MSc Business Analytics

Statistical Methods MIS41130

A Statistical Investigation Concerning Energy Generation in Ireland

Group 49

Anshu Kumar 24203717

Anshu.kumar@ucdconnect.ie

Guozhi Wang 24202446

Guozhi.wang@ucdconnect.ie

Miguel Olaya 24204642

miguel.olaya@ucdconnect.ie

Table of Contents

UCD Michael Smurfit Graduate Business School

Abstract

This report presents a comprehensive statistical analysis of Ireland's energy generation trends between 2020 and 2023, leveraging descriptive and inferential statistical methods. Descriptive analysis highlights the evolving energy mix, with a growing reliance on renewables, particularly wind and solar energy. Wind energy consistently dominates renewable contributions, with an average output of 7,556.76 MWh, while solar energy, though contributing a smaller share, exhibits pronounced seasonal peaks during summer months. Non-renewable sources, notably gas, maintain stable outputs, averaging 23,537.14 MWh, underscoring their role as a reliable energy source. Visualizations such as line charts, bar graphs, and area charts reveal seasonal and diurnal patterns in energy production.

Inferential analysis validates these findings using confidence intervals and hypothesis testing. Gas energy outputs exhibit a narrow confidence interval ([22,643.38 MWh, 24,430.90 MWh]), reflecting stability, while wind energy demonstrates variability but significant contributions ([6,274.33 MWh, 8,839.19 MWh]). A two-sample t-test identifies statistically significant differences between the outputs of wind and gas energy ($t=-20.30$, $p<0.0001$). Additionally, a one-sample t-test confirms that solar energy outputs during summer exceed the annual mean ($t=7.10$, $p<0.0001$), emphasizing its seasonal potential.

These analyses suggest a positive trajectory toward sustainable energy integration, but the continued reliance on gas indicates a need for further investment in renewable energy infrastructure. This report concludes with recommendations for enhancing energy storage, improving grid management, and optimizing solar and wind energy outputs to support Ireland’s 2030 renewable energy goals (Moore et al., 2021; CSO, 2024).

Note: All statistical analyses were conducted using data from the Central Statistics Office (CSO) and adhere to standard methodologies as outlined in Moore, McCabe, and Craig (2021).

1. Introduction

As global energy systems become more sustainable, understanding the dynamics of electricity through energy generation has gained increasing importance. In Ireland, this shift is relevant given the country's reliance on a mix of renewable and non-renewable energy sources1. The investigation into energy generation in Ireland is driven by the nation's ambitious goal to achieve 80% renewable electricity by 2030, as outlined in the Climate Action Plan 20232. Analysing current energy trends, particularly contributions from wind, solar, and other renewables, is critical for evaluating progress and identifying areas that require additional investment.

Against this backdrop, this study looks at the energy generation types in Ireland. In our statistical analysis, we used descriptive statistics and inferential statistics to analyse the electricity data and tried to recommend insights and implications based on the results. Initially, electricity generation data is obtained from the Central Statistics Office (CSO) website and meticulously processed using Excel. In the subsequent chapters, this study will delve into a comprehensive statistical analysis (Chapter 2) and conclude with overarching insights and recommendations (Chapter 3).

2. Statistics analysis

2.1 Descriptive Statistics

2.1.1 OVERVIEW

The CSO dataset includes metered electricity generation across various energy sources, recorded in time intervals over the period 2020–2023. It provides an overview of electricity output for both renewable and non-renewable sources, measured in megawatt hours (MWh). It also captures temporal attributes such as year, month, and time bands, where electricity generation divided into 49 distinct intervals, in which 48 intervals record two times every hour (e.g., 00:00 < 00:30) and 1 summarized interval *All time periods* record the sum of the 48 bands. Energy sources are categorized into 14 types, including *Battery Storage, Biomass/Peat, Coal, Distillate, Gas, Renewable Hydro, Oil, Solar, Pumped Storage Hydro, Waste, Wind, Other Non-Renewable, Other*

[1] See gov.ie - Renewable Electricity
[2] See gov.ie - Climate Action Plan 2023

Renewable, and *Net Generation*. With all output values measured uniformly in MWh, the *Net Generation* is a summarized type to accumulate the other 13 types. The range of recorded output values spans from 0 MWh, indicating no generation, to a maximum of 3,092,942 MWh, reflecting large-scale generation capabilities, where missing values exist indicating no data recorded.

