FORECASTING THE FUTURE: A GLOBAL SHIFT TOWARDS RENEWABLES

REPORT

Team A-05:

Name	Student ID
Sreemayi Dontha	0821928
Varun Kumar Boggula	0824044

Table of Contents

1. Introduction

1.1. Background and Motivation

The world today faces unprecedented environmental and economic challenges. The need for renewable energy transformation to replace fossil fuels stands as the top priority since global warming meets resource scarcity with geopolitical uncertainties (Gielen et al., 2019). Governments together with industry leaders and researchers now put their attention on renewable energy because it represents an effective climate change mitigation method alongside sustainable development needs (Pietzcker et al., 2014). Both the International Energy Agency (IEA) and International Renewable Energy Agency (IRENA) confirm that reaching 11,000 GW of renewable energy worldwide by 2030 represents the best method to stay under the 1.5°C limit (International Energy Agency, 2023). The current study analyses historical energy data of countries across the globe and forecasting the future renewable energy generation.

1.2. Project Objectives

Develop an interactive interface which displays renewable together with non-renewable power generation data per region for every nation. The developed tool enables stakeholders to track energy transitions by presenting dynamic energy trend visuals that enhance their analytical capabilities.

The research investigates the historical changes in renewable energy generation (Solar and Wind and Hydro and Bioenergy) between 2000 and 2023 across 97 different nations. The analysis monitors present renewable goals compared to previous performance records in order to show each nation's collective success and obstacles.

The forecasting processes employ Holt-Winters Exponential Smoothing as a time-series method for predicting future production levels to determine target achievement rates of countries until 2030. An analysis of clustering enables researchers to form distinct groupings of nations by integrating years of energy production statistics along with renewable resource percentages and fossil fuel utilization patterns in order to examine different international energy management approaches.

We will use both statistical tests and validation metrics that include T-tests and ANOVA as well as Pearson correlation alongside silhouette scores and Calinski-Harabasz Index and Davies-Bouldin Index to evaluate our model's effectiveness.

1.3. Purpose and Importance

A revolutionary change occurs in the energy industry. Global agreements (like Paris Agreement) and the global competition towards reducing greenhouse emissions would create immense pressure on the countries to frame the policies without compromising economic growth factor (Bataille et al., 2018). The implementation of renewable energy follows different paths for each country because individual nations experience influences from economic principles and technological strengths and natural resources and policy structures (Zhang et al., 2010).

Multiple significant reasons call for a thorough comprehension of renewable energy generation mechanisms. Investors together with financial institutions use renewable energy production forecasts to recognize profitable market opportunities for technology investments (Ameli & Kammen, 2012). The efficient analysis of the renewable energy trends aid in tracking the CO2 emissions and promote the renewable energy project developments. The future outlook of energy production supplies grid operators with necessary information for developing essential strategies regarding infrastructure planning and storage solutions and grid integration systems (Knezović et al., 2021).

2. Literature Review

In the recent past, the urgency of transition from non- renewable energy to the renewable ways has grabbed the attention of the academic researchers. The scholars have tried implementing number of the AI models to predict the future of renewable energy production, evaluate the impact of policies put out governments, as well as to understand the climatic impact on the evolution of renewable energy. Researchers have also spent adequate amount of time on understanding the trends of renewable energy generation over the years and help the policy makers to take appropriate actions promptly.

Models that are robust to rapid changes and trends are the most important ones in the renewable energy forecasting. Hyndman and Athanasopoulos (2021) have developed one such model which implements the time- series forecasting of energy data using the Exponential smoothing methods, which is beneficial in capturing the complex trends. Earlier, Chatfield, Koehler, Ord, and Snyder (2001) have developed a more refined exponential smoothing model which could capture structural shifts and complex trends in the data and performing significantly well at the same time. For instance, Li et al. (2024) have implemented the forecasting model using the data from wind power, which was incorporated with various climatic data correlated with supply- demand points, and concluded that it is high time to develop a more robust and advanced forecasting model which could perform exceptionally well when trained on energy data.

Understanding the synergies between different renewable energies would be handy for policy makers in framing the policies for the future. Regression is the most common and easy way to understand it. In the year 2019, the National Institute of Advanced Industrial Science and Technology (NIAIST) has released a report of their findings. It is more focused on the importance of statistical comparisons between the renewable energies in order to assess the efficiency. These kind of reports or findings emphasize the need of regression analysis in researching the inter relationship between the renewables and help finding the optimal ways for their productions.

In any historical data analysis, i.e., renewable energy (in our case), clustering analysis plays a crucial role in segmenting the data points w.r.t their trends and behaviours. For instance, in the year 2023, Caromba, Schutte, and van Laar have emphasized the role of clustering techniques by applying it to segment the deep- level mining operations based on their energy performances. As a further support, Viera et al. (2023) employed online clustering technique to detect the underlying data patterns in energy consumption. In addition, the traditional clustering techniques (like K-means, Hierarchical) come in very helpful for the governments or the international unions to understand the similarities between various regions and frame the policies which target the development of renewable energy sources.

For any task to be accomplished, the actions are more important in achieving the target set. In other words, greenhouses could not be achieved by the nations by just setting up the target goals, they also need to make policies and take actions to reach it. The same point of interest is illustrated in the REN21's Renewables in Cities: 2021 Global Status Report (Ranalder et al., 2021). The findings from the report suggests that there is a significant gap between the goals set and the actions taken to achieve them. Findings by Hernandez and Morales (2020) supported the above statement i.e., they have implemented a model to compare the future targets and the till data implementation trends of renewable energy. The main motto behind these studies is to enlighten the nations to know where they lag currently and help framing the strategies to achieve the target renewable energy that they have set for themselves.

The study conducted emphasizes on all the above discussed analytical techniques to gain insights on the renewable energy. The framework developed evaluates the historical trends, future forecasting and enhances the chances of policy makers to take effective decisions in order to achieve the targets set.

3. Data Set

3.1. Data Sources and Description

The research required two main datasets for a thorough analysis and prediction of renewable energy patterns:

3.1.1 Historical Energy Generation Data (2000–2023)

Over 24 years the dataset includes energy generation statistics from 97 nations during each year. The database contains complete information about renewable alongside non-renewable energy activities.

