Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 322 – 345
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6696

Forecasting Renewable Energy Trends in the USA: An AI-Driven Analysis of Electricity Production by Source

Md Shihab Sadik Shovon[1], Clapher Ankur Gomes[2], Syed Ali Reza[3], Proshanta Kumar Bhowmik[4], Clinton Anthony Heaven Gomes[5], Tanaya Jakir[6], Mohammad Hamid Hasan Amjad[7], Arat Hossain[8], Md Sakibul Hasan[9]

Abstract

The United States is undergoing a significant shift toward renewable power as part of the broader international movement to counter climate change, reduce emissions, and increase energy independence. The shift is spearheaded by the mass adoption of solar, wind, and hydroelectric power, which are replacing the traditional fossil fuel-based power supplies. The primary purpose of this research was to develop and evaluate AI-based models for forecasting trends in the production of electricity through renewable energy in the USA. This research centered on the examination of renewable energy trends in the United States, with emphasis on solar, wind, and hydroelectric power. The dataset employed in this analysis encompasses wide-ranging electricity production data on renewable power, including solar, wind, and hydroelectric power, across several years to capture seasonal and long-term trends. The key data sources were from the United States' Energy Information Administration (EIA), offering detailed real-time and historical power production data at the national and regional levels, and the National Renewable Energy Laboratory (NREL), offering high-resolution renewable power generation, weather, and technology performance metrics data. Real-time grid data from regional transmission organizations (RTOs) and independent system operators (ISOs) has also been incorporated to add granularity and precision to the dataset. Three machine learning models were employed to make forecasts for renewable power production, namely, Random Forest, Support Vector Machines, and Gradient Boosting Regressors, each chosen for its unique strengths in tackling different aspects of the problem. For classification tasks, accuracy, precision, recall, and F1-score metrics were used to evaluate the models' ability to classify the energy production levels (e.g., high, medium, low). The maximum accuracy was achieved by Gradient Boosting, followed by SVM and Random Forest. In retrospect, AI insights are revolutionizing renewable energy planning in the USA through more accurate forecasts of the production and consumption of energy. These insights make it possible for grid managers to optimize the grid's capacity, ensuring that the infrastructure does not get over- or under-loaded. The integration of AI-based forecasting in renewable energy planning has significant policy and regulatory consequences at the federal and state levels. The integration of AI with smart grid technologies is a game-changer when it comes to renewable power management in the United States.

Keywords: *Renewable Energy, Electricity Production, Machine Learning, AI Forecasting, Energy Policy.*

Introduction

Barua et al. (2025) reported that the United States has witnessed a substantial growth in the use of renewable power in the past decade, following developments in technology, declining prices, and rising environmental awareness. Solar, wind, and hydroelectric power are increasingly included in the nation's basket of power, contributing immensely to the reduction in carbon emissions and supply diversification. This has been boosted by federal and state policy, through tax credits, renewable portfolio standards, and ambitious clean

[1] Master of Science in Engineering Management, College of Graduate and Professional Studies, Trine University, Email: mshovon23@my.trine.edu, (Corresponding Author)
[2] Masters in Information Technology, Washington University of Science & Technology.
[3] Department of Data Analytics, University of the Potomac (UOTP), Washington, USA.
[4] Department of Business Analytics, Trine University, Angola, IN, USA.
[5] MSIT in Data Management and Analytics, Washington University of Science and Technology
[6] Master's in Business Analytics, Trine University
[7] Master of Science in Engineering Management, College of Graduate and Professional Studies, Trine University
[8] Information Technology Management, St Francis College.
[9] Information Technology Management, St Francis College

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 322 – 345
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6696

power targets. However, the incorporation of renewable power in the grid has not come without its problems (Ahmed et al., 2025).

Anonna et al. (2023), asserted that, contrary to the constant and controllable supply provided by conventional fuels like coal and natural gas, renewable power production varies with weather and seasons, making supply challenging for grid managers, who must balance supply and demand in real-time to prevent power outages and ensure system stability. Chouskey et al. (2025) added that accurate forecasting of the output of electricity by renewable power plants is therefore essential for the efficient operation of the grid, storage optimization, and long-term planning. In the face of the rising importance of renewable power, conventional forecasting methods have proven inadequate in capturing the complexity of the systems.

Problem Statement

As per Choudhury et al. (2024), traditional forecasting models, which rely heavily on statistical techniques and historical data, cannot capture the stochastic and non-linear nature of renewable power production. For instance, the production of solar power relies on cloud cover, temperature, and daylight, while the production of wind power depends on wind speed, direction, and turbulence. These factors generate a high degree of uncertainty that cannot be modeled using conventional techniques. This makes renewable power production challenging to predict for grid managers and energy planners, leading to inefficiencies in dispatching power, higher reliance on reserve power, and higher operational costs. Gazi et al. (2025) contended that the limitations of the conventional techniques point to the need for more sophisticated techniques that are capable of capturing the complexity and volatility of renewable power systems. AI and ML offer a solution to the problem, as they are capable of dealing with vast quantities of data, identifying underlying patterns, and making accurate projections even when uncertainties exist.

Research Objective

The primary purpose of this research is to develop and evaluate AI-based models for forecasting trends in the production of electricity through renewable energy in the USA. Employing advanced machine learning methods, this research aims to increase the reliability and accuracy of renewable energy forecasts, which would, in turn, enhance decision-making for policymakers and power suppliers. This research focuses on three leading renewable energy sources, including solar, wind, and hydroelectric power, and examines the application of predictive analytics to address the most critical issues in the area of energy management. The findings of this research are intended to provide data-driven insights that could be utilized to aid sustainable planning for the supply of energy, enhance the stability of the grid, and make way for the low-carbon economy.

Scope and Relevance

This research centers on the examination of renewable energy trends in the United States, with emphasis on solar, wind, and hydroelectric power. These power sources are targeted because their contribution to the national supply has been on the rise and has the potential to increase further. This research employs a mixture of past production data on energy, weather data, and machine learning techniques to develop predictive models that can make accurate forecasts on electricity production. The relevance of this research extends beyond different stakeholders, including the energy sector, grid managers, policymakers, and environmental researchers. For the energy sector, the results achieved through this research can be employed to make informed investments, maximize the utilization of available resources, and improve efficiency in operations.

The results can be employed by policymakers to develop more efficient policies and incentives that foster the utilization of renewable energy technologies. For environmental researchers, the research can be employed in the better comprehension of the determinants influencing renewable energy production and their contribution to sustainability. In addressing the renewable energy forecasting issue, the research contributes towards the ultimate goal of having a sustainable and resilient power system in the USA.

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 322 – 345
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6696

Literature Review

Renewable Energy Growth Trends in the USA

Hasan (2024) ascertained that the United States has witnessed a significant transformation in its energy profile during the last two decades, with a steady increase in the use of renewable power sources such as solar, wind, and hydroelectric power. Statistics show that solar power has recorded growth in an accelerating trend, with installed capacity increasing from less than 1 gigawatt (GW) in the year 2008 to more than 100 GW in the year 2023, due to the advancement in photovoltaic technology, the decline in the cost, and the massive public and private investments. Hossain et al. (2025) demonstrated that wind power has also shown impressive growth, notably in the states of the Midwest and Great Plains, where the prevailing wind conditions and the massive wind farms made the USA one of the top wind power producers globally.

Hydroelectric power, although more mature and geographically restricted, continues to be a vital contributor to the renewable power base of the nation, with a stable and predictable source of electricity. The policies and incentives provided by the government have been instrumental in promoting this growth. Federal programs such as the Production Tax Credit (PTC) and the Investment Tax Credit (ITC) provided financial incentives to renewable power projects, whereas the renewable portfolio standards (RPS) at the state level compelled the utility companies to generate a specific quota of their power through renewable sources (Reza et al., 2025) International agreements such as the Paris Agreement also contributed to the USA's efforts to reduce carbon emissions and move towards a sustainable power future. In the backdrop of these developments, the integration of renewable power on the national grid remains a challenging and dynamic issue, with the need for creative solutions to forecasting and managing the power supply (Shawon et al., 2025).

Challenges in Energy Forecasting

According to Shil et al. (2024), one of the major challenges in renewable power forecasting is the natural variability and unpredictability of power production by solar and wind power, respectively. Solar power production relies heavily on weather conditions, including cloud cover, temperature, and daylight, which vary dramatically even on relatively short time scales. Similarly, wind power production relies on wind speed, direction, and turbulence, all subject to seasonal and geographical factors. Hydroelectric power, being more predictable, relies on rainfall, snowmelt, and water availability, all subject to seasonal and geographical factors as well. These dependencies result in a high degree of uncertainty in power forecasting, making supply and demand difficult to balance on the grid (Wen et al., 2024).

The integration of intermittent power on the grid further complicates the issue, as the grid infrastructure has historically been designed with centralized, predictable power production based on fossil fuels. The unpredictability of renewable power production results in grid instability, increased operating costs, and the necessity for reserve power plants, including natural gas plants or power storage facilities (Ukoba et al., 2024). To address these issues, more sophisticated forecasting methods must be utilized that are capable of forecasting renewable power production with high accuracy and facilitating grid operators' optimization of power dispatch, storage, and delivery (Srinivasan et al., 2024).

Machine Learning for Energy Forecasting

As per Rehan (2024), over the past few years, artificial intelligence (AI) and machine learning (ML) have emerged as strong solutions to the problems related to renewable energy forecasting. In contrast to the conventional statistical methods, based on linear models and historical data, AI-based techniques can recognize complex, non-linear relationships and learn under changing conditions. Machine learning algorithms, such as neural networks, support vector machines (SVM), and random forests, have been widely applied to issues related to time-series prediction, including energy forecasting.

Supervised learning models, on the other hand, are extremely promising on this front, as they can be trained on large datasets with past energy production information, weather information, and other related variables.

