WIND POWER GENERATION DATA ANALYSIS REPORT

Insights and Trends from 2014 Dataset

ABSTRACT

This report presents an analysis of wind power generation data collected throughout the year 2014. Through exploratory data analysis and seasonal decomposition techniques, key insights into the trends and patterns of wind power generation have been uncovered. The findings offer valuable information for understanding the dynamics of renewable energy production and informing strategic decision-making in the renewable energy sector.

Alain Destin Nishimwe Karasira

27 February,2023

Contents

Introduction 1
Dataset Overview 2
Objectives 2
Data Preprocessing 2
Reading the Dataset 2
Evaluating the presence of missing values. 2
Data analysis 3
Time Series Plot 3
Plotting the Autocorrelation Series 4
Exploring Seasonal Trends in Wind Power Generation. 4
Visualizing Monthly Power generation.. 5
Average daily energy demand by hour. 6
Weekly Energy generation profile 6
Hypothesis testing 7
Modeling and Forecasting 7
Persistence Forecasting 7
Conclusion 8

Introduction

This report presents the findings of an in-depth analysis conducted on wind power generation data collected at 15-minute intervals throughout the year 2014. The objective of this analysis is to gain insights into the patterns and trends of wind power generation, which are essential for renewable energy planning and management.

Dataset Overview

The dataset used for this analysis comprises wind power generation measurements taken every 15 minutes for the entire year of 2014. Each data point includes the timestamp and corresponding wind power generation value.

Objectives

- Analyze the overall trend of wind power generation throughout the year.
- Identify any seasonal patterns or fluctuations in wind power generation.
- Explore the relationship between weather conditions and wind power generation.

Data Preprocessing

Reading the Dataset

```
# Read a dataset into a dataframe
Grid_df=pnd.read_csv('../Dataset/EirGridSystemDemand2014.csv',parse_dates=[['Date', 'Time']])
Grid_df.info()
```

Two parameters are included in the above function read_csv, one being the **data set file path**, and the **parse_dates** which specifies which columns should be parsed as datetime objects. In the given code, it combines the columns "Date" and "Time" into a single datetime column. This can be useful for time series analysis where the date and time information needs to be combined and treated as a single entity.

Evaluating the presence of missing values.

During the data preprocessing stage, one crucial task was to manage missing values within the dataset. To accomplish this, we employed a technique known as linear interpolation. Linear interpolation involves estimating the values that are missing based on the neighboring data points. This method is particularly useful for time series data analysis, as it allows us to fill in the gaps in the dataset by calculating intermediate values.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 35040 entries, 0 to 35039
Data columns (total 2 columns):
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   Date_Time  35040 non-null  datetime64[ns]
 1   Demand     35035 non-null  float64
dtypes: datetime64[ns](1), float64(1)
memory usage: 547.6 KB
```