The dataset contains four renewable energy sources which include solar power and wind power alongside hydroelectric and biomaterials. The variables measure operational capabilities through units which use gigawatts (GW). The data incorporates information on Coal together with Gas and Nuclear energy and additional fossil fuel sources. Alongside, gathered data points function as assessment tools through their presentation of conventional energy production statistics. Dataset contains metadata information which consists of country codes together with area classifications and continental listings as well as OECD/G7/G20/EU membership tags. Additionally, the built-in attributes simplify both data evaluation methods and classification analysis techniques. The whole data spans from 2000 to 2023 yearly and enables extensive trend assessment with seasonality analysis and future forecasting possibilities. Moreover, the dataset includes three numerical value tables that show energy generation metrics with their associated annual changes and time-based percentage variations. The designated database structure enables users to study both precise energy production developments with year-by-year comparison data as well as long-term production development evaluations.

3.1.2 2030 Renewable Energy Targets

The dataset provides run-time renewable energy capacity targets that individual nations have specified for 2030. We analyse the target values exclusively from the individual sheet even though the dataset includes multiple sheets. Also, the entry per country displays their renewable capacity target which breaks down into energy types including Solar, Wind, Hydro and Bioenergy. Additionally, the targets derive their basis from national energy policies and strategic plans which provide benchmarks to evaluate historical performance. On top of that. data appears in a tabular format with distinctive identifiers which include country codes and country names to enable ease of merging it with historical information.

3.2 Data Preparation and Exploration

Analytical evaluation required data cleaning together with pre-processing procedures for achieving reliable and consistent results. Multiple essential steps included in the analysis method included

3.2.1. Data Filtering

Country Selection: The analysis retained member countries from key organizations EU, G7, G20 and OECD as these countries maintain established energy policies. The filtering process achieved two objectives: it eliminated data variability at the hands of inconsistent reporting practices while narrowing the analysis toward countries which yielded reliable energy data.

3.2.2. Handling Missing Values

Missing Data Imputation: We substituted missing data points in essential numerical fields that include energy values and YoY (Year-over-Year) changes with zero values because precise forecasting and clustering depends on high data quality. The step played a vital role to stop data distortions from occurring during statistical processing.

Consistency Checks: The datasets underwent technical tests to confirm that Year was recognized as a numerical variable while all Value columns displayed numerical values. The consistent structure across the analysis flow benefited from this step.

3.2.3. Aggregation and Feature Engineering

Aggregation by Country and Year: To facilitate analysis of the historical data the researchers aggregated it into country-year totals. The data included computations for yearly national energy generation which distinguished renewable from non-renewable energy types.

Derivation of New Metrics: Firstly, analysis calculated multiple performance indicators known as key performance indicators (KPI). Secondly, the total renewable energy output includes all renewable energy resources such as solar, wind, hydro and bioenergy for each country during the year. Thirdly, each country-year received a determination of their non-renewable energy totals by including Coal, Gas, Nuclear and Other Fossil fuel usages. Also developed a renewable share indicator represents the ratio of renewable against total energy production which shows how much a nation depends on renewable energy. Additionally, the inverse relationship between fossil-based energy consumption and renewable transition progress can be measured through fossil dependence since it shows the percentage of non-renewable (fossil) forms against the total energy mix.

3.2.4. Pivoting and Merging

Reshaping Data

New clustering methods required a time-series structure which was achieved through data pivoting. The data organization step transformed the dataset to show each country-year combination as unique rows that contained individual energy type columns.

Dataset Integration

Common country identifiers within the datasets were used for merging purposes to enable historical trend analysis alongside future target assessments. In order to evaluate country performance against their 2030 renewable energy targets this merging procedure plays an essential role.

3.2.5. Exploratory Data Analysis (EDA)

Visualization

An initial EDA required different visual tools to analyse the data. The histograms allowed researchers to display how energy values distributed themselves throughout multiple countries and time periods. Scatter Plots illustrated the energy type relationships in order to detect potential correlations between Solar and Wind energy. The analysis used line plots together with time-series charts to display renewable and non-renewable energy delivery patterns throughout time.

Preliminary Insights

The Energy Data Analysis showed that energy outputs demonstrated substantial variations throughout the different geographical areas together with temporal period changes. The study identified statistical anomalies and established trends to direct predictive model creation as well as subsequent statistical analysis.

All product and research stages with these datasets create a reliable base that enables further statistical evaluation using predictive models. Our work on data quality as well as feature extraction allows us to deliver a solid examination of renewable energy patterns that serves present performance measurements while generating future prediction capabilities.

4. Methodology

The research approach contains multiple linked analytical methods which build a complete picture of renewable energy patterns. The research relies on statistical hypothesis testing together with predictive time-series modelling and clustering analysis methods. Researchers used various methods which addressed different research questions while generating strong observations for present behaviour along with projected developments.

4.1 BI Dashboard

An interactive dashboard which served to improve data accessibility along with visualization capabilities was created. Through the dashboard users can examine renewable together with non-renewable energy generation data from countries and regions which are separated into different energy source categories. The dashboard presents a convenient management system through dynamic charts combined with filtering tools and geographic maps which enables users to monitor energy patterns and make knowledge-based choices. The application functions as a hands-on instrument that enables policymakers together with researchers along with investors to analyse energy climate profiles dynamically and discover transition types while monitoring clean energy development goals.

4.2 Statistical Analysis and Hypothesis Testing

Statistical testing methodology allowed for proper assessment of renewable energy production differences along with relationship analysis in the data.

4.2.1 T-Tests

Objective: To compare the means of renewable energy production across different regions.

Method: Eligible for international energy data characteristics we applied Welch's T-test as this method provides effective results when group variances are unequal.

Implementation: The analysis of Bioenergy production employed two combined groups based on region between North American countries and Asian nations. The calculated p-value together with T-statistic helped determine if the identified differences between groups held statistically important significance.

4.2.2 One- way ANOVA (Analysis of Variance)

Objective: Statistical analysis provides information on whether differences exist significantly between energy production between multiple regions.

Method: The research team performed one-way ANOVA tests to analyse Solar, Wind, Hydro and Bioenergy production amounts. We examined the comparison between three or more group means to identify significant differences between any of the groups.

Implementation: ANOVA received dataset evaluation after the researcher applied region-based groupings from the original dataset design. The analysis measured inter-group variability through the calculated F-statistic together with its corresponding p-value.

4.2.3 Correlation Analysis

Objective: The investigation evaluated the extent and orientation of linearity that exists between different renewable energy sources.

Method: The application of Pearson correlation analysis delivered insights about Solar energy and Wind energy manufacturing relations.

Implementation: The analysis revealed an almost perfect positive relationship through the computed coefficient which demonstrated that countries with stronger Solar power tend to establish Wind power operations.

4.2.4 Linear Regression

Objective: To predict one form of renewable energy production using another as a predictor.