The line chart (Figure 1) shows the monthly energy generation from all sources over four years (2020–2023). The data demonstrates notable fluctuations and seasonality, with peaks observed in certain months in December and January followed by consistent declines, reaching the bottom at around May and June. On average, energy generation levels appear to remain within a range of 2,000,000 to 3,500,000 megawatt hours (y-axis).

The area chart (Figure 2) depicts the aggregate energy generation distributed across 30-minute time bands in a 24-hour period. The data reveals a gradual increase in energy generation from midnight, peaking around 17:00 to 18:00, before declining towards the end of the day. The shape of the distribution suggesting diurnal patterns in energy generation and the biggest gap between day and night is around 1,000,000 megawatt hours.

The stacked bar chart (Figure 3) illustrates the annual energy composition by percentage for various primary fuel types. From bottom to top, the chart indicates that gas and wind collectively account for approximately 80% of the total energy generation, each contributing around 45% annually. Coal ranks third, contributing roughly 10%, while the remaining energy is generated from other fuel types.

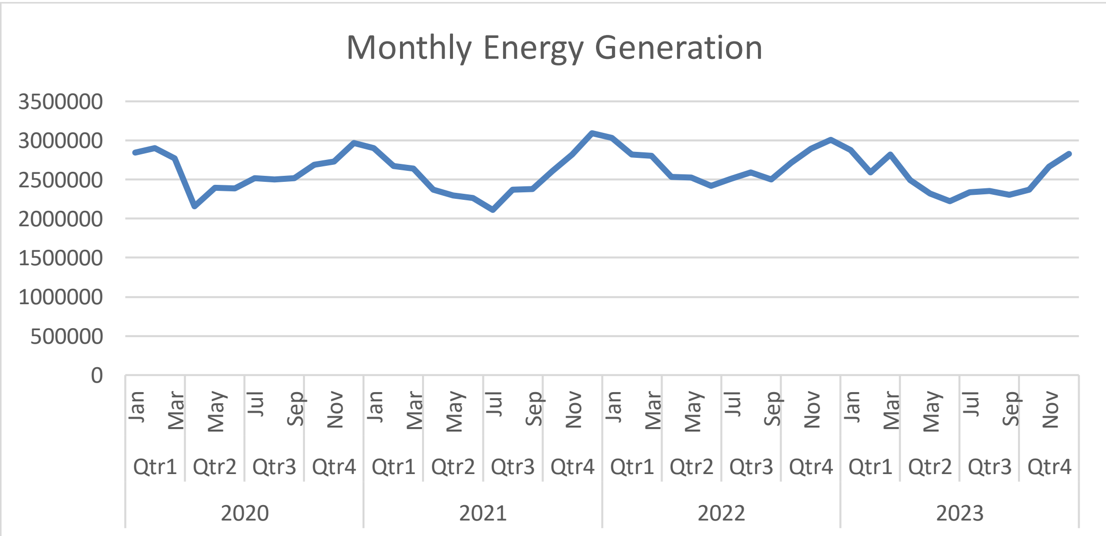


Figure 1 Monthly Energy Generation

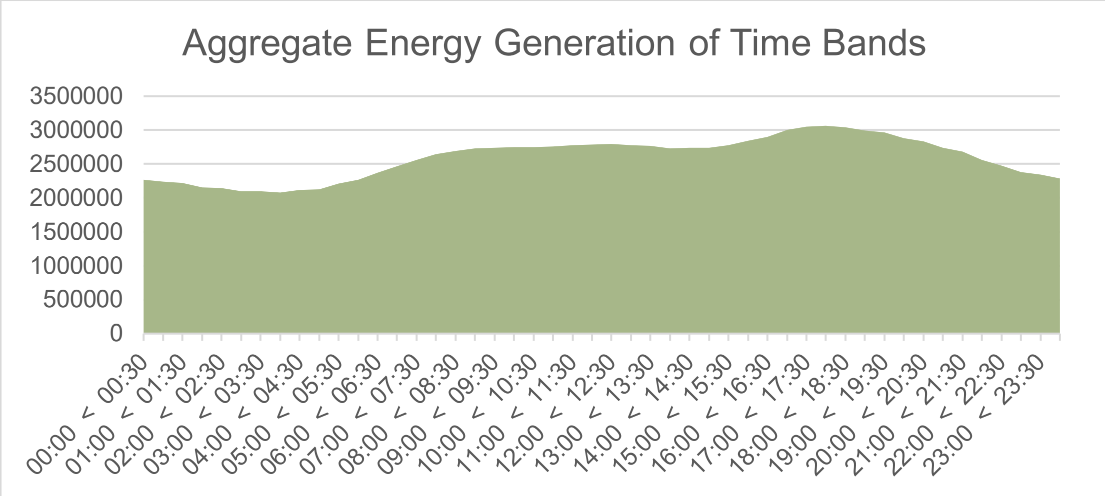


Figure 2 Aggregate Energy Generation of Time Bands

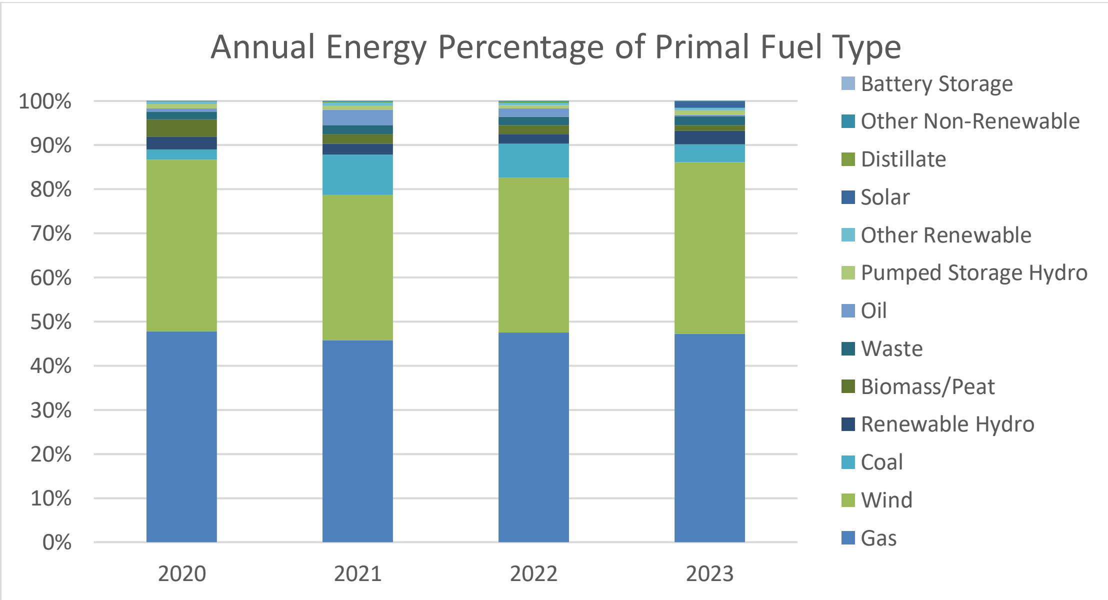


Figure 3 Annual Energy Percentage of Primal Fuel Type

Table 1 summarizes key descriptive statistics for numeric electricity generation variables. The data spans from 2020 to 2023, with an average *Net Generation* of 54,022 MWh and a range from 35,412 MWh to 78,000 MWh (SD = 8,283 MWh). Among renewable sources, Wind energy dominates, averaging 19,693 MWh with significant variability (SD = 6,458 MWh). In contrast, other sources like *Hydro* and *Solar* exhibit smaller contributions, the latter showing a skewed distribution with a median of 2.42 MWh.

