Spain's Energy Trends and Resource Utilization Solutions

Arohi Shiraskar, Eddie Dew, Sophia Hamid, Baran Narravula

The University of North Carolina at Charlotte

DTSC 2302

Marco Scipioni, Christopher Dong

May 06, 2024

Spain's Energy Generation Trends: Cost-Effective Utilization of Resources 1.0

Spain faces the challenge of meeting energy demands while managing environmental impact and ensuring economic sustainability. With growing concerns about climate change and the need for energy security, the efficient utilization of resources becomes primary. This paper delves into Spain's energy landscape, focusing on the efficient utilization of resources to meet the nation's growing energy demands. It seeks to understand the dynamics of energy production, its associated costs, and the implications for sustainability and economic viability. This is achieved through examining the sources' load-to-price ratios, comprehensive data analysis, prediction and forecasting techniques to obtain a cost-effective solution for Spain's energy production. Thus, the central idea of the study is to investigate how Spain can optimize its energy generation strategies to achieve cost-effective utilization of the renewable and nonrenewable energy resources while meeting environmental and policy objectives.

Context and Implications 2.0

Stakeholders, including government agencies, energy companies, environmental organizations, and the general public, are deeply impacted by the findings of this study. Government policymakers rely on such analyses to formulate effective energy policies that balance economic interests with environmental sustainability. Energy companies stand to benefit from insights into market trends and cost-efficiency measures, which can enhance their competitive edge and profitability. Environmental organizations advocate for sustainable energy practices, making the project's findings crucial for advancing environmental goals. Lastly, the general public, as consumers and citizens, are directly affected by energy prices and environmental impacts, making them important stakeholders in the energy discourse.

The relevance of such research is further underscored by a 2020 study exploring the influence of climate on energy consumption and CO2 emissions from the MDPI Journal. This study explored the relationship between renewable and nuclear energy consumption, carbon dioxide emissions, and economic growth in Spain from 1970 to 2018. The analysis revealed that there is a positive correlation between the use of nuclear energy and economic growth, but economic growth only increases through the use of renewable energy. Although nuclear power and renewable energy help reduce greenhouse gas emissions, increased economic activities offset this reduction. (Piłatowska et al., p. 1). This research informs the need for a balanced approach to energy production and its impact on the environment and economy.

Exploring more specifically on which energy sources were able to achieve the balance of usage and pricing, it was studied that the cities in severe conditions use more energy depending on the severity of the seasons. For example, in cooler climates, thermal energy is used more frequently for heating, whereas in warmer climates, electric energy is used more frequently for both heating and cooling. In particular to production of the C02 emissions, an article about climate influences stated, “climate zones with higher consumptions produce higher emissions; climate zones with severe climates produce higher emissions” (Zarco-Soto et al., 2020, p. 15). With Spain being a Mediterranean like climate, they will experience seasonal changes with hot dry summers and colder winters, which can result in higher emissions at those times.

Additionally, the relationship between energy consumption and costs, can also provide insight into market and energy behavior. For example, a study by ScienceDirect on the relationship between energy consumption and prices. The study analyzed the energy market in Iberia, with a particular focus on the relationship between energy prices and consumption in Portugal and Spain. Their findings show a connection between market prices and energy consumption. The factors that lead to price and consumption fluctuation are market concentration, regulatory limitations, and a poor economic recovery (Gil-Alana et al., 2020, p. 3). Spain has a thriving economy which means their prices won't fluctuate as much, resulting in a more balanced distribution of price. This can help our model focus more on the actual load without concerning the price variability.