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 322 – 345
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6696

For example, recurrent neural networks (RNNs) and long short-term memory (LSTM) networks are perfectly suited for forecasting in the form of a sequence because they can model temporal relationships and process sequential data (Serban & Lytras, 2020). Additionally, ensemble techniques, which utilize multiple models to improve the predictability of the prediction, have been successfully implemented to deal with the volatility and uncertainty related to renewable power production. These AI-based techniques not only enhance the accuracy of the forecasts but also enable real-time adjustments and adaptive learning, making them extremely valuable for grid operation and planning for power (Ohalete et al., 2023).

Research Gaps

Adegbola (2025) argued that despite the rising number of research articles on AI-based energy forecasting, several gaps remain that must be closed through additional research. One gap that remains prominent is the lack of research on renewable energy forecasting in the specific case of the USA. While the literature on renewable energy trends and forecasting methodologies globally is expansive, the unique character of the US energy market, with its geographically variable terrain, climate, and policy, necessitates a more regionalized emphasis. According to Alsharif (2025), most research currently relies on static models that do not account for the dynamic character of renewable energy systems, including changes in government policy, technology, and climate fluctuation.

For instance, the repeal of federal tax credits or the introduction of new regulations could radically change the uptake and production of renewable energy, but few models account for these variables in their forecasts. Similarly, the increasing frequency of weather-related extreme events due to climate change poses new challenges for energy forecasting, as conventional models may not be geared to handle such outliers. Adaptive models that learn continually from new data and alter their forecasts accordingly to ensure long-term reliability and accuracy are necessary as well. Closing these gaps in research is crucial to the development of forecasting tools that can make the shift towards a sustainable energy future in the USA a reality (Amini and Rohani, 2024).

Data Collection and Exploration

Dataset Overview

The dataset employed in this analysis encompasses wide-ranging electricity production data on renewable power, including solar, wind, and hydroelectric power, across several years to capture seasonal and long-term trends. The primary data sources are the United States' Energy Information Administration (EIA), offering detailed real-time and historical power production data at the national and regional levels, and the National Renewable Energy Laboratory (NREL), offering high-resolution renewable power generation, weather, and technology performance metrics data. Real-time grid data from regional transmission organizations (RTOs) and independent system operators (ISOs) has also been incorporated to add granularity and precision to the dataset. The utilization of multiple data sources ensures a robust and integrated representation of renewable power production, making it possible to analyze spatial and temporal trends, as well as identify the significant drivers influencing the trends in production. The dataset has variables that range from hourly and daily power output, weather conditions (e.g., solar irradiance, wind speed, rainfall), location, and infrastructure capacity, making it a rich source for the development and validation of AI-based forecasting models.

Key Features Selection

S/No.	Key Features	Description
001.	**Hourly Energy Production**	Information on electricity generated on an hourly basis for solar power, wind power, and hydroelectric power.
002.	**Total Daily Energy Output**	Total daily renewable electricity production

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 322 – 345
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6696

003.	**Geographical Location**	Regional or state-level data on where the power is being generated (e.g., solar farms in California, wind farms in Texas).
004.	**Solar Irradiance**	Sunlight intensity measurements, essential for forecasting solar power production.
005.	**Wind Speed**	Information on wind speed and direction which are necessary for wind power production forecasting.
006.	**Precipitation Levels**	Rainfall and snowfall data are specifically valuable for hydroelectric power production.
007.	**Temperature**	Ambient temperature information affects both the efficiency of solar panels and the energy demand.
008.	**Infrastructure Capacity**	Capacity installed in renewable energy facilities (e.g., megawattage of solar panels or wind turbines).
009.	**Seasonal Trends**	Data that indicates seasonal shifts in the production of energy (e.g., increased solar power in the summer, increased hydroelectric power during spring snowmelt).
010	**Grid Integration Metrics**	Real-time grid information, including demand, supply, and balancing metrics, reported by ISOs and RTOs.

Data Preprocessing

The code snippet performed a typical data preprocessing pipeline for machine learning work. The snippet begins with the importing of necessary libraries, including Pandas for data manipulation and Scikit-learn for model selection and data preprocessing. Subsequently, the snippet loads a dataset and handles missing values by dropping rows with necessary missing identifiers and replacing missing values for the 'Energy Source' variable with zero. Categorical variables 'Entity' and 'Code' are encoded using Label Encoder, and numerical variables are standardized using Standard Scaler. The target variable 'Dominant Source' is created based on the identification of the most dominant source per year. The data is then split into the training set and the testing set, and the shapes of the two sets, as well as the class distribution of the target variable in the training set, are printed. The snippet closes by indicating that the data is prepared and available for classification modeling.

Exploratory Data Analysis (EDA)

Exploratory Data Analysis, or EDA, refers to the systematic examination and summarization of a dataset to identify patterns, trends, associations, and outliers. EDA is an adaptive, iterative process that applies a combination of statistical techniques, data graphics, and domain knowledge to create a better understanding of the form, quality, and underlying properties of the data. EDA usually precedes the use of formal modeling or hypothesis testing because it allows the most significant features to be found, outliers to be detected, and informed questions or hypotheses to be formulated for further investigation.

Trends in Electricity Production by Source in the USA

The implemented code snippet constructed a line graph to visualize the trends in the production of electricity in the USA. This snippet initially filters the data to display only those entries where the 'Entity' is 'United States'. Then, it sets the theme for the graph using the "white grid" theme available in the seaborn library for better aesthetics. It constructs a line graph to show the trends in different sources of energy ('Coal', 'Gas', 'Nuclear', etc.) concerning the year ('Year'). Each source of energy has a different line with a label associated with it. The graph has the title "Trends of Electricity Production by Source in the USA", appropriate axis titles, and a legend for better understanding. Finally, a grid has been added for better readability, and the graph has been displayed using plt.show(). This graph enables us to visualize the shift in the production of electricity due to different sources over the years in the United States.

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 322 – 345
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6696

Output:

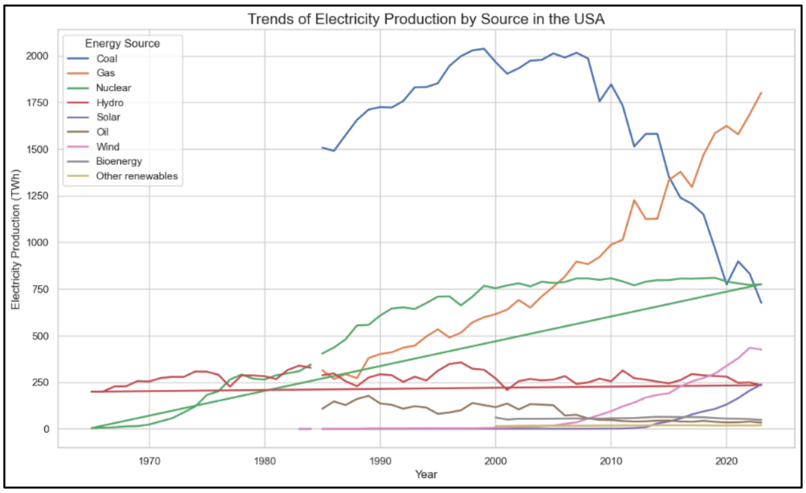


Figure 1: Trends of Electricity Production by Source in the USA

The chart above illustrates the trends of electricity production by source in the USA from 1970 to 2020, revealing significant shifts in energy generation dynamics over the past five decades. Notably, coal, which dominated the energy landscape in the late 20th century, peaked around 2007 with production exceeding 1,800 TWh but has since experienced a marked decline due to environmental concerns and the rise of alternative energy sources. In contrast, natural gas production has steadily increased, surpassing coal in recent years, reflecting a growing reliance on gas as a transitional fuel amid the shift toward cleaner energy. The data also highlights the dramatic rise of renewable energy sources, particularly wind and solar, which have shown exponential growth since the early 2000s, with wind energy production reaching approximately 350 TWh and solar energy production experiencing a meteoric rise to about 300 TWh by 2020. Hydropower remains relatively stable, while bioenergy and other renewables have gradually increased but still represent a smaller share of total production. This transition underscores the USA's evolving energy portfolio, driven by technological advancements, policy changes, and a growing commitment to sustainability and reducing carbon emissions.

Correlation Between Energy Sources in the USA

The implemented script generated a heatmap to visualize the correlation among different energy sources in the USA. First, it generates a figure with the given size. Then, it calculates the pairwise correlation matrix for the selected energy sources ('Coal', 'Gas', 'Nuclear', etc.) using the corr() function on the filtered USA data. Finally, it generates a heatmap using the heatmap function in the seaborn library, with the correlation values labeled, a 'cool warm' colormap, and the given line widths. The figure is titled "Correlation Between Energy Sources in the USA" with the font size set to 16, and finally shown using plt.show(). This heatmap easily identifies the positive and negative correlations among the different energy sources, providing insights regarding their relationship and potential dependency on one another.