Method: Simple linear regression models were developed where, for instance, Wind energy production was predicted based on Solar energy production.

Implementation: An evaluation of the model performance relied on both the R-squared statistics to measure predictor-based dependent variable variation and the slope coefficient's associated p-value. The predictive relationship demonstrated strong power because of these R-squared values exceeding the threshold value.

4.3 Predictive Modelling

Time-series forecasting played a critical role in forecasting future renewable energy production when assessing 2030 targets of countries.

4.3.1 Data Splitting

Objective: To ensure model validity and avoid overfitting.

Method: The historical data was split into two segments

Training Set: Data from 2000 to 2020 was used to fit the model.

Validation Set: Data from 2020 onward was held out for validation.

Test Set: Target data for 2030 for testing.

Rationale
The data separation enabled the model to receive comprehensive historical training data followed by unvisited target prediction evaluation.

4.3.2 Model Fitting Using Holt-Winters Exponential Smoothing

Objective
The method needs to detect seasonal patterns and trends that exist within renewable energy production data.

Method
The Holt-Winters approach proved suitable because it enables modelling of both level components and trend patterns and seasonal variations.

Implementation
The model received its parameters through calibration of the training data and validation data before generating forecasts for the 2030 target values.

Forecasting and Evaluation

Forecasts were generated for the year 2030. Model performance was evaluated using:

Mean Absolute Error (MAE): Measures average absolute prediction error.

Root Mean Squared Error (RMSE): Provides insight into error magnitude by penalizing larger errors.

R^2 (Coefficient of Determination): Indicates the proportion of variance in the test data that is predictable from the model.

4.4 Clustering Analysis

A clustering analysis method categorizes different nations according to their current energy conditions because it reveals the distinct approaches various countries employ to manage energy resources.

4.4.1 Feature Extraction

Feature Extraction:

Objective: Method builders extracted a strong feature collection that conveyed the complete energy characteristics of each nation.

Method: The analysis calculated important features based on historical data records.

Average Renewable Capacity: Mean renewable energy generation over the years. The renewable energy production trend can be demonstrated using linear regression applied on time series data to obtain its slope value.

Non-Renewable Capacity: Mean production of fossil fuels and other non-renewable sources. The Renewable Share indicates the percentage relation between renewable energy and entire energy production in a country.

Fossil Dependence: Complementary metric to Renewable Share.

Implementation: The data needed a pivot transformation to create time-series records by each country. The team generated feature values through aggregation procedures combined with trend measurement systems.

4.4.2 K-Means Clustering

Objective: To group countries into clusters with similar energy profiles.

Method: The K-Means clustering process operated on the standardized feature variables.

Steps:

Standardization: The implementation of StandardScaler normalization enabled all features to obtain equivalent weights for distance computation.

Cluster Implementation (Three-Cluster Approach): A refined analysis which groups countries into three distinct clusters named “**Fossil-Dependent**” and “**Transitioning**” and “**Renewable-Focused**” clusters.

4.4.3 Cluster Validation

Objective: To ensure that the clusters produced are robust and well-separated.

Method: Several metrics were used:

Silhouette Score: Measures the cohesion within clusters and separation between clusters. The clustering results indicated moderate distinction between the groups.

Calinski-Harabasz Index: Clusters detect as compact groups.

Davies-Bouldin Index: The clusters display low intra-cluster variance.

Cross-Validation: The 5-fold cross-validation process evaluated the stability of cluster solutions by checking their assignment robustness to different data splits.

5. Results

This segment demonstrates the results from our evaluations along with the interactive dashboard and the assessment of statistical methods and predictive modelling along with clustering methods' performance. These findings reveal important information about fundamental energy production operations for future analytical and government policy considerations.

5.1 Interactive BI Dashboard

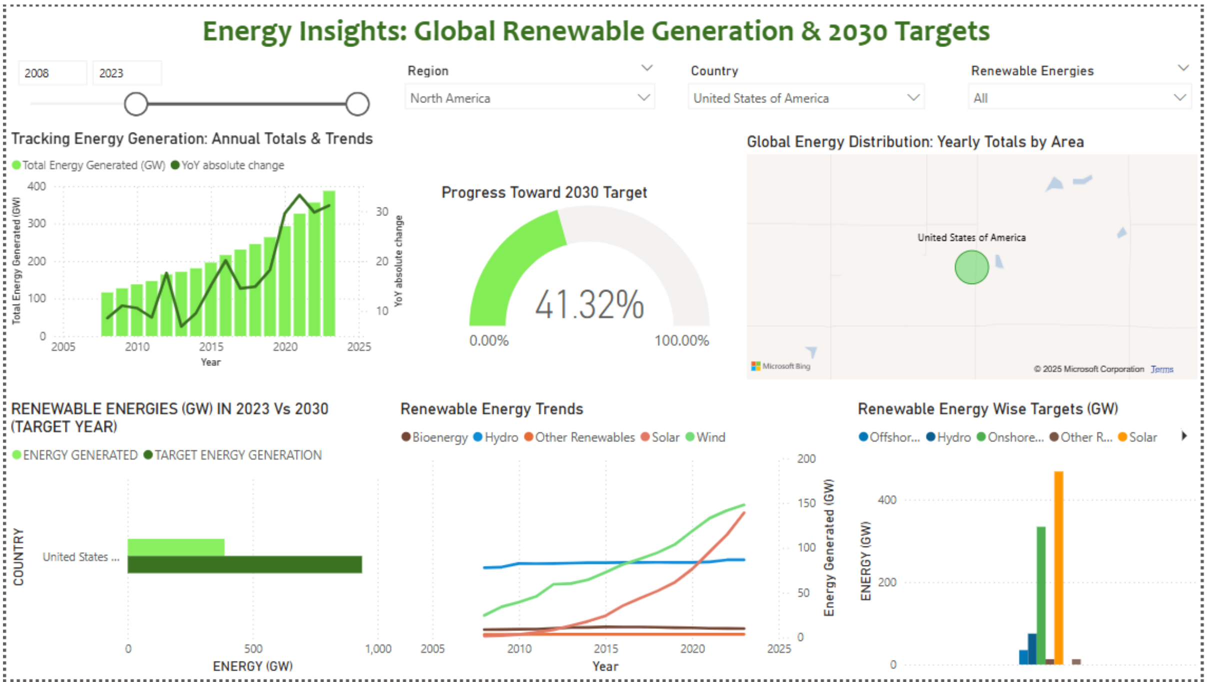

Fig 1: Interactive BI Dashboard

5.2. Statistical Test Outcomes

5.2.1. T-Test Analysis (North America vs. Asia)

The analysis used Welch's T test to determine whether renewable energy production means between North America and Asia differed for individual energy types.

