppliers. Over the years, traditional forecasting techniques have been developed to predict energy demand and have historically been reliable. However, the increasing complexity of the energy system and the advent of new technologies may render these traditional methods insufficient. These traditional mathematical models, based on the Box and Jenkins method, are mainly statistical and they are categorized as linear methods that employ a linear functional form for the time-series models. It encloses the Auto-Regressive Integrated Moving-Average (ARIMA) model (Wang et al., 2012; Yukseltan et al., 2017; Amini et al., 2016; Debusschere et al., 2012), exponential smoothing model (Chen et al., 2010), linear model (Zhou, 2017), and regression analysis methods (Fumo and Biswas, 2015; Amber et al., 2015).

Moreover, some mature nonlinear methods, such as Artificial Neural Networks (ANNs) (Tian and Hao, 2018; Ganesan et al., 2015) and Support Vector Machines (SVMs) (Zhang and Wang, 2018), have been employed. For instance, some studies showed that a range of energy demand forecasting models for time series, such as regression and soft computing techniques (including fuzzy logic, genetic algorithms, neural networks, and support vector regression), are extensively used for demand side management (Singh et al., 2012; Avami and Boroushaki, 2011; Suganthi and Samuel, 2012). The ANNs are limited by insufficient data for accurate forecasting (Tian and Hao, 2018). Neural networks need many control parameters, which have difficulties getting a stable solution, risks of over-fitting, and restraint by insufficient data (Tian and Hao, 2018; Ganesan et al., 2015). Alongside machine-learning algorithms, Deb (Deb et al., 2017) thoroughly reviewed existing machine-learning techniques and ANNs for forecasting time-series energy consumption. In 2017, the BRE Trust Centre reviewed machine-learning algorithms like artificial neural networks, support vector machines, and time series analysis for short and very short-term prediction, evaluating their performance using several metrics (Kuster et al., 2017). More recently, a study examined eight methods for predicting electricity demand in supermarkets, schools, and residential buildings at the individual structure level, employing statistics, machine learning, and a median ensemble technique (Groß et al., 2021). Additionally, researchers at the University of Brasilia utilized regularized machine learning models to predict short- and medium-term energy consumption in Brazil, comparing them against standard criteria such as the Random Walk and the ARIMA(Albuquerque et al., 2022).

Seasonality is a characteristic of a time series in which the data experiences regular and predictable changes that recur every calendar year. These changes are often influenced by the seasons of the year, holidays, and other recurring events. Seasonality
can significantly impact the patterns and trends observed in time-series data. In 2008, Lam considered the effect of seasonal variations on used energy, mainly caused by air-conditioning requirements changes (Lam et al., 2008). Additionally, holiday periods can cause spikes in energy consumption due to increased activities in homes and commercial establishments. A study by the U.S. Energy Information Administration highlighted the significant seasonal variation in energy consumption across the country, attributing it to weather-related factors, holiday patterns, and even the academic calendar (Outlook et al., 2010). Furthermore, the length of the day can also play a role, with longer daylight hours in the summer leading to reduced lighting needs.

Several researchers have also explored methodologies to address seasonality in energy consumption forecasting. In a study by Rashedul, the seasonal decomposition of time series (STL) approach was proposed to model and predict energy consumption, effectively capturing the seasonality patterns (Haq and Ni, 2019). In another study, Xiong significantly improved the accuracy and speed of forecasting energy consumption (Xiong et al., 2021). Moreover, several studies focused on seasonal SARIMA. For instance, Wang presented a combination of PSO optimal Fourier method models with seasonal ARIMA for energy consumption prediction (Wang et al., 2012). This research aims to compare traditional and advanced techniques for electrical load forecasting to assist suppliers in selecting efficient methods, considering impact of seasonality, and ensuring long-term sustainability.

3 METHODOLOGY

The methodology outlined in this study involves several crucial steps designed to compare energy consumption prediction results while accounting for seasonality’s influence. We have categorized our approach into two groups as follows:

• Group A: Non-seasonal forecasting models
• Group B: Seasonal forecasting models

In each of these groups, A and B, we have included two distinct forecasting models. Group A comprises the XGBoost (Wang et al., 2021) (Phan et al., 2021) and LSTM algorithms, while Group B integrates the SARIMA and Seasonal-LSTM models. We have applied these four forecasting algorithms (XGBoost, SARIMA, Seasonal-LSTM, and LSTM) to the two distinct datasets previously mentioned - see Sec. 4.1 below. To ensure precise predictions, we have customized the models to align with the unique characteristics of each dataset, using Python. The subsequent section will provide a comprehensive breakdown of the methodology employed for forecasting energy consumption.

3.1 System Model

In this section, we offer a summary of the fundamental elements and procedures required to create a system model that integrates various machine learning algorithms. Figure 1 presents a holistic perspective of the steps involved in time series forecasting as conducted in this study.