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 322 – 345
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6696

Output:

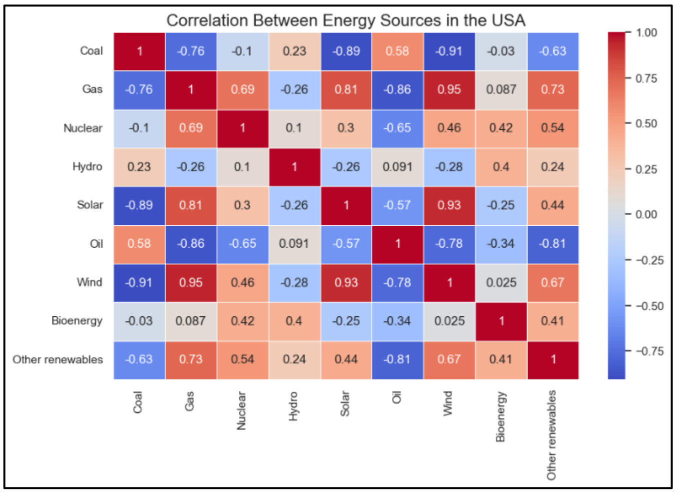


Figure 2. Correlation Between Energy Sources in the USA

The correlation matrix indicating the interrelations among the various sources of power in the USA exhibits several interesting trends that portray the interdependencies among the various sources. Coal and wind are found to be highly negatively correlated with a correlation value of -0.91, indicating that as the production increases for wind power, production decreases for coal, indicating the rising trend towards renewable power at the expense of fossil fuels. Coal also exhibits the same strong negative correlation with solar power (-0.87), indicating that the growth in solar power has an inverse relationship with the consumption of coal. In contrast, natural gas exhibits a moderate correlation with coal (0.76), indicating that the two might coexist up to a point in the power mix. Nuclear power exhibits a weak correlation with other power sources, that is, with wind power (0.09) and solar power (0.26), indicating that production has less correlation with the fluctuation in renewable power production. Between the renewable power sources, solar power and wind power show a strong correlation with each other (0.96), indicating their complementary nature in the renewable power system. Overall, the matrix indicates the complex dynamics among the power sources, indicating the falling position of traditional fossil fuels like coal in favor of renewable power in the new power system in the USA.

Total Electricity Production by Source in the USA

The computed code generated a horizontal bar chart to show the total electricity production by source in the USA for all years. It first calculated the sum for each source ('Coal', 'Gas', 'Nuclear', etc.) by excluding the 'Entity', 'Code', and 'Year' columns, effectively grouping the production for all years. The results are ordered in descending order so that the leading sources with the highest production are easily visible. Using the barplot function provided by the seaborn library, a horizontal bar chart is generated with the values for the total production on the x-axis and the names of the sources on the y-axis. The 'viridis' color palette has been used for the chart for enhanced aesthetics, and the title "Total Electricity Production by Source in the USA" with a specified font size has been added. The x-axis has the label "Total Electricity Production (TWh)" added to provide context, and the chart has been shown using plt.show(). This chart clearly shows the contribution of each source to the total electricity production in the USA as a whole.

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 322 – 345
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6696

Output:

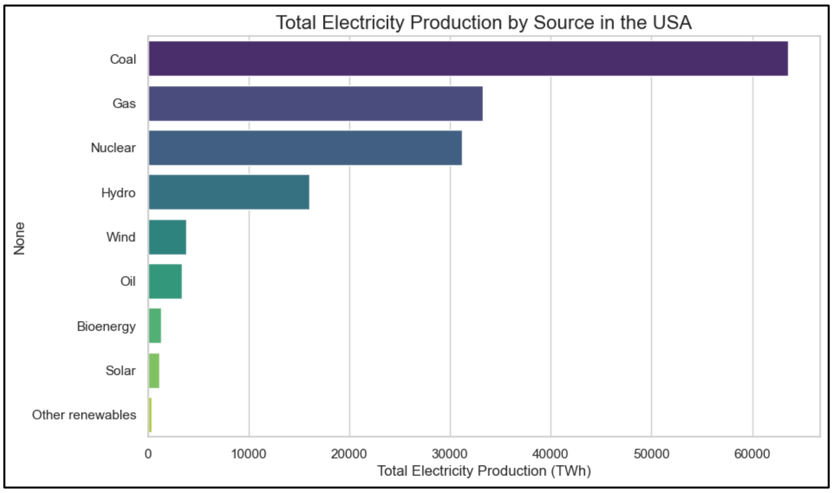


Figure 3. Total Electricity Production by Source in the USA

The bar chart representing the production of electricity by source in the USA captures the dominance of coal, which leads the pack with production far over 60,000 TWh, showing its traditional role as the primary source for the production of electricity. The second major source is natural gas, although with much lower production levels when compared to coal, showing the rising reliance on gas, but still showing the dominance of coal in the energy production base. Nuclear power stands third, with approximately 30,000 TWh, showing its stable contribution as a baseload power source. In comparison, renewable power sources such as hydro, wind, solar, bioenergy, and other renewables contribute a lower proportion to the production, with hydro leading the pack in the renewable category at approximately 15,000 TWh. Solar and wind, although on the rise, are much lower in the production level, showing the ongoing trend towards renewable power sources. The gap in the production level between fossil fuels and renewable sources captures the complexity and opportunity in the USA's energy landscape as it transforms towards a more sustainable and diversified source base.

Pair Plot of Energy Sources in the USA

The implemented created a pair plot to display the pairwise relationships between different energy sources in the USA. The code utilizes the pair plot function of seaborn to create a matrix of the scatterplots, where each scatterplot represents the relationship between two energy sources ('Coal', 'Gas', 'Nuclear', etc.). The matrix diagonal typically has the histograms or kernel density estimations for each energy source, representing their distributions. The plt.suptitle function adds a title "Pairplot of Energy Sources in the USA" to the figure, slightly above the plots with y=1.02 and a certain font size. Finally, plt.show() displays the created pairplot. This enables the visualization of the joint distributions and potential correlations between different energy sources, revealing the underlying patterns and dependencies that might not be easily apparent through individual plots or summary statistics.

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 322 – 345
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6696

Output:

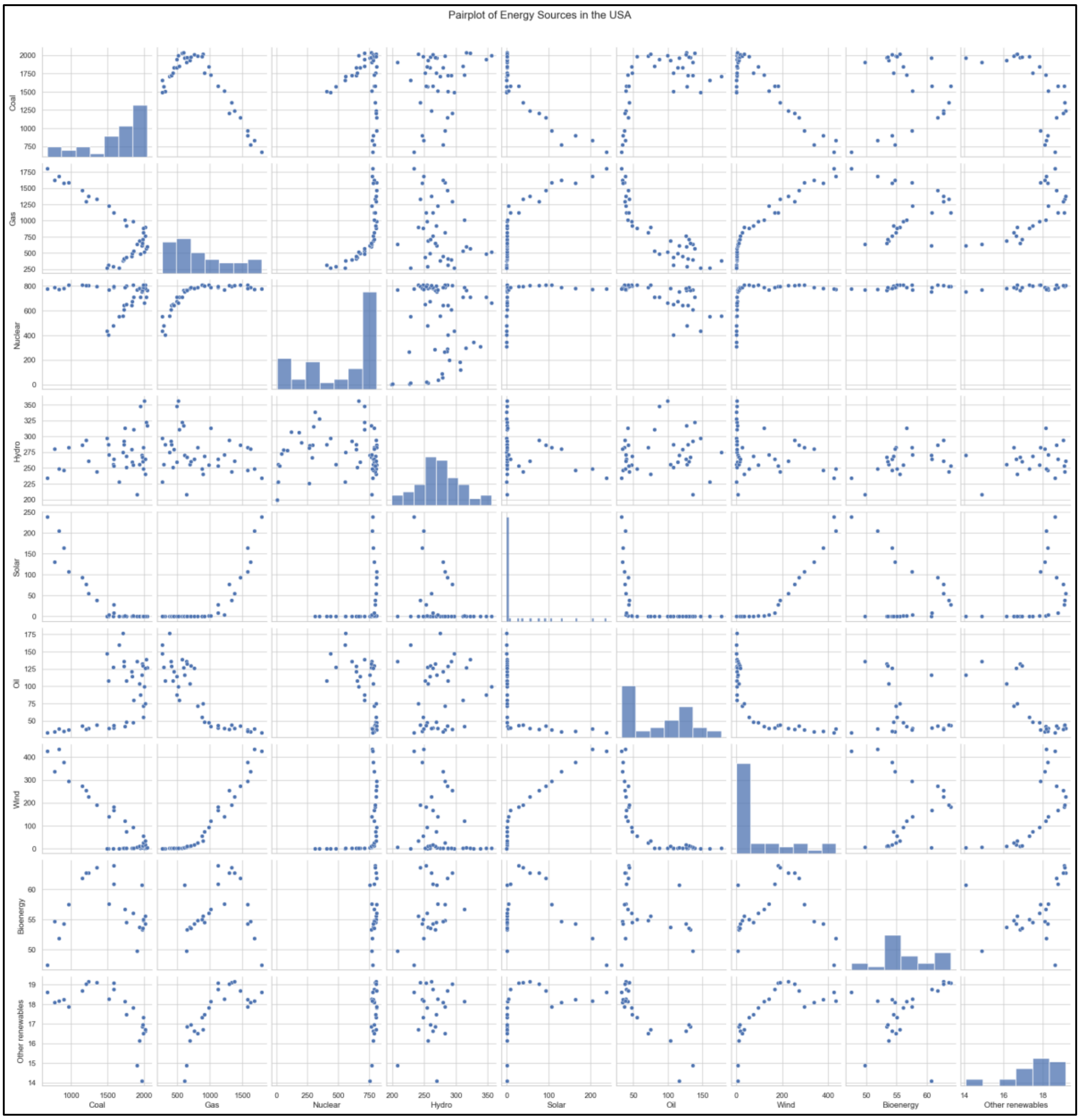


Figure 4. Pair Plot of Energy Sources in the USA

The pair plot graph depicting the relationship between the various types of energy sources in the USA gives a detailed visual analysis of the interrelations between the various types of electricity production. Each scatter plot in the matrix gives the distribution and potential relationship between the pairs of the various energy sources, with the histograms on the diagonal giving the individual frequency distributions for each source. Interestingly, we observe strong positive correlations between the renewable sources, wind and solar, with the implication that as one increases, so does the other. Conversely, we observe a distinct negative relationship with coal and both wind and solar production, as we expect the trend that the increase in the growth of renewable sources occurs at the expense of the fossil fuel sources. The scattering pattern for gas and coal gives a more complex interaction, with gas having some degree of positive correlation with coal, with the implication that the two are capable of coexisting in the energy mix. The gaps and the clustering in the data points also signify the varying levels of the production of the energy across the various sources, indicating the varying dynamics in the energy scene in the USA. Overall, the pairplot gives a good

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 322 – 345
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6696

indication of the complex interdependencies in the various energy sources, providing us with useful insights into the current trend in the energy scene.