```
Grid_df['Demand']=Grid_df['Demand'].interpolate(method='linear')
```

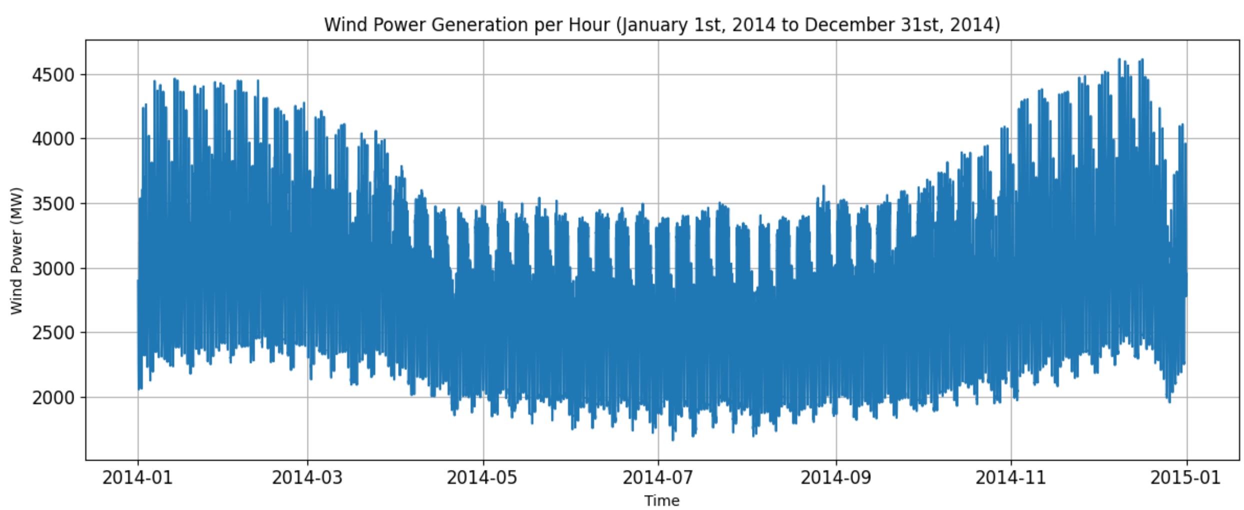
- Changing the Index to DateTime

The dataset was restructured by setting the date-time column as the index. This transformation optimizes time-based analysis and facilitates streamlined data retrieval for specific time periods.

Data analysis

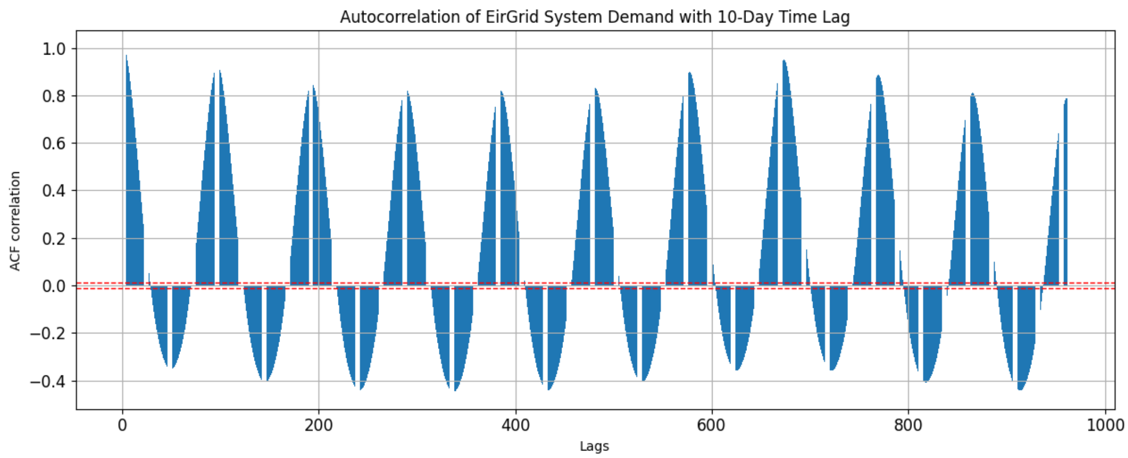
Time Series Plot

In this data analysis process, the time series data were visualized by plotting it on a graph. This graphical representation allowed us to observe the trend and patterns of the dataset over time. By plotting the time series, we gained insights into the fluctuations and changes in the data, providing a clear visualization of its behavior throughout the specified time.



The observed pattern of wind power generation aligns with typical seasonal weather effects. From March to July, the slight decline in wind generation may coincide with the transition from spring to summer, characterized by more stable weather conditions and potentially lighter winds. As summer progresses into July, wind speeds may increase again, possibly due to the onset of summer storms or stronger wind patterns associated with atmospheric changes. This uptick in wind generation continues until December, mirroring the seasonal transition into fall and winter months, which typically bring stronger winds and more variable weather patterns.

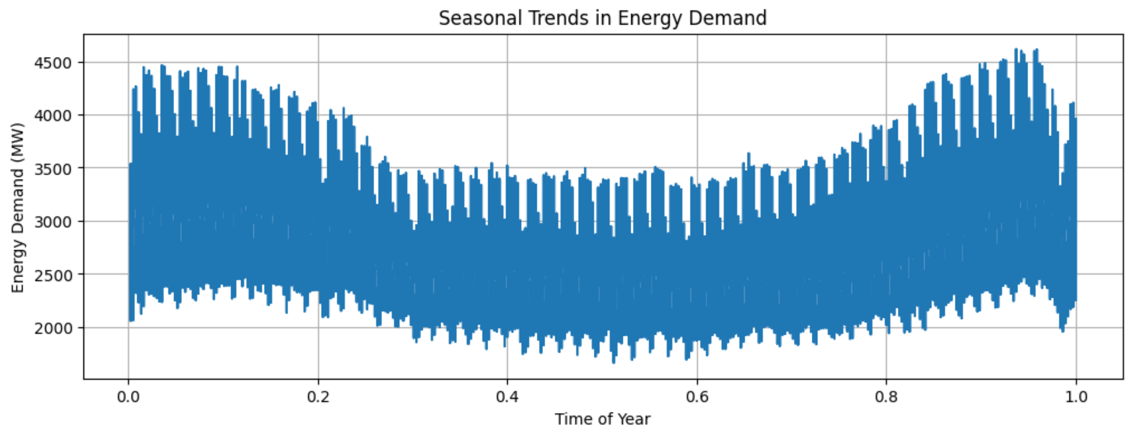
Plotting the Autocorrelation Series