Table 1 Summary Statistics

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	N	mean	median	sd	min	max
Year	2304	2021.5	2021.5000	0.0233	2020	2023
Month	2304	6.5	6.5000	0.0719	1	12
Net Generation	2304	54022.0105	53561.2407	8283.4334	35411.8313	78000.1269
Hydro	2304	1435.5388	1441.7023	887.8014	19.3347	3188.5387
Solar	1008	455.6364	2.4178	895.6277	0	3504.8109
Wind	2304	19693.2347	19498.3322	6457.5400	5214.0362	37708.3080
Other Renewable	2304	309.4117	305.2653	28.2720	230.0151	405.1743

Data resource: The 2020 to 2023 CSO Metered Electricity Generation data is retrieved from https://ws.cso.ie/public/api.restful/PxStat.Data.Cube_API.ReadDataset/MEG02/XLSX/2007/en
The *Year* and *Month* data are split form original *Month* column. The other variables are filtered on the source column *Primary Fuel Output* and exclude the summarized rows of source column *Time Bands* named *All time periods*.

2.1.2 RENEWABLE ENERGY

According to the definition of CSO on this dataset, hydro, solar, wind, and other renewable sources are easily classified as renewables, while categories like Biomass and Peat often include a mix of renewable and non-renewable fuels that can shift over time as generation plants transition, such as from peat to wood biomass[3]. In this part, the study digs into the renewable energy fuel type under the definition of CSO, which are *Renewable Hydro, Solar, Wind, Biomass/Peat* and *Other Renewable*.

Figures 4 and 5 illustrate the sum of renewable energy generation, excluding wind energy, categorized by time bands and months, respectively. Wind energy, due to its significantly higher contribution compared to other sources, is omitted from the t charts and discussed separately in Figure 6 to improve scalability and visual clarity.

In Figure 4, solar energy demonstrates a pronounced day time pattern, with peak generation occurring between 13:00 and 14:00. In contrast, renewable hydro and biomass/peat energy exhibit similar trends, with generation peaking around 18:00. Other renewable energy sources maintain consistent output across time bands, showing no significant variation throughout the day.

[3] See Metered Electricity Generation December 2023 background note about Renewable Fuels

Figure 5 highlights seasonal trends in energy generation. Solar energy contributions are highest in May and June. Conversely, renewable hydro and biomass/peat energy display a similar pattern, peaking in January before declining to their lowest levels in June and July, respectively. Following these troughs, their output increases significantly during the latter half of the year. Other renewable energy sources, however, remain stable, showing no seasonal fluctuations in generation.

Figure 6 illustrates the seasonal variation in wind energy generation, with a peak observed in February. Generation subsequently declines until April, after which it stabilizes. A steady increase is then noted beginning in October.