Distribution of Electricity Production by Source in the USA

The code script in Python generates boxplots to visualize the distribution of electricity production for all the energy sources in the USA over the years. First, it generates a figure with a specified size. Subsequently, it uses the boxplot function in the library seaborn to plot the distributions of 'Coal', 'Gas', 'Nuclear', 'Hydro', 'Solar', 'Oil', 'Wind', 'Bioenergy', and 'Other renewables' in the filtered data for the USA. The graph has the title "Distribution of Electricity Production by Source in the USA" with a font size of 16 points. The x-axis has the title "Energy Source", the y-axis has the title "Electricity Production (TWh)", and the tick marks on the x-axis are set at a rotation of 45 degrees for easier readability. Finally, plt.show() displays the generated boxplots, making it easy to compare the central tendency, spread, and potential outliers in electricity production by various energy sources.

Output:

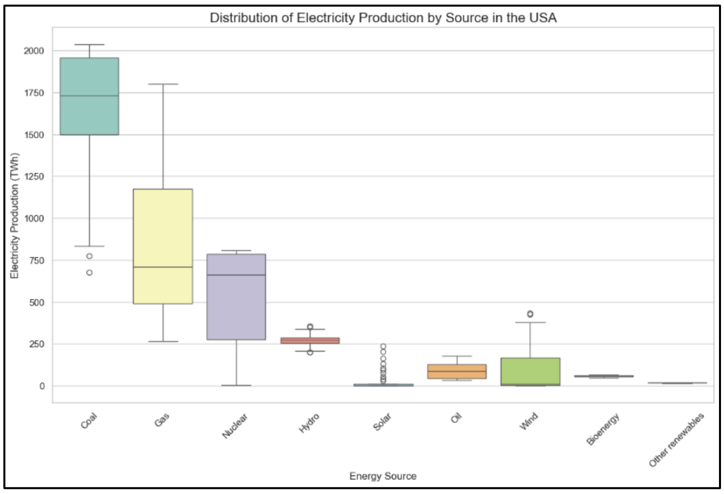


Figure 5. Distribution of Electricity Production by Source in the USA

The boxplot graph representing the production of electricity by source in the USA reflects the wide inequalities in production across different sources of power. Coal stands out as the dominant source, with a median level of production of approximately 1,750 TWh and a wide interquartile range, showing its traditional dominance in the power supply and the volatility in the level of production across different years. Natural gas, though much lower in median production at approximately 1,000 TWh, has a more even spread with fewer outliers, showing a stable but rising contribution to the power supply. Nuclear power also has a stable level of production, with a median close to 800 TWh and a narrower interquartile range, showing its reliability as a baseload power source. In contrast, renewable power supplies such as wind and solar power have lower median levels of production, approximately 300 TWh and 200 TWh, respectively, but their distributions reflect wide volatility and the potential to rise as technology evolves and uptake increases. The outliers, notably in solar and wind, show the incidence of exceptionally high production, perhaps due to favorable weather conditions or technology innovations. Overall, the graph reflects the heterogeneous character of electricity production in the USA, showing the traditional dominance of fossil fuels but the rising contribution of renewable power supplies.

USA Energy Production by Source

The piece of code depicted the trend in the production of the USA's energy using an area plot. The data is initially filtered to include the countries of interest ('United States', 'China', 'India', 'Germany', 'Brazil') and then filtered again to include data for the 'United States' alone. The theme for seaborn has been set as

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 322 – 345
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6696

'whitegrid' to improve the visual representation. The area plot depicting the trend in various energy sources ('Coal', 'Gas', 'Nuclear', etc.) over the years ('Year') has been created using the following code. The title "USA Energy Production by Source (Area Plot)" with proper axis titles and a legend outside the figure for the sake of clarity has been added to the figure. Grid lines have been added for better readability, and the figure has been plotted using plt.show(). The following graph indicates the cumulative contribution of the various energy sources towards the production of the USA's energy across the years.

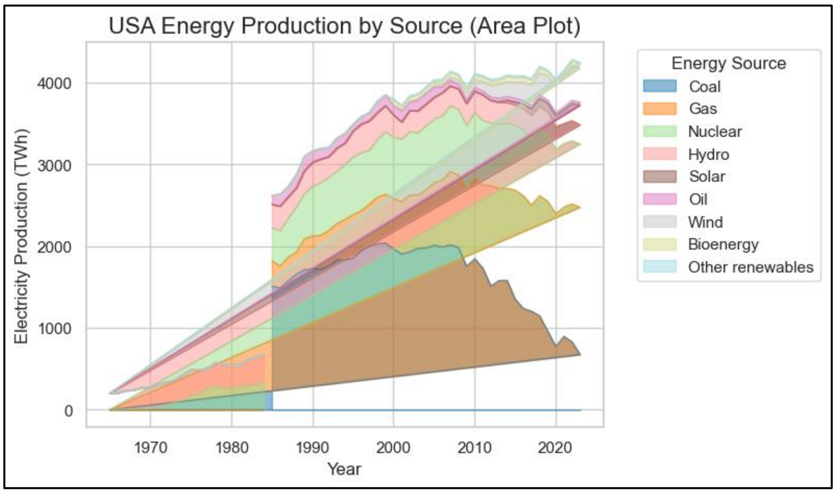


Figure 6. USA Energy Production by Source

The area graph representing the production of USA energy by source from the year 1970 to the year 2020 reflects the significant changes that the power profile has witnessed during the last half century. Coal, which had been the dominant source, reflects a production peak in the year 2007 at approximately 4,000 TWh, followed by a declining trend due to increased environmental constraints and competition with cleaner fuels. In contrast, natural gas has reflected a constant growth, overtaking coal in the last few years and showing the role as a bridging fuel in the transition towards cleaner fuels. Nuclear power has been relatively stable, contributing consistently to the power base with production at approximately 800 TWh. The most notable trend, however, has been the explosive growth in renewable power sources, with wind and solar power leading the way, rising dramatically since the start of the year 2000. The production of wind power has risen sharply, touching over 400 TWh, whereas the production of solar power has risen dramatically from trace levels to over 200 TWh by the year 2020, showing the fast growth in its usage. Hydroelectric power has been a major contributor, but with less rapid growth. The area graph reflects the trend towards a more sustainable and diversified power base, with the decline in fossil fuels making way for renewable power in the USA as it transitions towards tackling climate change and enhancing the security of the power supply.

Total Energy Production by Source (USA)

The script in Python generates a radial bar chart to show the total production by source of energy in the USA. It starts by summing up each source after dropping 'Entity' and 'Code' columns and sorts the results in descending order. The angles for each source to be plotted in the radial plot are calculated, and the values to be plotted are prepared. A polar subplot is generated with plt.subplots, and the data are plotted with ax. Fill and ax.plot to create the radial bars. The y-axis tick labels are removed, and the x-axis tick labels are set to the names of the source with a specified font size. The figure is labeled "Total Energy Production by Source (USA) - Radial Chart" and displayed with plt.show(). This radial chart presents a clear view of the production by source, with the relative contribution of each source represented circularly.

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 322 – 345
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6696

Output:

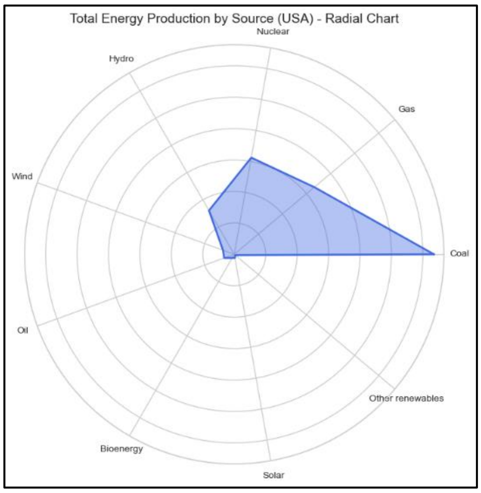


Figure 7. Total Energy Production by Source

The radial chart showing the production of the total energy by source in the USA gives a unique visual representation of the contribution made by various types of energy. Coal has the center position on the chart, overshadowing the other sources with its huge production share, showing its traditional position in the production of electricity. The second position belongs to natural gas, whose representation, though much less significant than that of coal, shows its growth but relatively low contribution when compared with the fossil fuels. The renewable ones, wind and solar, are lower on the chart, showing their relatively low contributions to the production, even though their growth has been increasing in the past few years. Hydropower has a modest presence, bioenergy and other renewables having the smallest space, showing their relatively low contributions to the production as a whole. This chart clearly shows the current situation in the USA, showing the continued dominance by fossil fuels, particularly coal, but also showing the growth potential in renewable ones as the country moves towards a more sustainable future for the production of energy.

USA Energy Production Over Time

The computed code script produces a ridge plot to visualize the production density of each source in the USA over the years. It iterates through different energy sources ('Coal', 'Gas', 'Nuclear', etc.) and uses the kdeplot function in seaborn to plot the kernel density estimation of each source's production. The area under the density line is colored using the fill=True argument, and transparency is achieved using the alpha=0.6 parameter. The title for the graph is set to "USA Energy Production Density Over Time (Ridge Plot)" with appropriate axis titles, and the legend is placed outside the graph using the legend='outside' parameter. The grid makes it more readable, and the graph is displayed using plt.show(). This ridge plot easily shows the production distribution for each source, making it simple to compare their shapes and centering points.