The presence of a repeating pattern at a 10-day lag interval, equivalent to 960 periods in our dataset, suggests significant temporal dependency within the wind power generation data. This observation indicates consistent behavior over this specific lag period, implying a strong correlation between values at different time points. Additionally, all peaks in the autocorrelation function surpassing the significance lines further validate the existence of statistically significant correlations, highlighting the importance of considering time-based relationships in our analysis for improved forecasting accuracy and understanding of wind power dynamics.

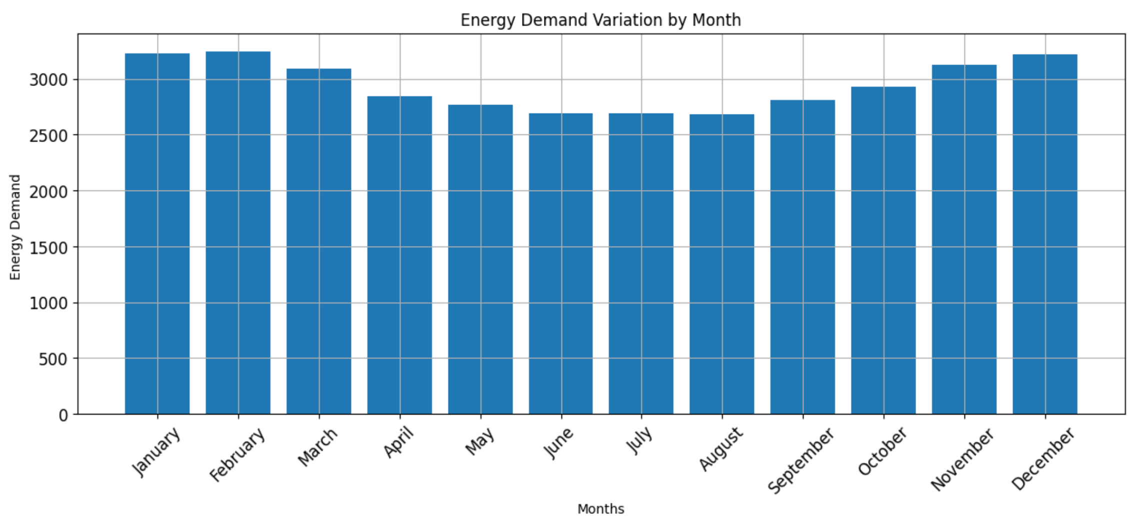
Exploring Seasonal Trends in Wind Power Generation.

In this analysis, the seasonal variability of wind power generation throughout the year was investigated. The focus is on understanding how wind energy production fluctuates over time, represented by the day of the year.



The analysis reveals clear seasonal patterns in wind power generation, with fluctuations corresponding to different seasons. Plotting the time series of wind power generation against the day of the year highlights peaks and valleys that align with expected seasonal variations.

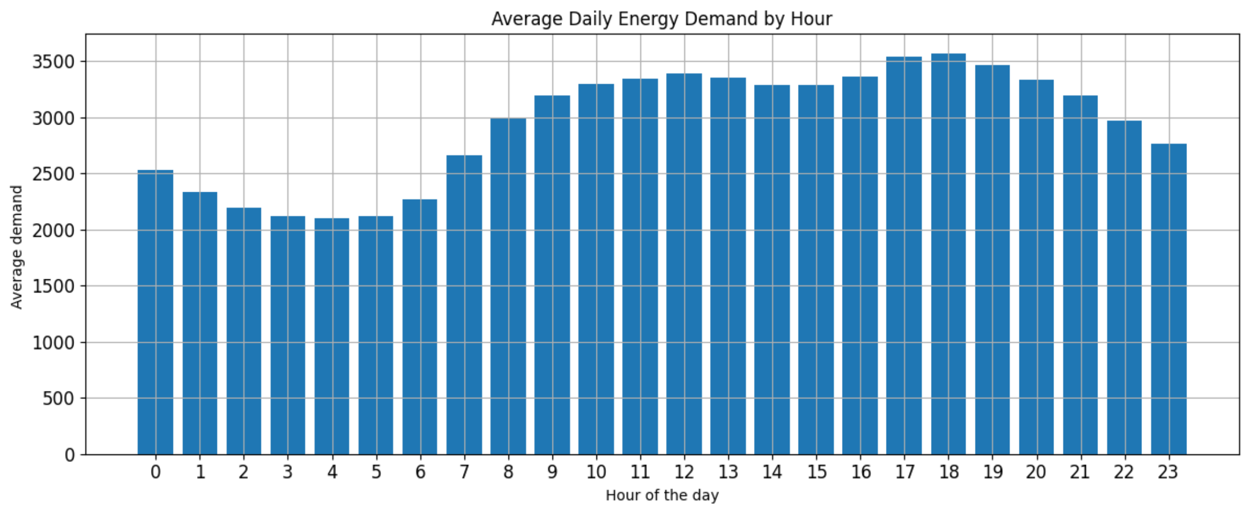
Visualizing Monthly Power generation



Throughout the year, wind power generation exhibits distinct seasonal patterns. In the early months, particularly January and February, wind power production remains consistently high, indicating sustained output. As spring approaches, there is a slight dip in generation, though overall stability persists. During the summer months, from June to August, wind power generation remains steady. However, as the year progresses into the fall and winter, there is a gradual increase in generation, attributed to stronger winds associated with seasonal weather changes. This rise in wind power production towards the end of the year underscores the impact of seasonal weather patterns on renewable energy generation.