Bioenergy:
T = 0.332,p = 0.7409.
Interpretation: The results indicate that North America and Asia share equivalent average bioenergy capacity levels between 2000 and 2023 since the p value exceeds 0.05 level. On average the production levels of bioenergy match between North America and Asia.

Solar:
T = –2.179,p = 0.03397
Interpretation: The outcome demonstrates that Asia generates more solar capacity than North America (p value <0.05). The analysis results through the negative T statistic demonstrate Asia exceeds North America in solar energy generation by several megawatts per day.

Wind:
T = –1.999,p = 0.05018
Interpretation: The comparison of wind capacity between Asia and North America shows a slight difference which showed borderline statistical significance at p=0.05. There exists a possible spatial discrepancy in wind power development which barely falls outside the standard 5% threshold for statistical significance.

Hydro:
T = –3.063,p = 0.00368
Interpretation: The analysis demonstrates Asia possesses significantly greater hydroelectric capacity (25-40 GW) than North America thus indicating its larger river ecosystems together with its investments in massive hydroelectric plants.

5.2.2. ANOVA Results (Across All Regions)

One-way ANOVA tests were conducted to assess whether mean production differs among all defined regions (e.g., Europe, Latin America, Africa, etc.):

Energy Type	F-Statistic	p-Value	Significance
Bioenergy	6.648	0.0019	p < 0.01
Solar	4.448	0.0138	p < 0.05
Wind	5.753	0.0042	p < 0.01
Hydro	6.350	0.0024	p < 0.01

Table 1: One- way ANOVA results of 4 different Renewable Energy types

Interpretation: Each F statistic demonstrates high significance which establishes that at least one region maintains different mean renewable capacity metrics than the others when examining each energy type. The geographical location and climate influence hydroelectric power generation to the point where it demonstrates significant regional differences.

Overall Implications:

Bioenergy demonstrates similar patterns of distribution between North America and Asia although its global patterns become distinctly different when including all regions which indicates European and Latin American bioenergy uses drive global changes.

The comparison between NA and Asia and multi region testing of Solar and Hydro shows distinct findings which emphasize geographic solar irradiance and hydrological resources influence.

Wind demonstrates significant differences when analysing multiple regions despite showing minimal pair-wise fluctuations thus indicating that North America shares similarities with Asia but other areas such as Europe may differ largely.

Geographic locations alongside climatic conditions along with developmental characteristics all influence the worldwide renewable energy capacity distribution according to quantitative research results.

5.2.3. Correlation and Regression Analysis

Pearson Correlation (Solar vs. Wind)
Coefficient: 0.9503
P-Value: 4.17×10^{-59}

Interpretation
The statistical relationship between Wind and Solar capacities explains more than 90% of Wind power variance. Countries pursuing substantial Solar infrastructure development simultaneously increase their Wind capacity installations because of combined strategic planning and investment motivations.

Linear Regression (Wind ~ Solar)
R^2: 0.9031
P-Value: 4.17×10^{-59}

Interpretation
The model predicts that various factors related to Solar capacity account for 90.3 % of Wind capacity development. The excellent relationship between Solar and Wind capacities can be validated through the high R^2 value combined with a significant slope in our data which emphasizes the complementary nature of these renewables.

5.3. Predictive Modelling Performance

We applied Holt Winters Exponential Smoothing models to both energy types of each country across two decades using the metrics MAE, RMSE with R^2 for evaluation during 2021–2023. Below we highlight key findings:

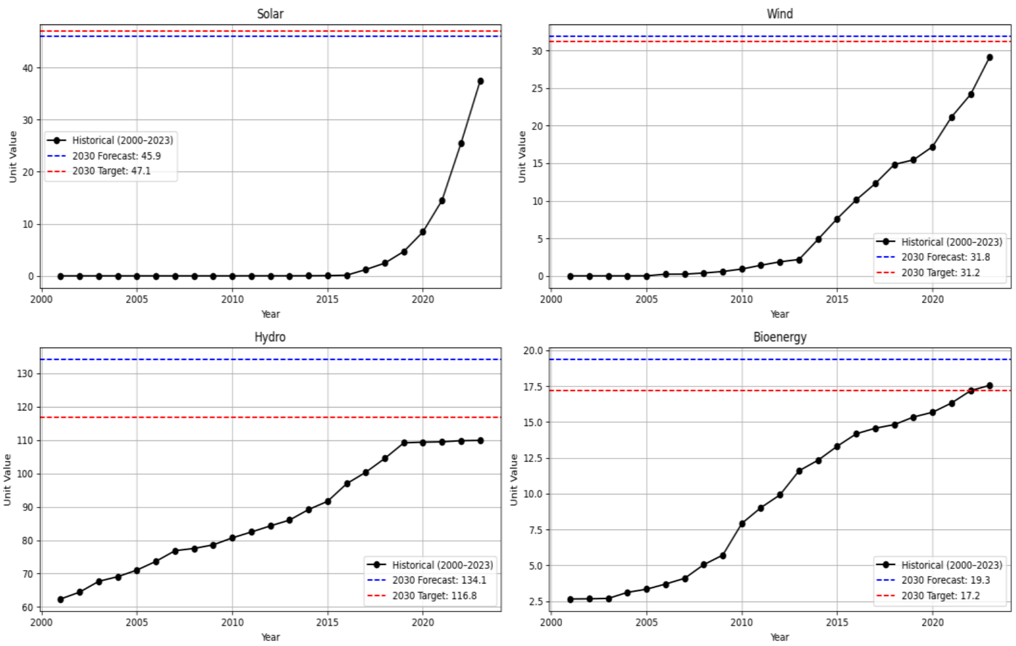

Fig 2: Renewable Energy (2000-2023) and 2030 Forecast vs Target – BRAZIL

Brazil (BRA)

Bioenergy: MAE = 0.61 GW, RMSE = 0.65 GW, $R^2 = -0.60$

Solar: MAE = 9.87 GW, RMSE = 11.72 GW, $R^2 = -0.56$

Wind: MAE = 4.73 GW, RMSE = 5.18 GW, $R^2 = -1.48$

Hydro: MAE = 4.57 GW, RMSE = 4.93 GW, $R^2 = -683.42$

Interpretation: The test data based mean prediction outperformed all models for each energy source because all R^2 scores were negative especially for Hydro which could not properly model the extreme seasonal variations.