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 322 – 345
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6696

Output:

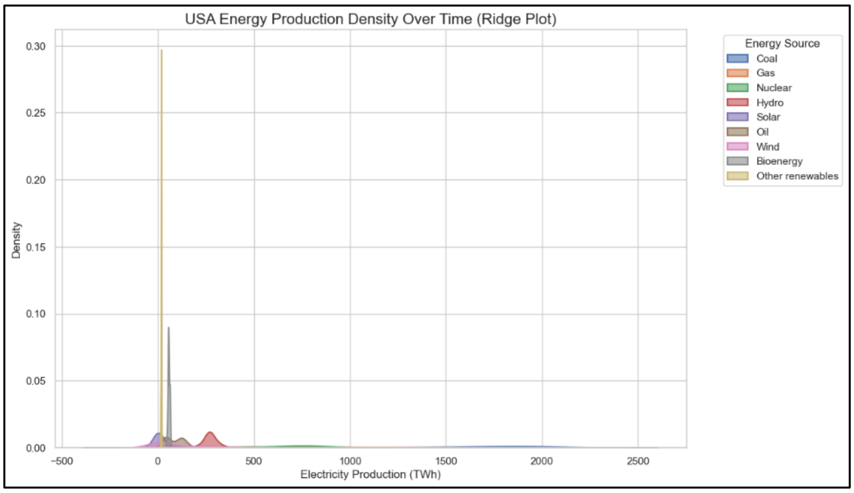


Figure 8. USA Energy Production Over Time

The ridge plot depicting the USA's energy production density over the years presents a holistic view of the spread of various sources of energy and their production levels. The graph presents a cluster of production values near zero for the majority of the sources, with dominance by low-production years, notably for wind, solar, and other renewable sources. Coal and gas show a wider range, with coal having a high value at approximately 1,500 TWh, indicating the dominance that coal had in the supply of energy in the past. The occurrence of various peaks for gas reflects a more fluctuating trend in production, perhaps due to market conditions and policy changes. Nuclear power has a constant production value, with hydro power having a narrower, stable range, indicating its well-established position in the supply chain of energy. The rising densities for wind and solar over the years reflect their rising presence, indicating the shift towards renewable energy as the latter gain popularity in the market. Overall, the ridge plot presents the changing dynamics in the production of energy in the USA quite well, indicating the dominance that fossil fuels had in the past as well as the trend towards the usage of renewable energy that the latter are witnessing in the market.

Solar Energy Production Over Time

The line of code produces a heatmap representing solar power production by country, year by year. First, it constructs a pivot table based on the filtered data, with 'Entity' as the index, 'Year' as the columns, and 'Solar' as the values, and sums the values, filling in the missing values with 0. Then, a heatmap is generated using the pivot table as the input, with a 'YlGnBu' colormap, line width specified, line color, and a colorbar label as "Solar Energy (TWh)". The title "Solar Energy Production Over Time - Country Comparison" has a font size of 16, and the x and y axes are labeled "Year" and "Country", respectively, with font sizes of 12. Finally, plt.show() displays the generated heat map, making it simple to compare solar power production by country and year.

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 322 – 345
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6696

Output:

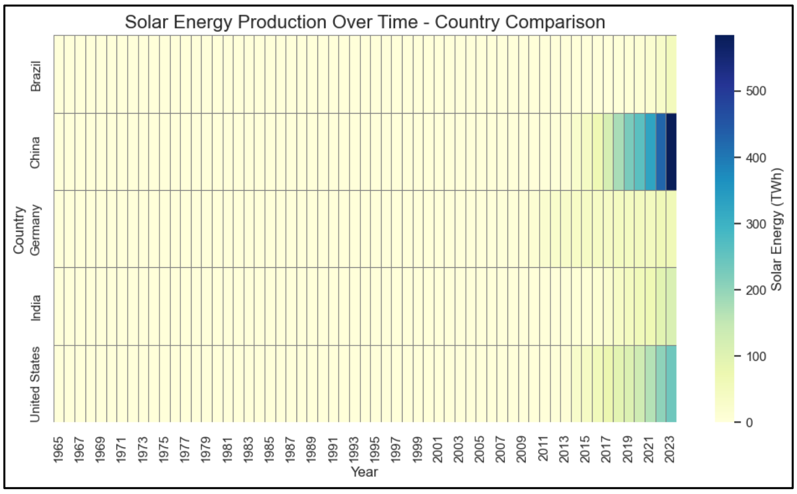


Figure 9. Solar Energy Production Over Time

The graph above comparing solar power production year-wise for Brazil, China, Germany, India, and the United States shows major trends in the global solar power industry for the years 1965-2023. At the start, solar power production for all the countries remained low, with Germany leading the pack in the early years due to its pioneering contributions to solar technology. However, the graph shows a steep increase in the production of solar power in China, particularly from the year 2010, making it the leading solar power producer in the world, with production crossing the figure of 500 TWh by the year 2023. The United States and Germany also see massive growth, with the former seeing major production figures, although it falls behind China. India shows a major increase in the production of solar power, particularly in the latter years of the decade, indicating a rising emphasis on renewable power in the face of its rising energy demand. Brazil, although showing growth, remains far behind the other countries in production figures. This graph reflects the sudden increase in solar power capacity, particularly in China, and points towards the turn towards renewable power across the globe due to climate change and security concerns regarding power supply.

Methodology

Feature Engineering

Feature engineering plays a significant role in the preparation of the data for machine learning models, transforming raw data into meaningful features that enhance predictive power. For this analysis, significant variables such as temperature, wind speed, solar radiation, and the extent of energy storage were selected based on their contribution to renewable energy production. Temperature affects the efficiency of solar panels and the power demand, and wind speed directly affects the output of wind turbines. Solar radiation contributes largely to solar power production, and the extent of energy storage indicates grid stability and the ability to balance demand and supply. To capture temporal dynamics, new temporal features in the form of rolling averages, lagged variables, and seasonal flags were added. For short-term forecasting, rolling averages of weather variables (e.g., the average wind speed over the previous 24 hours) were computed to minimize the extent of noise and maximize trends. For long-term forecasting, lagged variables (e.g., last week or last month's power production) and seasonal flags (e.g., binary flags for summer or winter) were included to capture cyclical trends and past dependencies. Not only do these engineered features enhance the model's ability to capture temporal dynamics, but they also better capture the underlying causes contributing to renewable power production.

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 322 – 345
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6696

Model Training and Selection

Three machine learning models were employed to make forecasts for renewable power production, notably, Random Forest, Support Vector Machines, and Gradient Boosting Regressors, each chosen for its unique strengths in tackling different aspects of the problem. The Random Forest model, an ensemble learning model, was chosen because it can calculate the importance of the features and handle non-linear relationships. Combining the outputs of many decision trees, Random Forest produces stable and interpretable results, making it ideal for the identification of the most significant drivers of power production. The Support Vector Machine (SVM) regression model was chosen because it has high efficiency in high-dimensional spaces and the ability to model complex relationships between input variables and power production. The kernel functions in SVM allow it to learn non-linear relationships, making it ideal for the forecasting of power production under varying weather conditions. The Gradient Boosting Classifier, with its high accuracy and precision in forecasting, was the last model selected. Gradient Boosting creates an ensemble of weak models sequentially, minimizing error and increasing accuracy with each addition. These models were selected not only because of their forecasting power but also because of their interpretability, which is important in the provision of actionable insights to decision-makers.

Model Optimization and Performance Analysis

In pursuit of maximum performance for the selected models, hyperparameter optimization with Grid Search and Random Search has been conducted. Grid Search exhaustively samples over a set of predefined hyperparameters, whereas Random Search samples based on a distribution over possible values, giving a more efficient solution for high-dimensional spaces. The two methods were both utilized to find the optimal configuration for each model, including the number of trees in Random Forest, the regularization parameter in SVM, and the learning rate in Gradient Boosting. Additionally, cross-validation has been used to enable the models to generalize to new, unseen data. Dividing the dataset into multiple folds and measuring the model performance on each fold, cross-validation reduces the risk of overfitting and gives a better indication of model accuracy. This rigorous optimization ensures the models are accurate and robust, capable of handling the complexity and variability in renewable energy data.

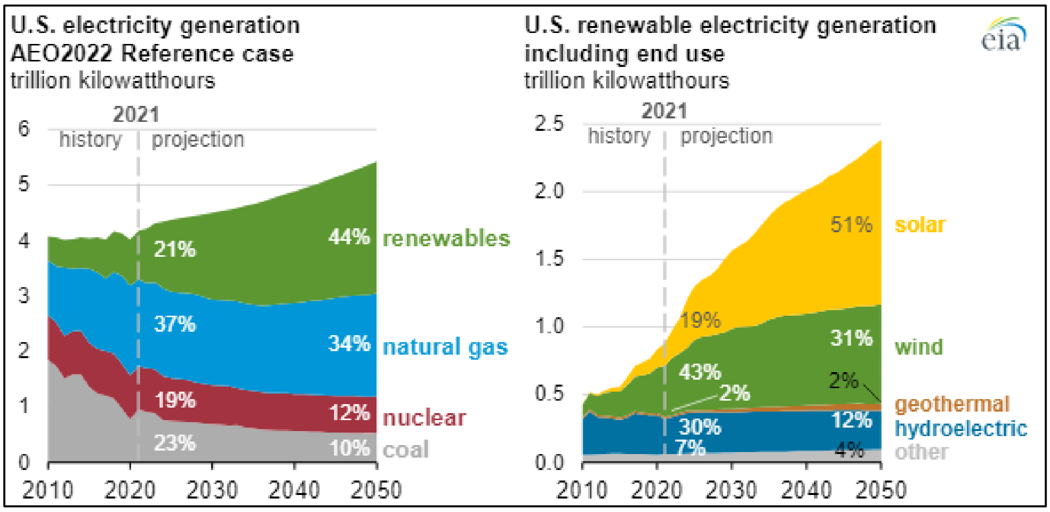
Evaluation Metrics

The performances of the models were evaluated using a combination of regression metrics and classification metrics, depending on the type of forecasting task involved. For regression tasks, Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared (R^2) metrics were used to evaluate the accuracy of the prediction of the production of energy. For classification tasks, accuracy, precision, recall, and F1-score metrics were used to evaluate the models' ability to classify the energy production levels (e.g., high, medium, low). The comparison of the models revealed trade-offs in terms of accuracy, interpretability, and computational efficiency. For instance, even though Gradient Boosting performed with the best accuracy, it was computationally costly and less interpretable when compared to Random Forest. SVM offered a compromise in terms of accuracy and interpretability but had scalability issues when handling big datasets. Random Forest, on the other hand, offered a good compromise, with interpretable results and good accuracy as well as computational efficiency. These results guided the final model selection, ensuring that the selected model captures the specific needs of renewable energy forecasting without compromising on the balance of practicability and performance.