To ensure this economic stability, Spain is faced with challenges and tradeoffs in its policies. Spain incorporates two main policies, the European Union Emissions Trading Scheme (EU ETS) and Electricity from Renewable Energy (RES-E) policies, each having different benefits. RES-S, in particular, enables renewable energy producers to earn additional revenue and guaranteed payments, encouraging further investment in renewable electricity production. In an article from the Taylor and Francis journal about the interactions of energy policies, it was stated, "A greater RES-E deployment positively contributes to local benefits, CO2 emissions reductions, and dynamic efficiency benefits." (Del Río, p. 13). This approach aligns emissions reductions with the deployment of renewable energy, promoting a more integrated policy framework. However, on the other hand, the European Union's ETS is designed to promote demonstrating greater cost-efficiency, which increases the consumer cost, discouraging social benefits. The coexistence of those policies with different objectives have created conflicts and made it difficult for Spain to reach their targets. Those policies are relevant to this research as they highlight the complex imbalance between energy production, emissions reduction, and regulatory frameworks.

While optimizing energy generation holds promise for reducing carbon emissions and fostering economic growth, it also raises concerns about environmental degradation, resource depletion, and social equity. Thus, ethical evaluation involves weighing the potential benefits of cost effective energy utilization against the risks of environmental harm and social injustice. The ethical framework in this project is based on principles of sustainability, equity, and accountability. Sustainability entails balancing present needs with future generations interests, ensuring that energy practices do not compromise environmental integrity or resource availability. Equity requires ensuring that costs and benefits are fairly distributed, protecting vulnerable communities from bearing disproportionate burdens or facing exclusionary policies. Accountability involves transparency and responsibility in decision-making, holding stakeholders accountable for the social and environmental impacts of energy generation initiatives.

Measurements 3.0

The research aims to explore the efficiency of renewable and non-renewable energy sources by analyzing their load-to-price ratio, and cost-effectiveness is operationalized as the amount of energy produced per unit cost. The goal is to identify which energy source offers the most load for the least cost, providing data-driven recommendations for budgeting and optimal energy resource allocation.

The given variables were measured from a dataset by the European Association for the Cooperation of Transmission System Operators for Electricity (ENTSO-E). It included hourly data on energy generation, categorized by production type in megawatts, spanning from 2015 to 2018, along with the respective energy prices. The dataset was then divided into renewable and non-renewable groups, and each category's energy load was calculated per Euro ($\frac{Load}{Total\ Price}$). Renewable energy refers to energy generated from naturally replenishing sources such as sunlight, wind, and water. Non-renewable energy, on the other hand, encompasses finite sources that are not self-replenishing, including coal, oil, and nuclear power.

To evaluate the energy resources' efficiency, the data was examined through a variety of visualizations, including line graphs, box plots, and bar graphs, to detect any significant trends or patterns in energy generation and pricing over the four-year period. Furthermore, a time-series analysis was used to forecast the renewable and non-renewable load/euro for 2019, incorporating other confounding factors such as seasons and holidays that could influence energy consumption. Once the projected energy load was calculated, various regression models will be experimented with to determine which one offers the best fit for predicting the corresponding energy price in 2019.

Looking into future predictions was relevant as it forecasted a possible scenario Spain could experience. The predictions would help to gain insights into potential price fluctuations and provide data-driven guidance for energy management and budgeting. Overall, this combined methodology offers a comprehensive outlook on energy generation and pricing, facilitating improved budgeting for upcoming years in Spain.

Data section 4.0

Cleaning and Visualizing

After identifying a variable of interest, the data set was analyzed to find out how the target variable is affected over time. Before the analysis, the data needed to be manipulated in order to be presented to a model error free. First, the feature names were not clean with unnecessary capitalization and spacing (ex "Generation biomass"). A python package called pyjanitor was used to convert the column names to snake case, making the dataframe much easier to manipulate. Next, missing values were checked as the model would not be able to handle missing values as it wouldn't be able to learn how to handle them. For this case, interpolation was used to linearly fill these missing values to continue our continuous line of data. With a cleaner dataset, the metric was computed. Figure 4.1 shows the statistical summary of the newly implemented features.