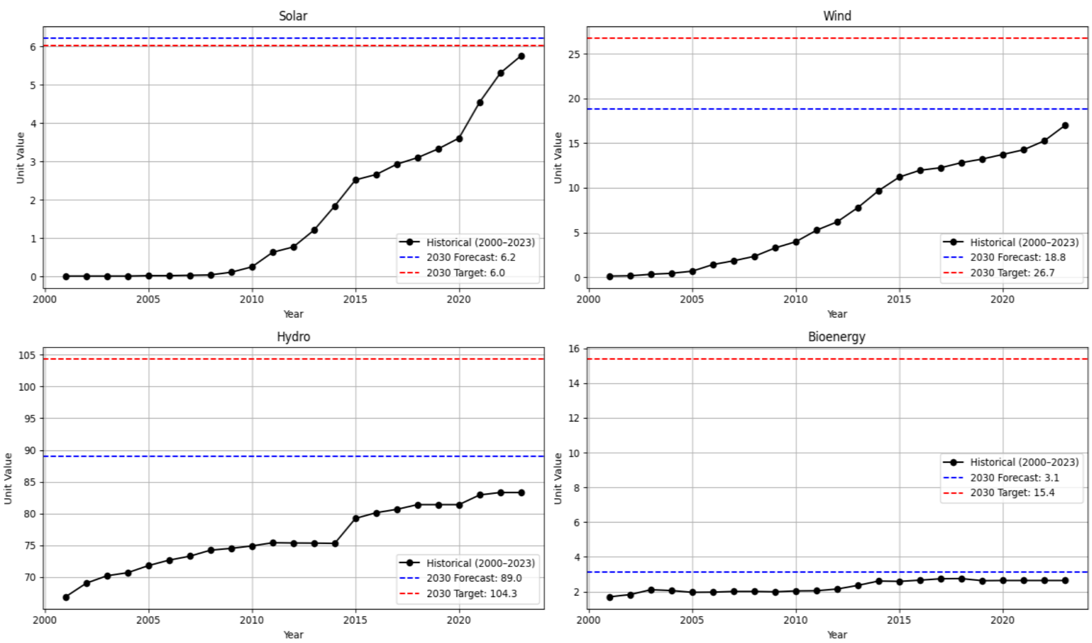

Fig 3: Renewable Energy (2000-2023) and 2030 Forecast vs Target - CANADA

Canada (CAN)

Bioenergy: MAE = 0.10 GW, RMSE = 0.11 GW, R^2 = 0.00
Solar: MAE = 1.08 GW, RMSE = 1.12 GW, R^2 = –3.99
Wind: MAE = 0.75 GW, RMSE = 1.04 GW, R^2 = 0.16
Hydro: MAE = 0.51 GW, RMSE = 0.54 GW, R^2 = –8.13

Interpretation: The predictive efficiency of wind power forecasts exhibits relatively average performance (R^2 = 0.16) yet hydro and solar models generate underwhelming results because Canadian hydroelectric operations remain highly erratic.

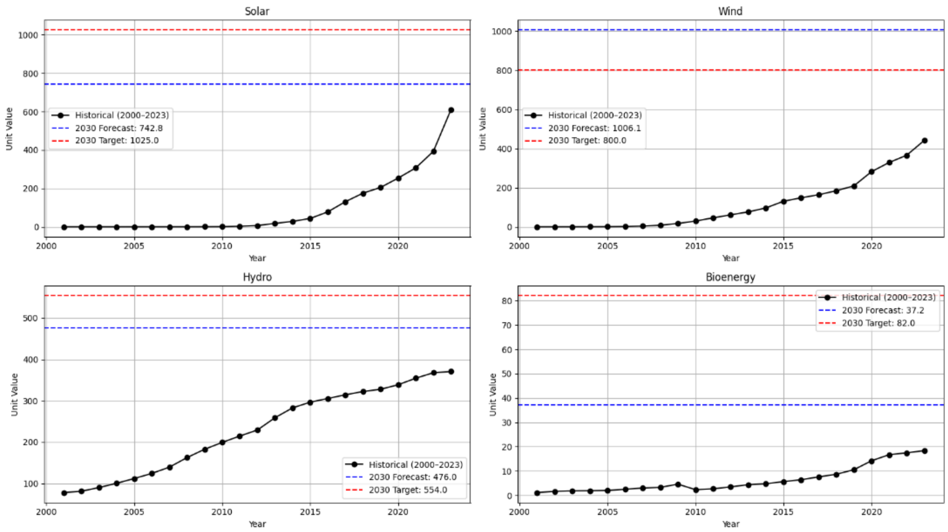

Fig 4: Renewable Energy (2000-2023) and 2030 Forecast vs Target - CHINA

China (CHN)

Bioenergy: MAE = 1.35 GW, RMSE = 1.57 GW, R^2 = –4.31
Solar: MAE = 85.00 GW, RMSE = 123.25 GW, R^2 = 0.065
Wind: MAE = 47.87 GW, RMSE = 50.45 GW, R^2 = –0.15
Hydro: MAE = 4.32 GW, RMSE = 5.57 GW, R^2 = 0.37

Interpretation: Solar power data reveals some predictive ability through R^2 approximately equal to 0.07 yet errors exceed 120 GW because of the swift development of Solar facilities across China. Hydro forecasts are most reliable (R^2 = 0.37).