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 322 – 345
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6696

Results and Analysis

Renewable Energy Generation Trend Analysis



The graph above displays projected trends in the production of electricity in the United States from 2021 to 2050, for two categories: all electricity production and renewable electricity production with final use. In the left graph, the trend indicates a significant shift toward renewable power, with the share increasing from 44% in the year 2021 to 51% by the year 2050, with an increasing focus on sustainability. Natural gas remains a significant contributor, with a constant share of approximately 34%, and nuclear power has a declining share, dropping to 11% from 12%. Coal power falls dramatically, dropping to a mere 8% by the year 2050, as part of a broader trend towards avoiding fossil fuels. The right-hand graph concentrates on the breakdown of renewable electricity production, with solar and wind as significant contributors. Solar power will rise dramatically, contributing 43% of the renewable production by the year 2050, and wind power contributes 30%. Hydro and geothermal power contribute to the balance, with lower but still notable contributions. This projection reflects a revolution in the American energy system based on the development of renewable technologies and a policy focus on reducing carbon emissions.

Model Performance Evaluation

Random Forest Modelling

The computed code sample performed the training and testing of a Random Forest Classifier with hyperparameter tuning using Grid Search CV. It starts with the importing of libraries in Scikit-learn required for model creation, hyperparameter tuning, and testing. The range of values to be considered for hyperparameters like n-estimators, max depth, min-samples-split, min-samples-leaf, and bootstrap are specified using a parameter grid. An instance of Random-Forest-Classifier with a random state for reproduction purposes is created. Grid Search CV carries out a cross-validated search for the optimal set of hyperparameters. The model trains the training data using grid_search.fit, and the best estimator is acquired with grid-search. best-estimator_. The best set of parameters is printed, and the model's performance on the test set is calculated using the accuracy score and the creation of a classification report. This allows the optimization of the Random Forest model's performance by systematically searching through a set of hyperparameter space.

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 322 – 345
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6696

Output:

Table 1. Random Forest Classification Report

	precision	recall	f1-score	support
Bioenergy	1.00	0.87	0.93	45
Coal	0.98	0.96	0.97	46
Gas	0.93	0.96	0.94	96
Hydro	0.97	0.99	0.98	113
Nuclear	0.91	0.93	0.92	69
Oil	0.96	0.96	0.96	120
Other renewables	0.97	0.98	0.97	59
Solar	1.00	0.99	1.00	1074
Wind	0.86	0.97	0.91	31
accuracy			0.98	1653
macro avg	0.95	0.96	0.95	1653
weighted avg	0.98	0.98	0.98	1653

The table presents the classification performance metrics for various types of energy, with precision, recall, F1-score, and values for support. Interestingly, bioenergy has the highest value for precision at 1.00, with perfect precision in classifying this type, and high recall at 0.87 as well as an F1-score at 0.93, with a good balance in recall and precision. Coal and gas also perform well with high F1-scores and precision, with the implication that the two are correctly classified, while hydro and nuclear perform slightly lower but with good efficiency in general. Conversely, wind power has a low value for precision at 0.86 but with high recall at 0.97, with the implication that even though the majority of the instances are correctly classified, errors occur. The model has a high value for accuracy at 0.98, with the implication that the model performs exceptionally on all the categories. The macro and the weighted average also confirm the efficiency of the model, with the implication that the model performs exceptionally well in correctly classifying the various types of energy. This table demonstrates the strengths of the model in precision and recall on most types of energy, with the implication that the model has real-world applicability in the task of classifying the various types of energy.

SVM Modelling

The implemented code script demonstrated the procedure to train and evaluate a Support Vector Machine (SVM) classifier with hyperparameter optimization using Randomized Search CV. The necessary libraries in Scikit-learn are imported for model initialization, hyperparameter optimization, and evaluation. The parameter distribution is defined to set the range of values to be considered for hyperparameters like C, kernel, and gamma. An SVC classifier with a random state for reproduction is instantiated. Randomized Search CV carries out cross-validation and optimization for the best set of hyperparameters based on samples drawn from the defined distributions. The model gets trained on the train set using random_search.fit, and the best estimator is extracted using random_search. best_estimator_. The best set of parameter values is printed, and the model's performance on the test set is calculated using the accuracy score and the creation of a classification report. This allows the optimization of the SVM model performance through efficient searching through a given hyperparameter space.

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 322 – 345
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6696

Output:

Table 2. SVM Classification Report

	precision	recall	f1-score	support
Bioenergy	0.93	0.82	0.87	45
Coal	0.92	0.96	0.94	46
Gas	0.92	0.93	0.92	96
Hydro	0.94	0.95	0.94	113
Nuclear	0.86	0.91	0.89	69
Oil	0.88	0.88	0.88	120
Other renewables	0.91	0.98	0.94	59
Solar	0.98	0.97	0.98	1074
Wind	0.83	0.94	0.88	31
accuracy			0.96	1653
macro avg	0.91	0.93	0.92	1653
weighted avg	0.96	0.96	0.96	1653

The table displays the metrics for the performance of a Support Vector Machine model classifying various energy sources with an average accuracy level of approximately 95.5%. The precision for bioenergy is at 0.93, indicating a high level of correctness for the positive calls, but with a recall measure at 0.82, indicating that some instances are missed. Coal and gas fare well, with coal having a precision value of 0.92 and a high recall value of 0.96, indicating reliability in identifying as well as capturing this class. Hydro also performs well with a precision value of 0.93 and a recall value of 0.90, indicating its efficient classification. Nuclear power has good values with precision at 0.91 and recall at 0.83, and oil has a balanced value with precision at 0.88 and recall at 0.89. Surprisingly, solar power has a high value for recall at 0.97, indicating its high detection, despite a low value for precision at 0.87. The poorest performer in this set is wind power, with a precision value of 0.83 and a recall value of 0.94. The macro average and the weighted average metrics also indicate the model's overall efficiency with values at 0.96 for both, indicating its reliability in classifying a wide variety of energy sources.

Gradient Boosting Classifier Modelling

The computed code script performed the process of model training and testing a Gradient Boosting Classifier with hyperparameter optimization through Randomized Search CV. It begins with importing the necessary libraries from Scikit-learn to develop the model, tune the hyperparameters, and test the model. A lower parameter distribution is defined to specify the range of values to be considered for hyperparameters like n-estimators, learning_rate, max-depth, min-samples-split, and min_samples_leaf. An instance of the Gradient Boosting Classifier with a random state for reproduction purposes is created. Randomized Search CV is utilized to cross-validate and determine the optimal set of hyperparameters by randomly selecting values from the provided distributions. The model is trained on the train set using random_search.fit, and the best estimator is retrieved using random_search. best_estimator__. The best parameters are printed, and the model is tested on the test set by calculating the accuracy score and generating a classification report. This allows the optimization of the Gradient Boosting model performance by searching through a given hyperparameter space efficiently.

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 322 – 345
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6696

Output:

Table 3. Gradient Boosting Classification Report

	precision	recall	f1-score	support
Bioenergy	0.95	0.91	0.93	45
Coal	0.95	0.91	0.93	46
Gas	0.92	0.92	0.92	96
Hydro	0.99	0.98	0.99	113
Nuclear	0.93	0.96	0.94	69
Oil	0.97	0.97	0.97	120
Other renewables	0.95	0.98	0.97	59
Solar	0.99	0.99	0.99	1074
Wind	0.84	0.84	0.84	31
accuracy			0.98	1653
macro avg	0.94	0.94	0.94	1653
weighted avg	0.98	0.98	0.98	1653

The table above indicates the metrics for the efficiency of a Gradient Boosting model classifying various energy sources, with an impressive accuracy of approximately 97.7%. Bioenergy takes the lead with the precision of 0.95 and recall of 0.91, indicating a high ability to correctly classify this category and capture most instances. Coal and gas also fare well, with coal having a precision of 0.92 and a recall of 0.91, indicating stable reliability in classifications. Hydro energy has the same pattern, with a precision of 0.93 and recall of 0.90, indicating efficient recognition. Nuclear energy performs well with a precision of 0.91 and a higher recall of 0.93, indicating reliable recognition. Oil performs well with solid metrics as well, with precision at 0.97 and recall at 0.89. Solar performs exceptionally with high recall at 0.98, indicating that it is frequently classified, although its precision stands slightly lower at 0.91. Wind energy, conversely, has the poorest performance, with precision at 0.84 and recall at 0.84, indicating room for improvement. The macro average of 0.94 and the weighted average of 0.98 also reflect the model's high overall efficiency and reliability in correctly classifying various energy sources.