	Renewable Load/Euro	Non-renewable Load/Euro
Count	35,064	35,064
Mean	0.35	0.21
Standard Deviation	0.21	0.09
Min	0.00	0.01
25%	0.20	0.15
50%	0.32	0.20
75%	0.47	0.26
Max	1.80	0.81

Figure 4.1 Summary Statistics of the Renewable and Non-renewable Load/Euro dataset

When comparing the two, the renewable load/Euro has a higher spread compared to the non-renewable dataset meaning that there most likely could be a certain time period where non-renewable output was more efficient than the renewable. Lastly, outliers were analyzed within the dataset. With a wide range of values, a lot of outliers were introduced, which can be a problem to the model since they could decrease the confidence and overfit to a particular feature. Outliers should be limited for maximum performance while not removing a decent chunk of the dataset, so only extreme outliers were removed, totaling 834 observations. While there were some outliers still in the dataset, model selection was deemed crucial to handle these values. More on that down below.

Feature Selection

Since this is a Time Series Analysis model, the features needed to align as such. The first set of features that were added were the simple time features including things like hour, day, year, month, season, etc. These features can help the model find patterns as to when the load output per euro could see an increase or decrease at a given point in time. Figure 4.2 shows an example of the efficiency for both resources on a per week basis coupled with the season.

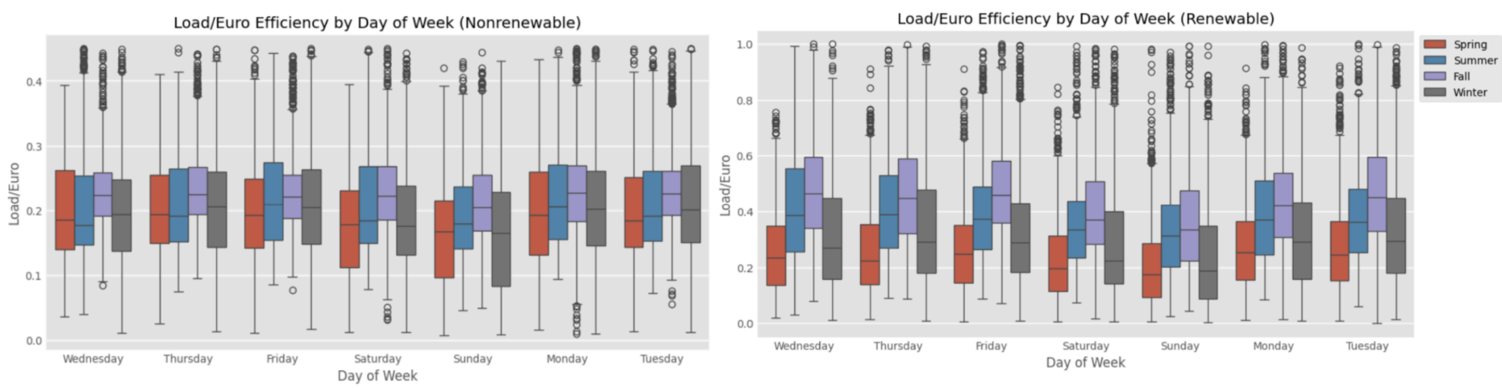

Figure 4.2 Seasonal patterns for each resource type

Outliers are in this chart, but they are pretty confined which is not too much of an issue for the purposes of this model.

The next feature that was added was Spain holidays. The holidays were extracted from the holidays package in python and merged with the main dataset. Spain doesn't have too many holidays compared to other countries, but it should still be considered in the model. The final feature that was added was lag. Lag is simply taking the previous value from the dataset at some given time and adding it to the same row, giving the model some trend to follow. The lag feature for this model has 3 lag features: each ranging from one to three years back at a given point in time. This totals to 13 features. Note: 2 features (lag3 and quarter) were removed as they had no impact on the model. Figure 4.3 shows the correlations between the features and respective target variables.