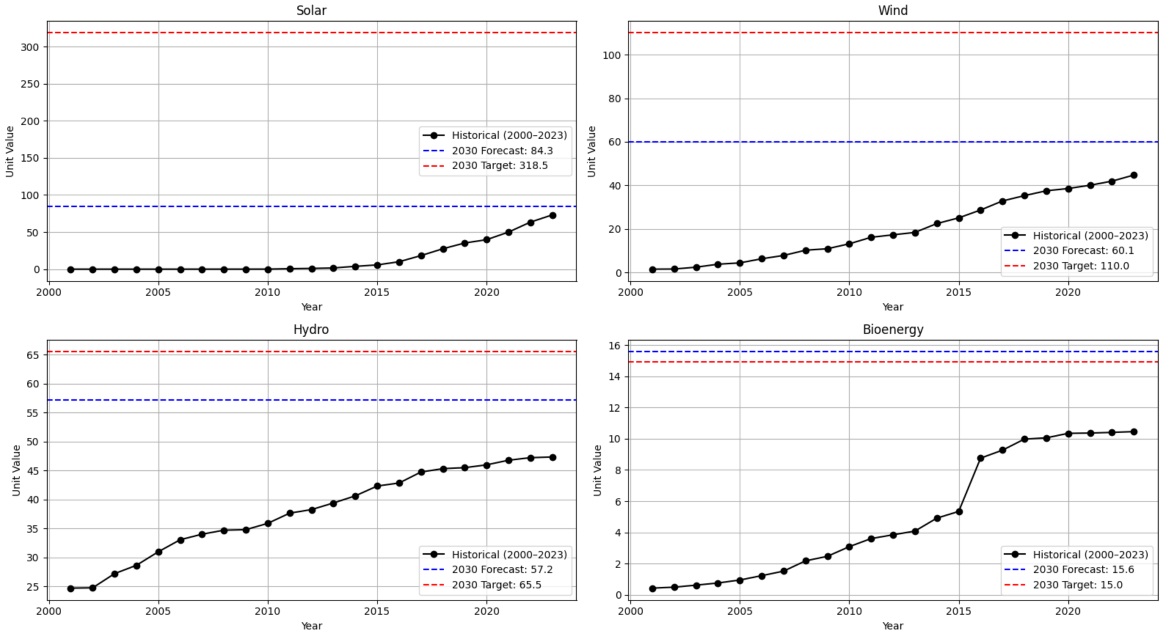

Fig 5: Renewable Energy (2000-2023) and 2030 Forecast vs Target - INDIA

India (IND)

Bioenergy: MAE = 0.98 GW, RMSE = 1.06 GW, R^2 = –822.48

Solar: MAE = 13.52 GW, RMSE = 14.74 GW, R^2 = –1.4

Wind: MAE = 0.61 GW, RMSE = 0.67 GW, R^2 = 0.88

Hydro: MAE = 1.11 GW, RMSE = 1.31 GW, R^2 = –29.46

Interpretation: Wind energy predictions performed exceptionally well in India based on R square value of 0.88 because the forecasts show continuous annual growth. The multivariate patterns from other sources exceed the smoothing model's capability to predict.

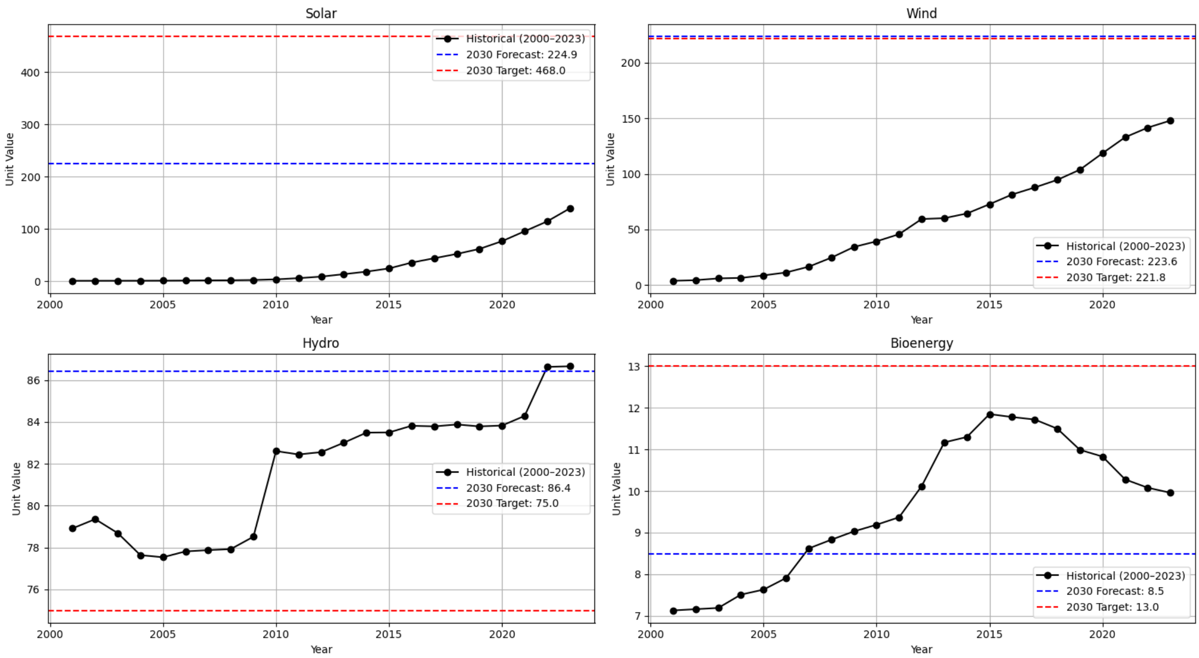

Fig 6: Renewable Energy (2000-2023) and 2030 Forecast vs Target - USA

United States (USA)

Bioenergy: MAE = 0.26 GW, RMSE = 0.26 GW, $R^2 = -2.99$

Solar: MAE = 10.18 GW, RMSE = 11.78 GW, $R^2 = 0.57$

Wind: MAE = 2.67 GW, RMSE = 2.80 GW, $R^2 = 0.79$

Hydro: MAE = 1.51 GW, RMSE = 1.78 GW, $R^2 = -1.56$

Interpretation: The forecast strength for Wind energy ($R^2 = 0.79$) and Solar energy ($R^2 = 0.57$) indicates fairly strong relationships which produce mature growth through incremental development patterns. The R^2 value for Bioenergy and Hydro comes out to be negative thus indicating unpredictable trends exist.

Key Insights:

On-Track: Brazil achieves solar and wind power estimation success at approximately 1 GW off their target while American wind power forecasting stays within 2 GW of the objective.

Under-Forecast: The current trends in Canada's Bioenergy and Wind power alongside China's Solar and Wind capabilities and India's Solar development and USA's Solar systems remain insufficient since they operate below targets by more than 7 GW.

Over-Forecast: The excessive 17 GW amount in Brazil's Hydro power forecast points to recent high-water levels probably contributing to this result.

5.4. Clustering Analysis and Profiles

An analysis using K Means clustering on Fossil Dependence together with Renewable Share and Renewable Trend data generated three distinct categories from 51 countries.