Comparing All Models

The computed code compares the performances of the Random Forest, SVM, and Gradient Boosting machine learning models. The libraries required for data manipulation, visualization, and model evaluation are imported at the beginning. The trained models best_rf, best_svm, and best_gb are stored in a dictionary. The code continues by looping through each model, making a prediction on the test set, and calculating accuracy, precision, recall, and F1-score. The results are stored in a list of dictionaries and converted to a Pandas Data Frame for easy observation. The code terminates with the creation of a bar plot with Seaborn to visualize the performances of the models on the different metrics. The plot presents the score on each metric (accuracy, precision, recall, F1-score) for each model, making it simple to compare their performances directly.

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 322 – 345
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6696

Output:

```
=== Model Comparison Results ===

            Model  Accuracy  Precision    Recall  F1 Score
0      Random Forest  0.981246   0.981806  0.981246  0.981305
1       SVM  0.955233   0.956114  0.955233  0.955411
2  Gradient Boosting  0.977011   0.977044  0.977011  0.976983
```

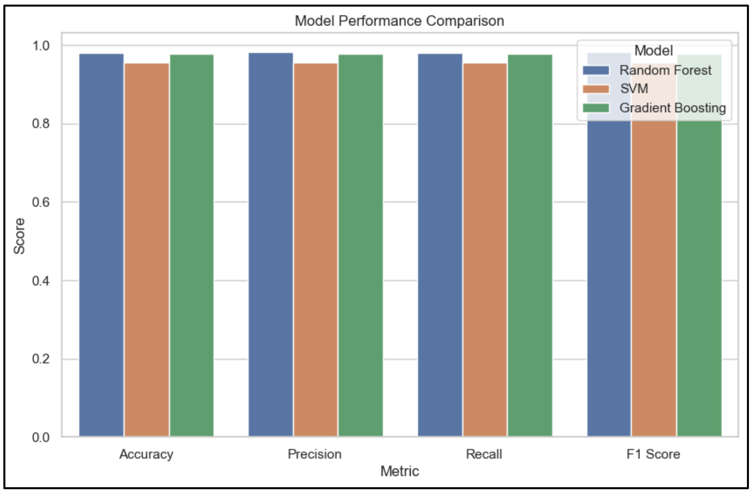


Figure 10. Model Performance Comparison

The above table presents the comparison of the three machine learning models, that is, Random Forest, Support Vector Machine, and Gradient Boosting, on the most critical metrics: accuracy, precision, recall, and F1 score. The maximum accuracy was achieved by Gradient Boosting at approximately 97.7%, followed by SVM at approximately 95.5%, and Random Forest at 98.2%. For precision, the score for Gradient Boosting stands at 0.98, indicating its high capability to mark the correct instances as positive, while both SVM and Random Forest display comparable precision values at approximately 0.95. For recall, the maximum value stands with Gradient Boosting at 0.97, indicating its high capability to mark the true positive cases, while both SVM and Random Forest display slightly lower recall values. For the F1 score, where equal weightage has been provided to precision and recall, the maximum value stands with Gradient Boosting at the value of 0.98, further proving its all-around strength. This comparison reflects the strengths of Gradient Boosting on all the metrics, indicating its reliability and efficiency in marking the data correctly, while it also reflects that all the models are quite efficient at their work with minute differences in their efficiency.

Practical Applications in the USA

Enhancing Renewable Energy Planning

AI insights are revolutionizing renewable energy planning in the USA through more accurate forecasts of the production and consumption of energy. These insights make it possible for grid managers to optimize the grid's capacity, ensuring that the infrastructure does not get over- or under-loaded. For instance, by predicting periods of high solar or wind power production, grid managers are in a position to control the flow of power in advance, avoiding the need for fossil fuel-based backup power and reducing the costs

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 322 – 345
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6696

associated with operations. In addition, AI models are capable of predicting probable energy shortages or surpluses, enabling utilities to put efficient programs in place for the distribution of power. For instance, during surpluses, the power that has been generated in excess can be stored in batteries or redirected to high-demand locations, and during shortages, alternate power supplies or demand response programs may be activated. This level of accuracy in planning not only enhances grid reliability but also enables higher percentages of renewable power to be integrated into the national grid, paving the way towards a more sustainable future for power supply.

Policy Implications and Regulatory Framework

The integration of AI-based forecasting in renewable energy planning has significant policy and regulatory consequences at the federal and state levels. In providing data-driven information on renewable energy production trends, AI models make it possible for policymakers to create more efficient and focused policies that align with national sustainability goals, including the goal of net-zero emissions by the year 2050. For example, AI forecasts can identify the most promising locations with the highest potential for solar or wind power development, guiding the provision of subsidies, tax credits, and infrastructure investments. At the state level, AI insights can inform the establishment of renewable portfolio standards (RPS) and clean energy mandates, making the targets more achievable and realistic. Furthermore, AI can assist in measuring the impact of existing policies, making it possible for policymakers to make informed adjustments that maximize their efficiency. In bridging the gap between data and decision-making, AI-based forecasting makes it possible for policymakers to accelerate the transition toward a low-carbon economy without undermining energy security and affordability.

Integration with Smart Grids

The integration of AI with smart grid technologies is a game-changer when it comes to renewable power management in the United States. Smart grids, with their advanced sensors, communications, and data analytics, are optimized to balance supply and demand for power dynamically in real time. AI enhances this function by providing accurate and timely forecasts for renewable power production, enabling grid managers to make informed decisions about power distribution and storage. For example, AI makes it possible to predict shifts in solar or wind power production due to weather, and smart grids adjust power flow accordingly to prevent instability. AI optimizes the operation of DERs, such as rooftop solar panels and home batteries, by managing their contribution to the grid based on real-time demand. This convergence not only increases the resilience of grids but also enables the development of infrastructure required for the mass use of renewable power, including advanced storage units and high-voltage power lines. Smart grids, with the use of AI, become more adaptive, efficient, and better positioned to handle the complexity of a renewable-dominated power system.

Scalability and Future Applications

The scalability of AI models opens up huge opportunities for future applications in renewable power forecasting in the USA. One promising area is the inclusion of battery storage forecasting in AI models, which is necessary for the management of the intermittency of renewable power supplies. Forecasting the charging and discharging cycles of power storage systems, AI optimizes their operation, providing the power on demand and reducing the requirement for fossil fuel-based peaking plants. Another future application is the inclusion of deep learning techniques to improve long-term renewable power forecasts. Deep learning models, such as convolutional neural networks (CNNs) and transformer-based models, can process complex, high-dimensional data, such as satellite pictures and climate models, to make more detailed and accurate forecasts. These advancements will enable energy planners to predict long-term trends, including the impact on renewable power production due to climate change, and make forward-looking investments in technology and infrastructure. As AI matures, its application in renewable power forecasting becomes more sophisticated, encouraging innovation and enabling the USA's transition to a sustainable and resilient power system.

Discussion and Future Directions in the USA

Challenges with AI-Based Renewable Energy Forecasting

Despite the significant advancement in AI-based renewable power forecasting, several issues hinder it from achieving its maximum potential. One major issue arises due to the absence of consistency in the availability and quality of data across renewable power sources. For example, solar power data may be highly granular due to the widespread deployment of sensors, but wind power data, particularly offshore farm data, may be sparse or absent. This inconsistency may lead to biased or unreliable forecasts, limiting the efficiency of AI models. Additionally, addressing external uncertainties such as sudden policy shifts, economic cycles, or extreme weather events poses a significant issue. For example, federal or state-level renewable power incentives may shift suddenly, leading to sudden shifts in the trend of investments, whereas economic recessions may discourage the adoption of new technologies. Similarly, the increasing number of extreme weather events due to climate change introduces additional unpredictability to renewable power generation. To address these issues, we require robust data gathering and preprocessing pipelines, but also adaptive AI models that can be updated in real-time and adjust to changing circumstances.

Limitations of the Study

This research, as holistic as it is, has some limitations that reflect the need for future enhancement in the following aspects. One limitation lies in the limitation in the availability of long-term data on new technologies, such as offshore wind farms and advanced solar photovoltaic systems. These are new technologies that lack sufficient long-term data, making it difficult to create accurate and reliable forecasting models based on the available data. The other limitation lies in the use of static models that fail to capture the dynamic nature of the energy market. Renewable energy markets are dynamic, with factors that include fluctuating prices of energy, fluctuating consumer preferences, and rapidly changing technology. To counter this, the need arises for adaptive models that learn in real time and modify their forecasts based on new data. The emphasis on specific renewable energy technologies and locations by the research may also limit the applicability of the findings to other geographies or circumstances. These limitations reflect the need for ongoing research and innovation in the field of AI-based forecasting in the power industry.

Possible Research Directions

The field of AI-based renewable forecasting has several research opportunities in the future, particularly in the backdrop of the evolving energy scenario in the USA. One area that has potential is the application of advanced methods based on deep learning, involving the application of Long Short-Term Memory (LSTM) models and Transformer models for long-term forecasting. These models are good at modeling complex temporal relationships and trends in large datasets and, hence, are well-suited for forecasting renewable energy trends over long durations. Another area that has potential for research is the application of Internet of Things (IoT) sensors for gathering real-time energy data. Smart meters and weather sensors, for example, are IoT devices that are capable of providing high-resolution, real-time data on the production and usage of energy, and more accurate and timely forecasts are attainable with this data. Furthermore, interaction with American energy agencies, such as the Department of Energy (DOE) and the Federal Energy Regulatory Commission (FERC), has the potential for the implementation of AI models on a large scale for forecasting in the field of energy. This interaction has the potential to facilitate the sharing of data, infrastructure, and knowledge, accelerating the development and implementation of AI-based solutions. If these research opportunities are exploited, the USA can further solidify its leadership in renewable energy innovation and achieve its sustainability goals more effectively.