![Figure 1: Flow chart of time series energy forecasting.](image)

3.2 Seasonal Effects

Seasonality in time series data refers to regular and predictable changes that occur annually, often tied to seasons, holidays, and recurring events. These patterns significantly influence the observed data trends. In the context of energy consumption, seasonality is pronounced due to factors like weather variations, leading to increased heating or cooling needs during extreme seasons. Holidays can also cause spikes in energy use, as can the length of daylight hours. Neglecting seasonality in energy consumption forecasting models can lead to inaccurate predictions. Existing research highlights that accounting for seasonality greatly improves forecasting model accuracy, ultimately aiding in more effective energy planning and management.

3.3 Algorithm Description

This section provides concise explanations of each forecasting algorithm utilized in this study.

Non-seasonal forecasting models:

• LSTM (Long Short-Term Memory), a type of recurrent neural network (RNN), is well-suited for
sequence prediction tasks, particularly in time series forecasting. This study configures the LSTM model with three LSTM layers, each containing 40 units. Dropout layers with a 0.2 dropout rate are incorporated to prevent overfitting. The model processes sequences of length 20, effectively capturing temporal dependencies to make single predictions. With 32,681 trainable parameters, the model is designed to decode complex time series trends, enhancing its capability to uncover intricate data patterns.

- **XGBoost (Extreme Gradient Boosting)** is a powerful machine learning algorithm for accurate forecasting by combining predictions from multiple decision trees. It’s known for its efficiency and versatility, suitable for various forecasting tasks. The model in use is configured as a regression model, with extensive hyperparameters, 100,000 boosting rounds, a learning rate of 0.05, max depth of 5, no gamma value, 80% subsampling, and objective function ‘reg:squarederror’.

**Seasonal forecasting models:**

- **SARIMA (Seasonal Autoregressive Integrated Moving Average)** enhances upon ARIMA for seasonal time series data, introducing three seasonal parameters (P, D, Q) and a seasonality factor ‘s’. The provided SARIMA model is adaptable to various data sets. Auto_ARIMA from pmdarima automatically tunes model parameters. SARIMA order is extracted, and SARIMAX is initialized for training. A rolling forecast predicts test data, repeatedly re-optimizing SARIMA.

- **Seasonality decomposition in LSTM forecasting** entails the dissection of a time series into its constituent parts: seasonal, trend, and residual components. This dissection is critical for gaining a deeper comprehension of the time series’ patterns and fluctuations, thereby enabling more accurate predictions. By separating these components, LSTM models can adeptly harness seasonality information, significantly enhancing their forecasting capabilities. To execute this decomposition of time series data, the seasonal decomposition method from the statsmodels library is employed.

4 EXPERIMENTAL RESULT

4.1 Data Sets

This research utilizes four distinct data sets sourced by Kaggle. Each data set offers unique insights into different energy consumption scenarios, providing a broad spectrum of data for the analysis. All data sets comprise two columns: the observation date and the corresponding energy consumption value. The date column captures the chronological progression of the data, while the energy consumption column measures the actual energy usage. These data sets were specifically selected for their relevance and potential to reveal valuable patterns and trends influencing energy consumption. In our training, we modified the energy consumption data from an hourly to a 6-hourly frequency. This change is because the original data set exhibited strong seasonality, leading to high complexity in the models. By aggregating the data into 6-hour intervals, we effectively reduced the complexity and produced a more condensed time series, allowing the models to fetch the underlying patterns and trends.We provide a detailed description of each data set. We focused on hourly data from the past two years of the following data sets:

- **PJM Hourly Energy Demand for the years 2016-2018** (Albuquerque et al., 2022) (Khan et al., 2022).
- **Italy’s Hourly Energy Consumption for 2020-2022** (Lisi and Edoli, 2018) (Rossi and Brunelli, 2013).

Furthermore, Figure 2 provides a clear presentation of the mean features extracted from the two data sets illustrated in Figure A (PJM) and Figure B (Italy). These features are derived using the XGBoost algorithm, which employs an iterative process of constructing decision trees and assessing the impact of each feature on the model’s predictive accuracy. This information holds significant value, as it aids in discerning the most pertinent and influential features within the data set (feature selection).

4.2 Evaluation Metrics

Common metrics used to evaluate forecast accuracy include Mean Absolute Error (MAP) and other evaluation metrics such as RMSE and MSPE, in addition to them we have used R-Squared ($R^2$) as a measure that compares the stationary part of the model to a simple mean model. $R^2$ can be evaluated by 1. This metric, also known as the coefficient of determination, measures the proportion of the variance in the actual ob-
In Eq. 1 $SS_{res}$ is the Sum of the Square of Residuals. Here, residual is the difference between predicted and actual values, and $SS_{tot}$ is the Total Sum of Squares.

\[ R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \] (1)

In this section, we conduct rigorous experiments on two sets of forecasting algorithms (XGBoost, SARIMA, LSTM, Seasonal-LSTM) using two years of hourly electricity demand data. Our main objective is to evaluate each model’s forecasting accuracy. The focus is on comparing algorithm outcomes concerning seasonality considerations.