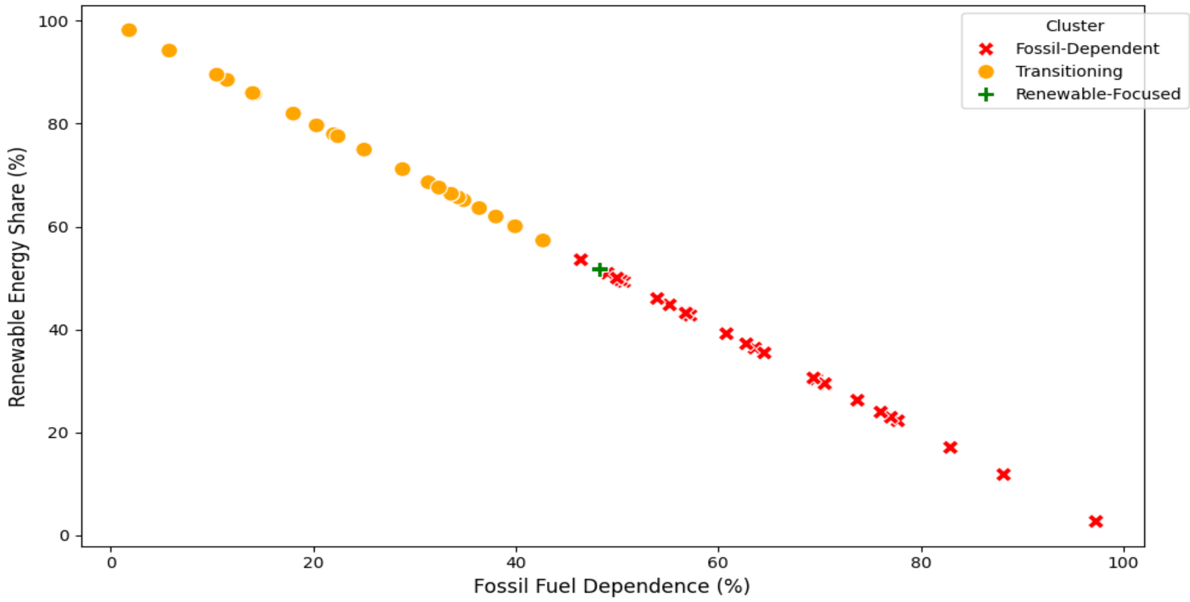

Fig 7: Country Clusters by Energy Profile

Cluster	Count	% of Countries	Avg. Fossil Dependence	Avg. Renewable Share	Avg. Renewable Trend (GW/yr.)
Fossil-Dependent	26	51.0 %	64.5 %	35.5 %	1.58
Transitioning	24	47.1 %	26.4 %	73.6 %	0.99
Renewable-Focused	1	2.0 %	48.2 %	51.8 %	52.39

Table 2: Clustering profiles using KPIs

Interpretation: These countries preserve their traditional dependence on fossil energy resources while their renewable energy development reaches minimal annual growth rates of approximately 1.6 GW/yr. (51% of all countries). And the transitional category comprises fifty-one percent of countries where renewable energy accounts for 73.6% of the power sector yet growth remains below one gigawatt per year. Additionally, China stands out from other nations because its remarkable annual installed capacity expansion in renewable energy reaches 52.4 GW while its renewable energy shares lags behind fossil fuel utilization at 51.8 percent.

5.4.1. Cross-Validation and Clustering Metrics

To validate cluster stability and quality, we used 5-fold cross-validation and three standard indices:

Metric	Value	Interpretation
Silhouette Score (CV average)	0.4494	Moderate separation; clusters are distinguishable but overlap exists.
Final Silhouette Score	0.4433	Consistent with CV—cluster cohesion is acceptable.
Calinski-Harabasz Index	68.23	High—clusters are compact and well separated.
Davies-Bouldin Index	0.5726	Low—clusters exhibit low intra-cluster variance.

Table 3: Performance Evaluation metrics for clustering analysis and their interpretations

Overall Interpretation: The clustering solution proves to be robust but its moderate silhouette scores (~0.44) indicate that certain countries exist within ambiguous cluster transition areas even though it demonstrates good compactness (CH = 68.23) and distinctness (DB = 0.57).

6. Discussion and Conclusion

6.1. Why Are the Results Important?

This research review presents diverse findings about renewable energy transitions across different nations which enable deep understanding of power generation evolution and future projections as well as national preparedness for energy transition goals. These research results hold critical value because they combine separate observations into comprehensive information about national energy policy and infrastructure development strategies.

Renewable energy production demonstrates significant differences between worldwide geographical areas according to findings from this analysis. Statistical hypothesis tests including t-tests and ANOVA demonstrate that particular continents together with sub-regions possess noticeable distinctions in their renewable energy capabilities. A profound difference between renewable energy production exists because of varying geographic settings and economic characteristics together with infrastructure availability. The regions which possess plentiful hydrological resources coupled with longer sun exposure times and powerful government support for clean energy operations demonstrate superior performance levels compared to other areas. Regional dynamics understanding enables stakeholders to optimize their investment decisions by using targeted capacity programs and directing financial and technical resources for better results.

Different power generation strengths across countries indicates that no single approach can achieve renewable energy transition success worldwide. Each country needs particular attention that focuses on its unique capabilities alongside required capabilities and constraints. The approach confirms that successful renewable energy adoption needs specific policies together with incentives for each region and infrastructure strategies that adjust to local needs.

The research data reveals how different renewable energy systems form an interconnected system with solar power dependent on wind power development. The analysis shows that nations investing in either renewable energy type tend to achieve similar investment growth in the other category. The co-development pattern exists because the two renewable techniques share technological domains while jointly using common infrastructure together with mutual economic benefits. Strategic planning can merge both power sources through the regression analysis results which show how one generation output correlates directly with another generation output.

Energy developers and urban planners together with public officials must consider this information for energy development and urban planning purposes. The plan-based unification of solar and wind power allows nations to optimize their infrastructure costs and remove unnecessary operations while strengthening their power grid reliability. A multi-layered energy policy emerges when renewable development matches each other leading to increased collective advantages within the entire power system.

Future trends were examined within the predictive modelling segment using Holt-Winters Exponential Smoothing. The prediction models produced different degrees of precision in their analysis between various countries and energy resources. The model showed success in some instances by generating predictions that precisely matched actual pattern movements. The model demonstrated difficulty in recognizing intricate energy dynamics especially when the energy source displayed non-linear behaviour or irregular growth trends.

The prediction of renewable energy output faces natural challenges because swift market transitions alongside policy changes and capital market shifts and technological breakthroughs cause difficulties. The value of exponential smoothing as a forecasting technique stands strong yet evidence indicates additional complex models with exogenous variable integration could boost the accuracy of predictions. The precision and use of models will improve through real-time data integration in addition to using economic performance indicators and policy initiatives and climate variables.

Among all study parts the clustering analysis emerged as most significant because it divided countries into groups that included the fossil-dependent, transitioning and renewable-focused sectors. This energy segmentation system demonstrates how countries use their resources currently and delivers a strategic guide for constructing future energy interventions.