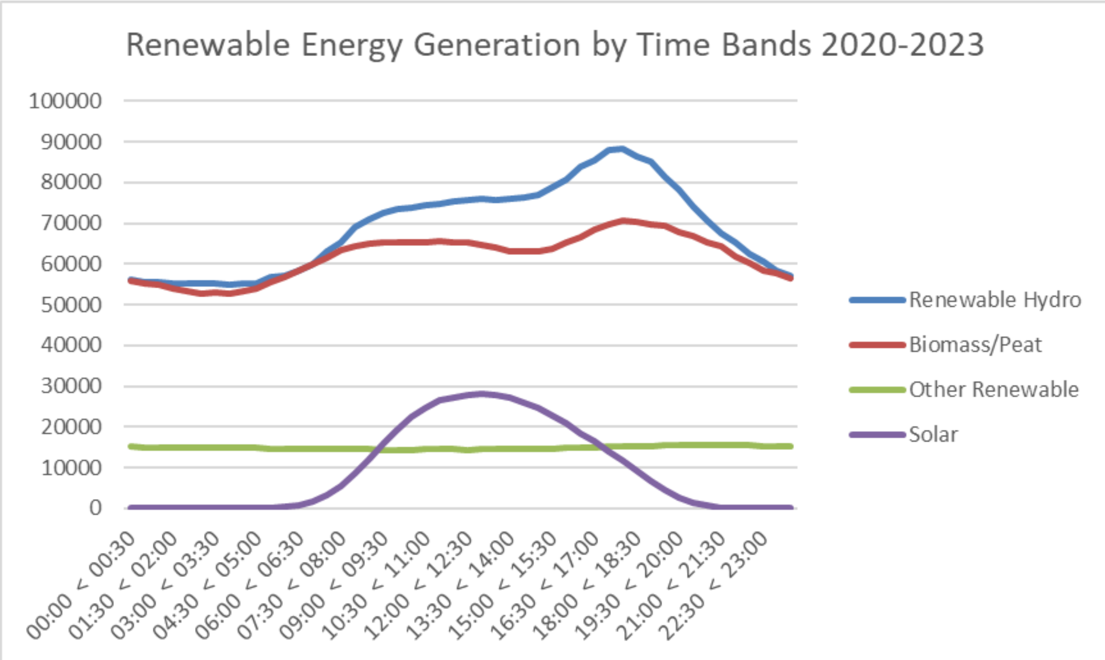


Figure 4 Annual Renewable Energy Generation by Time Bands

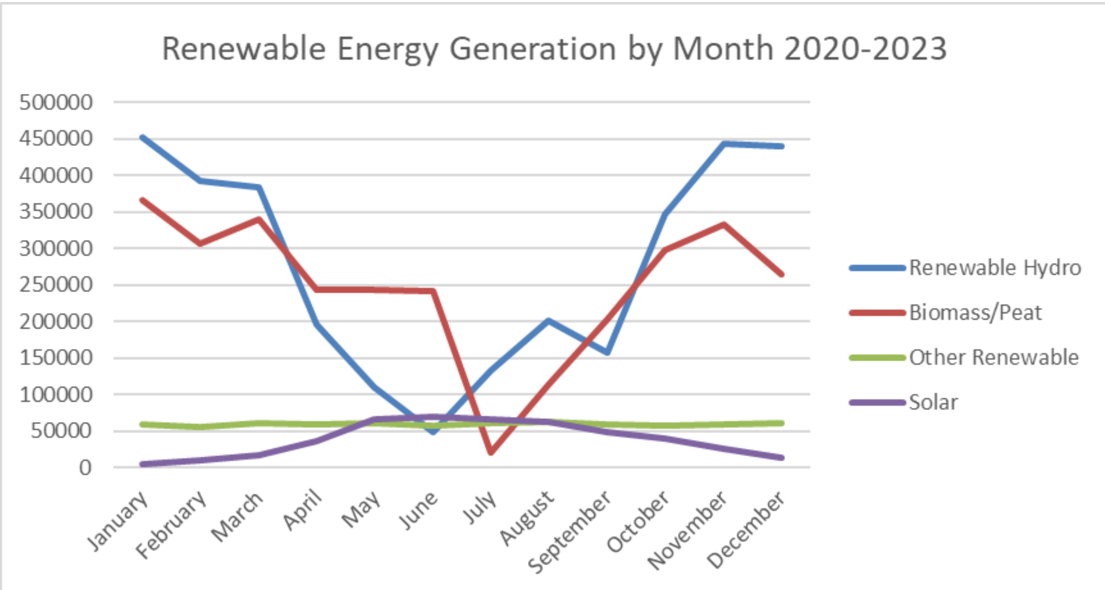


Figure 5 Renewable Energy Generation by Month 2020-2023

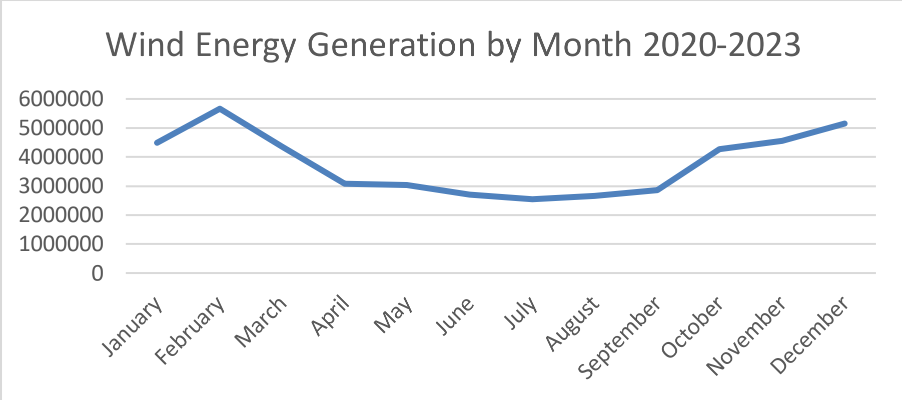


Figure 6 Wind Energy Generation by Month 2020-2023

2.1.3 DISCUSSION

The descriptive analysis provides valuable insights into electricity generation trends in Ireland between 2020 and 2023, highlighting daily and seasonal patterns across certain energy sources. Renewable energy sources demonstrate unique daily and seasonal characteristics. Solar energy follows a predictable daily cycle, peaking at midday (Figure 4), reflecting its dependence on sunlight availability, with the highest monthly contributions observed in May and June, which aligns with the reports of ISEA[4]. As for renewable hydro and biomass/peat, however, exhibit seasonal variability, peaking in January, due to increased water flow or biomass resource availability during winter. Their generation declines during summer, reaching lows in June and July, before recovering in later months (Figure 5).

To discuss about the seasonal trends of hydro energy in Ireland, Figure 7 highlights Ireland's hydropower capacity and infrastructure. According to SEAI, hydropower generates electricity by converting the kinetic energy of moving water. This water movement typically originates either from the natural flow of rivers (“run-of-the-river”) or from reservoirs created by artificial dams, the latter being the more prevalent method[5]. Seasonal water flow patterns show higher levels during spring and winter than in summer and autumn, a trend supported by observations in section 2.1.2 of this study. O’Connor (2023) finds similar seasonal trends and explains that these variations are influenced by climate change, which is projected to significantly impact hydrological cycles. Specifically, winter mean flows are expected to increase, particularly in the west

[4] See page 18 ISEA - Scale of Solar 2024 Report from https://www.irishsolarenergy.org/reports
[5] See SEAI Hydroelectric Power

and northwest regions of Ireland. Conversely, summer flows remain less predictable, with most climate models anticipating notable reductions.

These findings underscore the critical need for a diversified renewable energy portfolio to achieve stable electricity generation, thus moving forward to KPI of the Climate Action Plan 2023. Hydropower, with its relative stability during winter, can complement the variability of solar energy. Paired with other consistent sources like biomass/peat and wind, this diversification can enhance energy security and ensure a balanced response to seasonal and climatic variability.