Figure 3 visually illustrates the energy consumption patterns in Italy and the United States throughout the week, emphasizing a significant rise during weekdays and a subsequent decline as the week progresses toward the weekend. It is notable that the highest energy consumption occurs at the end of each weekday.

Figure 4 shows seasonality decomposition of PJM data set of energy consumption visualized in separate figures of trend, seasonality and residual.

Within the category of seasonality forecasting models, Figure 5a illustrates the prediction results of the SARIMA model, configured with non-seasonal orders of (1, 0, 4) and seasonal orders of (4, 1, [1, 2], 12), applied to the PJM data set. Conversely, Figure 5b showcases the prediction outcomes of the SARIMA model, specifically configured as SARIMA(3, 0, 4)x(4, 1, [1, 2], 12) in Italy data set.

Figure 6 depicts a comparison of the actual and predicted energy consumption in both PJM and Italy data sets, respectively. This comparison is based on the results generated by the seasonal LSTM forecasting algorithm. The graph provides a visual representation of how closely the predicted values align with the real data.

In the group without seasonality considerations, Figure 7 displays the results of the LSTM algorithm’s energy consumption forecasting. This figure offers a visual comparison between the actual energy consumption and the model’s predictions. The X-axis represents time, while the Y-axis represents energy consumption levels. This visual assessment allows us to gauge the accuracy and performance of the prediction model. Any disparities between the actual and predicted values may indicate areas where further improvements to the model are needed.
The final algorithm in the non-seasonal group is XGBoost. Figure 8 illustrates the energy consumption forecasting results generated by this model for two specific dates within a one-month period. The figure demonstrates that the forecasting model has achieved a reasonably successful alignment with the actual energy consumption patterns.

Table 1 showcases the outcomes of both seasonal and non-seasonal models, enabling a comparison of how seasonality impacts time series forecasting.

5 CONCLUSIONS

This study has provided valuable insights into the realm of energy consumption forecasting, a critical component for optimizing resource allocation and ensuring a dependable energy supply. We conducted a rigorous evaluation of four distinct forecasting algorithms, namely XGBoost, SARIMA, LSTM, and Seasonal-LSTM, utilizing two years’ worth of hourly electricity demand data from both Italy and the PJM region.

Our findings underscore the paramount impor-
Table 1: Result of two forecasting models groups for PJM and Italy data sets.

<table>
<thead>
<tr>
<th></th>
<th>ITALY</th>
<th>PJM</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE</td>
<td>MAE</td>
<td>$R^2$</td>
</tr>
<tr>
<td>Seasonality</td>
<td>SARIMA</td>
<td>0.0492</td>
<td>0.0302</td>
</tr>
<tr>
<td></td>
<td>S-LSTM</td>
<td>0.0087</td>
<td>0.007</td>
</tr>
<tr>
<td>Non-seasonality</td>
<td>XGBoost</td>
<td>0.0601</td>
<td>0.0424</td>
</tr>
<tr>
<td></td>
<td>LSTM</td>
<td>0.0228</td>
<td>0.0174</td>
</tr>
</tbody>
</table>

Importance of accounting for seasonality when forecasting energy consumption. Within the seasonality group, the SARIMA and Seasonal-LSTM models emerged as standout performers, exhibiting exceptional accuracy with $R^2$ values that nearly approached 1. These models adeptly captured the inherent seasonal patterns in energy consumption, showcasing their robust forecasting capabilities.

In contrast, the non-seasonality group witnessed competitive performances from the XGBoost and LSTM models. While their $R^2$ values were slightly lower, they still demonstrated strong forecasting prowess, with LSTM particularly noteworthy for achieving an impressive $R^2$ score of 0.98.

The results presented in Table 1 reinforce the superiority of seasonality-aware models, with SARIMA and Seasonal-LSTM outperforming others by achieving lower RMSE and MSPE values while securing higher $R^2$ scores. These models excelled in effectively capturing and leveraging the cyclic variations in energy consumption.

It’s also worth noting that the performance of each model varies depending on the data set. For instance, LSTM with seasonal decomposition exhibited a high value for MSPE in PJM data sets. This highlights the importance of data set-specific considerations when comparing the results of SARIMA and LSTM with seasonal decomposition in the seasonality model group. SARIMA inherently possesses seasonal properties within the algorithm, while LSTM with seasonal decomposition incorporates these properties externally.

Nonetheless, it’s essential to acknowledge that other factors, such as weather conditions, can significantly influence the forecasting outcomes, potentially impacting algorithm selection for a given time series data set.

In conclusion, this research serves as a valuable resource for selecting appropriate forecasting algorithms, considering seasonality and data characteristics. Its insights hold great potential for energy companies seeking to elevate their resource management practices, thereby contributing to sustainable energy strategies and intelligent urban development. The utilization of accurate forecasting models can substantially enhance the allocation and optimization of energy resources, establishing them as indispensable tools in today’s dynamic energy landscape.

ACKNOWLEDGEMENTS

This study was carried out within the MOST – Sustainable Mobility Center and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1033 17/06/2022, CN00000023). This manuscript reflects only the authors’ views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

REFERENCES