Average daily energy demand by hour.

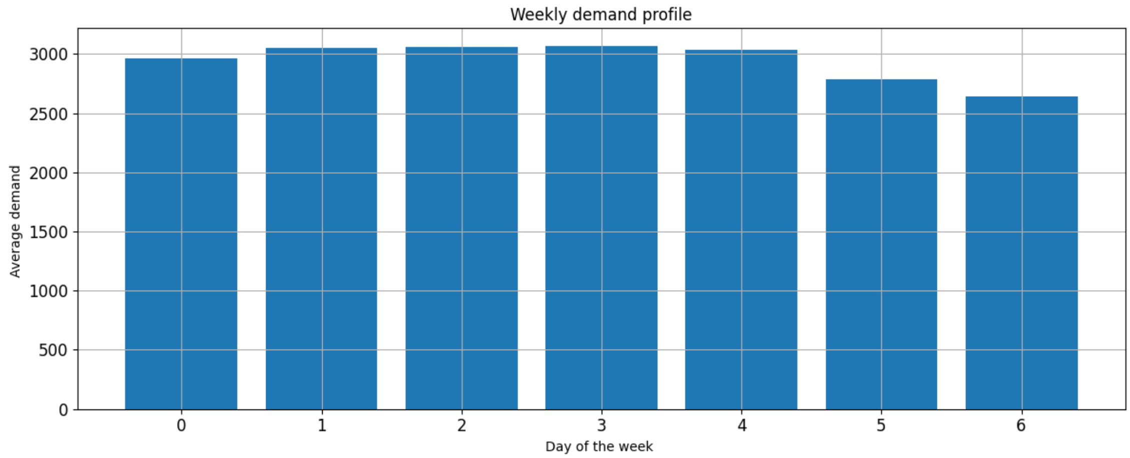
The daily profile of energy generation provides valuable insights into patterns of energy demand throughout the day. By examining the average energy demand across different hours, analysts can identify peak demand periods, periods of low demand, and any recurring trends or fluctuations.



The above graph illustrates the variations in wind power generation throughout the day. The graph reveals that wind power generation experiences spikes during the evening hours, particularly between 5 PM and 7 PM, in contrast to other times of the day. Moreover, it shows a gradual decrease in wind power generation during the morning hours, followed by a rise starting at 5 AM. These fluctuations likely coincide with daily routines, suggesting that wind power generation aligns with patterns of human activity throughout the day.

Weekly Energy generation profile

The objective of this analysis is to examine the patterns of energy generation on a weekly basis, aiming to gain insights into the fluctuations in demand over time. Understanding these patterns is essential for effective energy management and resource planning, as it allows energy providers to anticipate and respond to changes in demand more efficiently.



Illustrates the variation in energy generation throughout the week. Interestingly, there is a significant contrast in generation between weekends and weekdays. This could be attributed to reduced industrial and commercial activities on weekends, resulting in lower energy demand.

Hypothesis testing