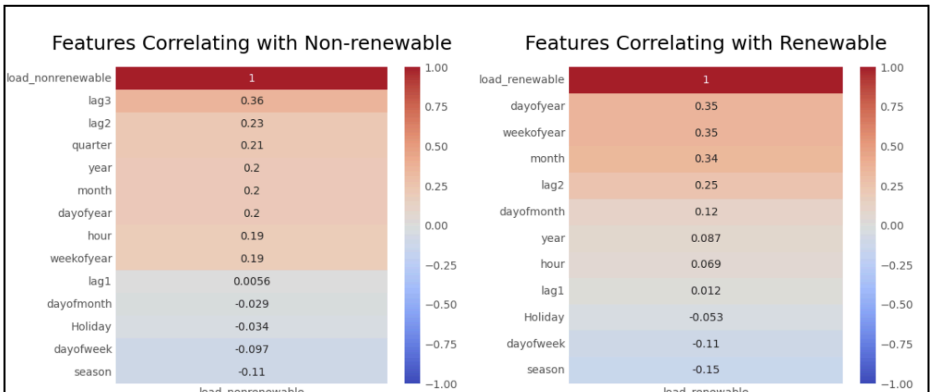

Figure 4.3

Building the Model

The common theme throughout preprocessing the data is its time series and it has two target variables, renewable and non-renewable. These are important to identify for the modeling process. Before the model is explained there are a couple of things that should be noted.

- Time Series Split: a feature in scikit learn that splits the data into time series intervals so the model can see different time variants of the data.
- Multi Output Regressor: a feature in scikit learn that is able to take one model and use it on multiple target variables.

The model uses Time Series Split by using "folds" of the data split into 3 years. So the model will train and test from 2015-2017, 2015-2018, and 2015-2019. Through each fold, the model uses a Gradient Boosting Regressor to make predictions. The reason why Gradient Boosting Regressor is used is it has the ability to use parameters that prevent overfitting, such as the learning rate, and its capacity to utilize tree-based estimators for accurately estimating the values of the target variable.

Model Results and Predictions

After some trial and error, the model was performed pretty well. For non-renewable, it was on average off by 0.10 loads/Euro with a Mean Absolute Percentage Error (MAPE) of 44.50% which is not too bad for this kind of model with high variation. Renewable was a little more difficult at predicting as the model was off on average of 0.19 loads/Euro with a MAPE of 81.94%. Definitely some room for improvement, but overall the model was able to predict the shape of the dataset well. Upon further inspection of the particular times the model was most off, weekends later in the year tended to be a little more difficult to predict. A feature not included in the model that could counteract this is adding sports dates, in particular soccer as the primary feature as it becomes popular later in the year with different Cup series occurring. Figure 4.4 shows a feature importance plot of the renewable and non-renewable load/Euro prediction from the testing and training sets. The model found day of the year the most useful feature for both target variables. Finally, the model was used to predict what might happen in 2019. Figure 4.5 shows the output for the resulting year. The model predicts that Renewable load/Euro output will increase over the course of the year while the non-renewable load/Euro will remain relatively constant. The main takeaway from these results is to invest in non-renewable resources at the start of each month and use renewable resources for the majority of the year to have the most efficient load output per Euro spent on the resource.