The fossil-dependent cluster contains nations which heavily depend on conventional energy sources because of their prior investments in fossil infrastructure or lack of funds and geographical boundaries. These countries need substantial policy support together with technical assistance coupled with international financing support to shift into renewable energy solutions. Achieving sustainability principles requires these countries to break free from established obstacles while dedicating resources to fundamental infrastructure.

The transforming group comprises countries which have notably increased their renewable energy utilization yet they experience obstacles concerning the expansion of sustainable power systems as well as energy storage solutions and unified policies. These nations provide essential research sites for evaluating policy tools which can enhance renewable energy penetration rates. These countries require well-planned strategic adjustments instead of major structural modifications to accomplish their continued development.

A tiny exceptional group of nations represents the renewable-oriented cluster which achieves results through aligned policies, technological innovation and investments. These countries function as valuable examples and R&D centres throughout the world. Such exceptional nations deliver critical knowledge and technologies enabling nations in different stages of energy transition to share information for expanding international cooperative practices. This small group has substantial influence on the creation of worldwide rules and principles regarding energy systems.

The study's findings enhance our knowledge about renewable energy forces while creating diverse planning frameworks which work together. The interactive dashboard allows real-time data tracking among decision-makers through its data-focused tool which helps them monitor global trends during active events. The in-depth interaction and detail of this system enables prompt energy management decisions so nations can take action as soon as their goals deviate from the intended path.

The research reveals that renewable energy transformation displays substantial complexity while being strongly related between different regions because of their individual characteristics. The data shows a path toward strategic planning which depends on rigorous information analysis and real-time analysis as well as complete understanding of global energy trends. The research advances knowledge through a combination of theoretical research with practical value for policymakers as well as investors and planners who want to achieve sustainable outcomes in the future.

6.2. Conclusion

The research performs an inclusive data-based assessment of worldwide renewable power generation clustering and future prediction analysis. The approach unifies three analytics approaches to create a robust system that studies and manages energy transition complexities.

This research presents its main achievement through multidimensional data analysis methods. The study applies multiple analytical tools which analyse macro-level regional inequalities and micro-level technology related systems within the same framework. The study introduces predictive models although these models need further improvement to reveal foundational information about future energy sector developments. Through clustering analysis stakeholders gain the ability to categorize countries beyond geographical or economic boundaries because it shows their real energy behaviour and future outlook.

These study results create consequences which affect both academic institutions and other sectors. Policy-makers apply research findings to set priorities among intervention strategies while delivering funds and developing regulatory policies that fit individual national requirements. Renewable energy investors alongside developers receive valuable information about profitable expansion areas as well as methods to maximize technology collaborations. International organizations leverage clustering outcomes to discover ways for regional knowledge sharing together with business opportunities.

The study presents an operational strategy that enables rapid growth of renewable energy across the global landscape. The study demonstrates why organizations need strategic connections between energy investment plans and resource development strategies. These analytical tools and frameworks become essential because the world is speeding up its pursuit of sustainability targets for 2030.

7. References

[1] Ameli, N., & Kammen, D. M. (2012). Clean energy deployment: addressing financing cost. *Environmental Research Letters*, 7(3), 034008. https://doi.org/10.1088/1748-9326/7/3/034008

[2] Bataille, C., Åhman, M., Neuhoff, K., Nilsson, L. J., Fischedick, M., Lechtenböhmer, S., ... & van Zyl, A. (2018). A review of technology and policy deep decarbonization pathway options for making energy-intensive industry production consistent with the Paris Agreement. *Journal of Cleaner Production*, 187, 960-973. https://doi.org/10.1016/j.jclepro.2018.03.107

[3] Gielen, D., Boshell, F., Saygin, D., Bazilian, M. D., Wagner, N., & Gorini, R. (2019). The role of renewable energy in the global energy transformation. *Energy Strategy Reviews*, 24, 38-50. https://doi.org/10.1016/j.esr.2019.01.006

[4] International Energy Agency. (2023). World Energy Outlook 2023. https://www.iea.org/reports/world-energy-outlook-2023

[5] Knezović, K., Marinakis, A., Evrenosoglu, C. Y., & Oudalov, A. (2021). Role of grid and bulk storage in the integration of variable renewable energy resources: Framework for optimal operation-driven multi-period infrastructure planning. *Energy*, 226, 120378. https://doi.org/10.1016/j.energy.2021.120378

[6] Pietzcker, R. C., Luderer, G., Kriegler, E., Schaeffer, M., Krey, V., & Riahi, K. (2014). Energy system transformations for limiting end-of-century warming to below 1.5 °C. *Nature Climate Change*, 4(6), 494-501. https://doi.org/10.1038/nclimate2248

[7] Zhang, X., Ruoshui, W., Molin, H., & Martinot, E. (2010). A study of the role played by renewable energies in China's sustainable energy supply. *Energy*, 35(11), 4392-4399. https://doi.org/10.1016/j.energy.2009.05.030

[8] Caromba, C., Schutte, C., & van Laar, J. (2023). Application of clustering techniques for improved energy benchmarking on deep-level mines. *Energies*, *16*(19), 6879. https://doi.org/10.3390/en16196879

[9] Chatfield, C., Koehler, A. B., Ord, J. K., & Snyder, R. D. (2001). A new look at models for exponential smoothing. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *50*(2), 147–159. https://doi.org/10.1111/1467-9884.00230

[10] Energy - Renewable Energy; Researchers at National Institute of Advanced Industrial Science and Technology (NIAIST) target renewable energy (Improvement of the 1,3-butadiene production process from lignin - a comparison with the gasification power generation process). (2019). *Energy & Ecology*, 1189.

[11] Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts. https://otexts.com/fpp3/

[12] Li, F., Liu, D., Sun, K., Hong, S., Peng, F., Zhang, C., & Tao, T., et al. (2024). Stochastic and extreme scenario generation of wind power and supply–demand balance analysis considering wind power–temperature correlation. *Electronics*, *13*(11), 2100. https://doi.org/10.3390/electronics13112100

[13] Ranalder, L., Busch, H., Hansen, T., Brommer, M., Couture, T., Gibb, D., & Guerra, F., et al. (2021). *Renewables in Cities: 2021 Global Status Report*. REN21.

[14] Viera, J., Aguilar, J., Rodríguez-Moreno, M., & Quintero-Gull, C. (2023). Analysis of the behavior pattern of energy consumption through online clustering techniques. *Energies*, *16*(4), 1649. https://doi.org/10.3390/en16041649

[15] Dataset 1: Yearly electricity data
[16] Dataset 2: global-renewable-power-sector-targets-2030