In the preceding analysis of energy generation fluctuations during weekends and weekdays, a hypothesis was formulated with the objective of demonstrating the similarity in energy generation during these periods. Specifically, the null hypothesis proposed that there is no difference in energy generation between weekends and weekdays, while the alternative hypothesis suggested otherwise.

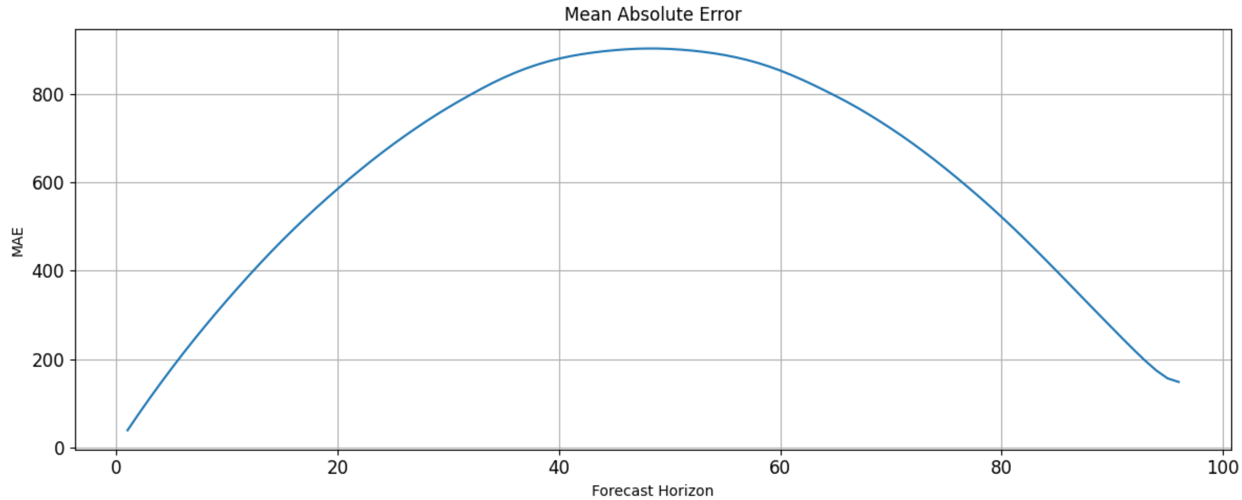
The P-value obtained is $4.55*10^{-26}$. As this p-value is significantly below the conventional significance level of 0.05, the null hypothesis was rejected, indicating a statistically significant difference in energy generation between weekends and weekdays.

Modeling and Forecasting

Persistence Forecasting

Persistence benchmarking involves using a simple forecasting method where the forecast for a future period is based on the observed value from the current period. In other words, it assumes that future values will remain the same as the most recent observed value.

The task in this section was to evaluate the persistence benchmark over the different forecast horizons starting from one hour to one day on the second half of the data set. It was done by calculating MAE for each horizon and the graph below indicates that the MAE has a curvilinear relationship with the horizon taken.



Conclusion

In conclusion, the analysis of wind power generation data for the year 2014 yielded valuable insights into seasonal patterns, temporal dependencies, and weekly energy demand cycles. The findings highlight the importance of considering weather conditions, time-based relationships, and weekday variations in energy generation for effective renewable energy planning and management. Moving forward, these insights can inform decision-making processes and facilitate the development of robust forecasting models to optimize renewable energy utilization and resource allocation.