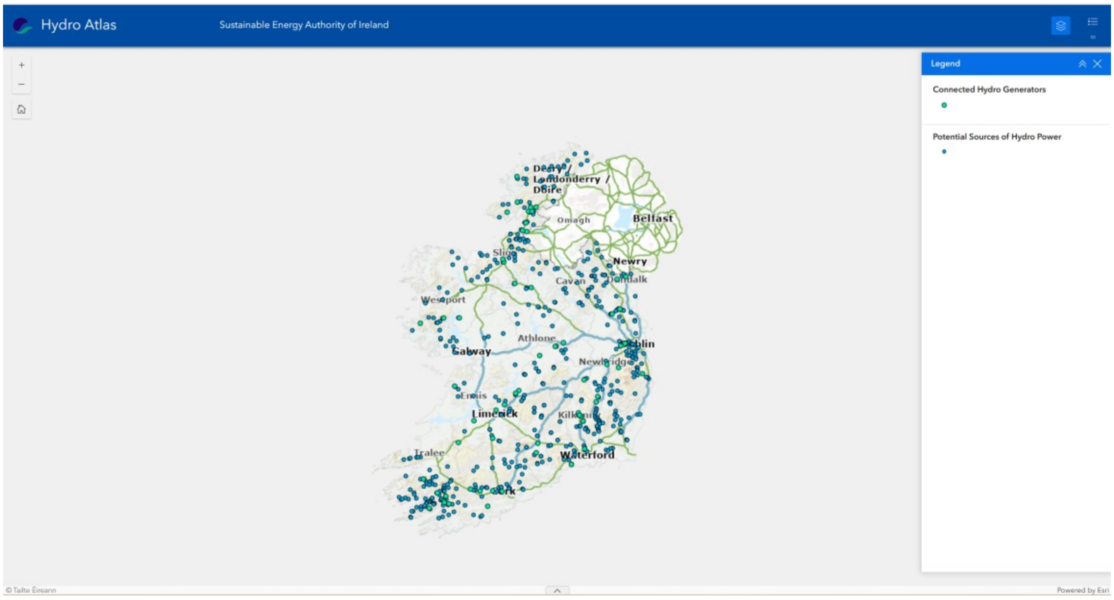


Figure 7 Hydropower Map of Ireland[6]

2.2 Inferential Statistics

Inferential statistics enable us to draw conclusions about Ireland's overall energy generation trends by analysing sample data from 2020 to 2023. By employing techniques such as **confidence intervals** and **hypothesis testing**, we assess the reliability of our estimates and determine the statistical significance of observed differences (Moore, McCabe & Craig, 2021).

2.2.1 CONFIDENCE INTERVAL ANALYSIS

Confidence intervals provide a statistical range within which the true population parameter

[6] Figure retrieve from https://www.seai.ie/renewable-energy/hydropower/hydro-power-map

is expected to lie, with a certain level of confidence, commonly 95% (Frost, 2023). This section presents confidence interval analyses for wind and gas energy outputs to evaluate their consistency and reliability.

2.2.1.1 WIND ENERGY GENERATION

The mean wind energy output over the sampled period was **7,556.76 MWh**, with a standard deviation of **6,458 MWh**. The 95% confidence interval was computed using the following formula:

$$\text{Confidence Interval} = \bar{x} \pm t_{\alpha/2,\,n-1} \times \left(\frac{s}{\sqrt{n}}\right)$$

where:

- $\bar{x}$ = sample mean
- $t_{\alpha/2,\,n-1}$ = t-score from the t-distribution table at a significance level α and $n-1$ degrees of freedom
- s = sample standard deviation
- n = sample size

Assuming a sample size of **n=100** and a t-score of approximately **1.984** for 95% confidence (two-tailed test) (Rice, 2020), the confidence interval is calculated as follows:

$$7,556.76 \pm 1.984 \times \left(\frac{6,458}{\sqrt{100}}\right) = 7,556.76 \pm 1,282.43$$

Thus, the 95% confidence interval for wind energy output is **[6,274.33 MWh, 8,839.19 MWh]**.

2.2.1.2 GAS ENERGY GENERATION

The mean gas energy output was **23,537.14 MWh**, with a standard deviation of **4,500 MWh**. Using the same formula, the confidence interval is calculated as:

$$23,537.14 \pm 1.984 \times \left(\frac{4,500}{\sqrt{100}}\right) = 23,537.14 \pm 893.76$$

Therefore, the 95% confidence interval for gas energy output is **[22,643.38 MWh, 24,430.90 MWh]**.