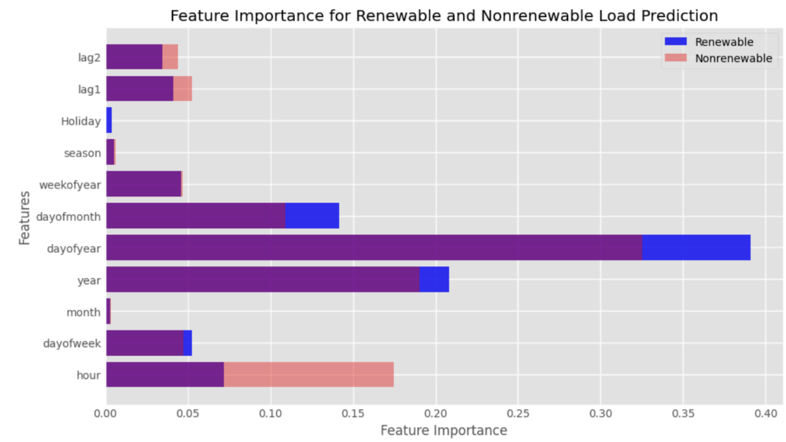

Figure 4.4 Feature Importance Plot

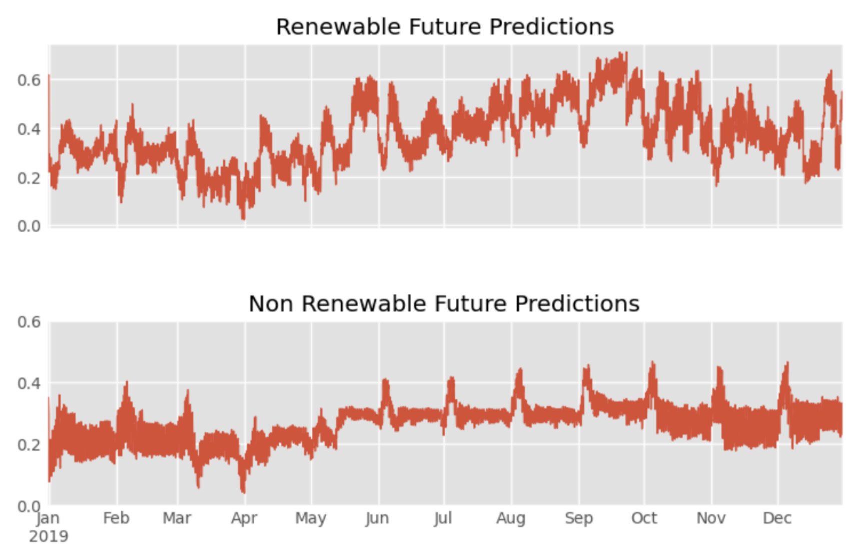

Figure 4.5 2019 Predictions

Conclusion 5.0

In conclusion, Spain is a well developed country with effective control over the energy output relative to the investment in renewable and non-renewable resources. Although its RES-E policy may be strict and increase production costs, they are potentially willing to optimize their system. A model has been built to help identify trends and recommend optimal resource use to make most of Spain’s energy budget. While the model isn't perfect, some features can suggest efficient energy patterns across the year. It proposes that renewable energy should be used later in the year, with non-renewable energy be used sparingly at the beginning of each month to fully maximize the usage. If this approach is not feasible, another strategy is to start the year with heavy reliance on non-renewable, gradually moving to renewable as the year progresses. However, there are limitations to this solution. The data analyzed in the model consider energy production load and overall energy prices at a national level, but other factors such as location, climate, or specific events can also significantly impact energy production and cost. Therefore, it isn't always possible to derive accurate conclusions from generalized data. To improve the model, additional variables that account for specific locations and weather conditions should be considered. Despite these limitations, the suggested strategies could be helpful in guiding Spain's energy policies towards more sustainable practices and better resource management.

References

Del Rio, P. (2009). Interactions between climate and energy policies: the case of Spain. Climate Policy, 9(2), 119–138. https://doi.org/10.3763/cpol.2007.0424

Gil-Alana, L. A., Martin-Valmayor, M., & Wanke, P. (2020). The relationship between energy consumption and prices. Evidence from futures and spot markets in Spain and Portugal. Energy Strategy Reviews, 31, 100522–100522. https://doi.org/10.1016/j.esr.2020.100522

Piłatowska, M., Geise, A., & Włodarczyk, A. (2020). The Effect of Renewable and Nuclear Energy Consumption on Decoupling Economic Growth from CO2 Emissions in Spain. Energies (Basel), 13(9), 2124-. https://doi.org/10.3390/en13092124

Zarco-Soto, I. M., Zarco-Periñán, P. J., & Sánchez-Durán, R. (2020). Influence of climate on energy consumption and CO2 emissions: the case of Spain. Environmental Science and Pollution Research International, 27(13), 15645–15662. https://doi.org/10.1007/s11356-020-08079-7