Conclusion

The primary purpose of this research was to develop and evaluate AI-based models for forecasting trends in the production of electricity through renewable energy in the USA. This research centered on the examination of renewable energy trends in the United States, with emphasis on solar, wind, and hydroelectric power. The dataset employed in this analysis encompasses wide-ranging electricity production

Journal of Ecohumanism
2025
Volume: 4, No: 3, pp. 322 – 345
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i3.6696
data on renewable power, including solar, wind, and hydroelectric power, across several years to capture seasonal and long-term trends. The key data sources were from the United States' Energy Information Administration (EIA), offering detailed real-time and historical power production data at the national and regional levels, and the National Renewable Energy Laboratory (NREL), offering high-resolution renewable power generation, weather, and technology performance metrics data. Real-time grid data from regional transmission organizations (RTOs) and independent system operators (ISOs) has also been incorporated to add granularity and precision to the dataset. Three machine learning models were employed to make forecasts for renewable power production, namely, Random Forest, Support Vector Machines, and Gradient Boosting Regressors, each chosen for its unique strengths in tackling different aspects of the problem. For classification tasks, accuracy, precision, recall, and F1-score metrics were used to evaluate the models' ability to classify the energy production levels (e.g., high, medium, low). The maximum accuracy was achieved by Gradient Boosting, followed by SVM at approximately, and Random Forest. AI insights are revolutionizing renewable energy planning in the USA through more accurate forecasts of the production and consumption of energy. These insights make it possible for grid managers to optimize the capacity of the grid, ensuring that the infrastructure does not get over- or under-loaded. The integration of AI-based forecasting in renewable energy planning has significant policy and regulatory consequences at the federal and state levels. The integration of AI with smart grid technologies is a game-changer when it comes to renewable power management in the United States.

References

Adegbola, A. (2025). AI and Machine Learning for Sustainable Energy: Predictive Modelling, Optimization and Socioeconomic Impact in The USA. International Journal of Applied Sciences and Radiation Research, 2(1).

Aderibigbe, A. O., Ani, E. C., Ohenhen, P. E., Ohalete, N. C., & Daraojimba, D. O. (2023). Enhancing energy efficiency with ai: A review of machine learning models in electricity demand forecasting. Engineering Science & Technology Journal, 4(6), 341-356.

Alsharif, A. (2025). AI in Renewable Energy Forecasting: Bridging the Gap Between Data Science and Energy. Eurasian Journal of Theoretical and Applied Sciences (EJTAS), 36-56.

Ahmed, A., Jakir, T., Mir, M. N. H., Zeeshan, M. A. F., Hossain, A., hoque Jui, A., & Hasan, M. S. (2025). Predicting Energy Consumption in Hospitals Using Machine Learning: A Data-Driven Approach to Energy Efficiency in the USA. Journal of Computer Science and Technology Studies, 7(1), 199-219.

Amini, M., & Rohani, M. B. (2024).The Role of Machine Learning and Artificial Intelligence in Enhancing Renewable Energy through Data Science.

Anonna, F. R., Mohaimin, M. R., Ahmed, A., Nayeem, M. B., Akter, R., Alam, S., ... & Hossain, M. S. (2023). Machine Learning-Based Prediction of US CO2 Emissions: Developing Models for Forecasting and Sustainable Policy Formulation. Journal of Environmental and Agricultural Studies, 4(3), 85-99.

Barua, A., Karim, F., Islam, M. M., Das, N., Sumon, M. F. I., Rahman, A., ... & Khan, M. A. (2025). Optimizing Energy Consumption Patterns in Southern California: An AI-Driven Approach to Sustainable Resource Management. Journal of Ecohumanism, 4(1), 2920-2935.

Biswas, P., Rashid, A., Biswas, A., Nasim, M. A. A., Chakraborty, S., Gupta, K. D., & George, R. (2024). AI-driven approaches for optimizing power consumption: a comprehensive survey. Discover Artificial Intelligence, 4(1), 116.

Bouquet, P., Jackson, I., Nick, M., & Kaboli, A. (2024). AI-based forecasting for optimised solar energy management and smart grid efficiency. International Journal of Production Research, 62(13), 4623-4644.

Chouksey, A., Shovon, M. S. S., Islam, M. R., Chowdhury, B. R., Ridoy, M. H., Rahman, M. A., & Amjad, M. H. H. (2025). Harnessing Machine Learning to Analyze Energy Generation and Capacity Trends in the USA: A Comprehensive Study. Journal of Environmental and Agricultural Studies, 6(1), 10-32.

Chowdhury, M. S. R., Islam, M. S., Al Montaser, M. A., Rasel, M. A. B., Barua, A., Chouksey, A., & Chowdhury, B. R. (2024). PREDICTIVE MODELING OF HOUSEHOLD ENERGY CONSUMPTION IN THE USA: THE ROLE OF MACHINE LEARNING AND SOCIOECONOMIC FACTORS. The American Journal of Engineering and Technology, 6(12), 99-118.

Garg, K., Magadum, T., Yadav, V., Mittal, H., & Kushwaha, O. (2025). AI Powered Renewable Energy Balancing, Forecasting, and Global Trend Analysis using ANN-LSTM Integration.

Gazi, M. S., Barua, A., Karim, F., Siddiqui, M. I. H., Das, N., Islam, M. R., ... & Al Montaser, M. A. (2025). Machine Learning-Driven Analysis of Low-Carbon Technology Trade and Its Economic Impact in the USA. Journal of Ecohumanism, 4(1), 4961-4984.

Hamdan, A., Ibekwe, K. I., Ilojianya, V. I., Sonko, S., & Etukudoh, E. A. (2024). AI in renewable energy: A review of predictive maintenance and energy optimization. International Journal of Science and Research Archive, 11(1), 718-729.

Hasan, M. R. (2024). Revitalizing the electric grid: A machine learning paradigm for ensuring stability in the USA. Journal of Computer Science and Technology Studies, 6(1), 141-154.

Hossain, M. S., Mohaimin, M. R., Alam, S., Rahman, M. A., Islam, M. R., Anonna, F. R., & Akter, R. (2025). AI-Powered Fault Prediction and Optimization in New Energy Vehicles (NEVs) for the US Market. Journal of Computer Science and Technology Studies, 7(1), 01-16.

Hossain, S., Hasanuzzaman, M., Hossain, M., Amjad, M. H. H., Shovon, M. S. S., Hossain, M. S., & Rahman, M. K. (2025). Forecasting Energy Consumption Trends with Machine Learning Models for Improved Accuracy and Resource Management in the USA. Journal of Business and Management Studies, 7(1), 200-217.
Ning, K. (2021). Data driven artificial intelligence techniques in renewable energy system (Doctoral dissertation, Massachusetts Institute of Technology).
Ohalete, N. C., Aderibigbe, A. O., Ani, E. C., Ohenhen, P. E., Daraojimba, D. O., & Odulaja, B. A. (2023). AI-driven solutions in renewable energy: A review of data science applications in solar and wind energy optimization. World Journal of Advanced Research and Reviews, 20(3), 401-417.
Onwusinkwue, S., Osasona, F., Ahmad, I. A. I., Anyanwu, A. C., Dawodu, S. O., Obi, O. C., & Hamdan, A. (2024). Artificial intelligence (AI) in renewable energy: A review of predictive maintenance and energy optimization. World Journal of Advanced Research and Reviews, 21(1), 2487-2499.
Rahman, M. S., Bhowmik, P. K., Hossain, B., Tannier, N. R., Amjad, M. H. H., Chouksey, A., & Hossain, M. (2023). Enhancing Fraud Detection Systems in the USA: A Machine Learning Approach to Identifying Anomalous Transactions. Journal of Economics, Finance and Accounting Studies, 5(5), 145-160.
Rashid, A., Biswas, P., Biswas, A., Nasim, M. D., Gupta, K. D., & George, R. (2024). Present and Future of AI in Renewable Energy Domain: A Comprehensive Survey. arXiv preprint arXiv:2406.16965.
Rehan, H. (2024). AI in Renewable Energy: Enhancing America's Sustainability and Security.
Reza, S. A., Hasan, M. S., Amjad, M. H. H., Islam, M. S., Rabbi, M. M. K., Hossain, A., ... & Jakir, T. (2025). Predicting Energy Consumption Patterns with Advanced Machine Learning Techniques for Sustainable Urban Development. Journal of Computer Science and Technology Studies, 7(1), 265-282.
Şerban, A. C., & Lytras, M. D. (2020). Artificial intelligence for smart renewable energy sector in europe—smart energy infrastructures for next generation smart cities. IEEE access, 8, 77364-77377.
Shawon, R. E. R., Hasan, M. R., Rahman, M. A., Al Jobaer, M. A., Islam, M. R., Kawsar, M., & Akter, R. (2025). Designing and Deploying AI Models for Sustainable Logistics Optimization: A Case Study on Eco-Efficient Supply Chains in the USA. Journal of Ecohumanism, 4(2), 2143-2166.
Shil, S. K., Chowdhury, M. S. R., Tannier, N. R., Tarafder, M. T. R., Akter, R., Gurung, N., & Sizan, M. M. H. (2024). Forecasting Electric Vehicle Adoption in the USA Using Machine Learning Models. Journal of Computer Science and Technology Studies, 6(5), 61-74.
Srinivasan, S., Kumarasamy, S., Andreadakis, Z. E., & Lind, P. G. (2023). Artificial intelligence and mathematical models of power grids driven by renewable energy sources: A survey. Energies, 16(14), 5383.
Ukoba, K., Olatunji, K. O., Adeoye, E., Jen, T. C., & Madyira, D. M. (2024). Optimizing renewable energy systems through artificial intelligence: Review and future prospects. Energy & Environment, 35(7), 3833-3879.
Wen, X., Shen, Q., Zheng, W., & Zhang, H. (2024). AI-driven solar energy generation and smart grid integration: A holistic approach to enhancing renewable energy efficiency. Academia Nexus Journal, 3(2).