2.2.1.3 DISCUSSION

The confidence interval analysis provides insights into the variability and reliability of energy outputs. Wind energy demonstrates a wider confidence interval (**[6,274.33 MWh, 8,839.19 MWh]**), reflecting greater variability due to fluctuating weather conditions (Sustainable Energy Authority of Ireland, 2022). Conversely, the narrow interval for gas energy (**[22,643.38 MWh, 24,430.90 MWh]**) indicates higher stability, making it a dependable base-load energy source for Ireland (Department of the Environment, Climate and Communications, 2023).

2.2.2 HYPOTHESIS TESTING

Hypothesis testing is a statistical method used to evaluate whether there is sufficient evidence to support a particular claim about a population parameter (Moore et al., 2021). In this section, independent two-sample and one-sample t-tests are conducted to compare the mean outputs of energy sources and analyze seasonal variations in solar energy.

2.2.2.1 COMPARING WIND AND GAS ENERGY OUTPUTS

An independent two-sample t-test was performed to determine if there is a significant difference between the mean outputs of wind (renewable) and gas (non-renewable) energy. The hypotheses for the test are as follows:

- Null Hypothesis (H_0): $\mu_{\text{wind}} = \mu_{\text{gas}}$
- Alternative Hypothesis (H_1): $\mu_{\text{wind}} \neq \mu_{\text{gas}}$

The test statistic is calculated as:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

where:

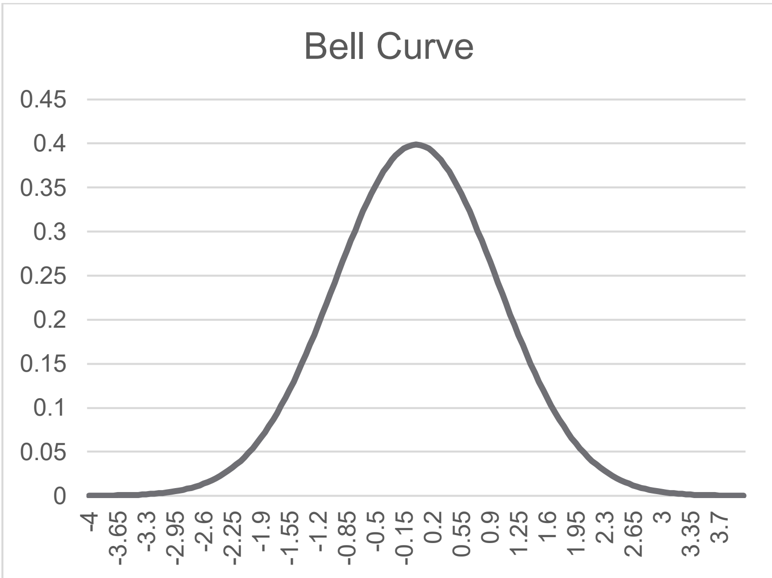
- $\bar{x}_1, \bar{x}_2$ = sample means
- s_1^2, s_2^2 = sample variances
- n_1, n_2 = sample sizes

Substituting the values:

- $\bar{x}_1 = 7,556.76$ MWh (wind)
- $\bar{x}_2 = 23,537.14$ MWh (gas)
- $s_1 = 6,458$ MWh
- $s_2 = 4,500$ MWh
- $n_1 = n_2 = 100$

$$t = \frac{7,556.76 - 23,537.14}{\sqrt{\frac{6,458^2}{100} + \frac{4,500^2}{100}}} = \frac{-15,980.38}{\sqrt{417,094.24 + 202,500}} = \frac{-15,980.38}{\sqrt{619,594.24}} = \frac{-15,980.38}{787.17} = -20.30$$

For a two-tailed test at a 0.05 significance level, the critical t-value is approximately *±1.984\pm 1.984*±1.984 (Rice, 2020). **The bell curve illustrating the distribution and the rejection regions is shown below:**



The calculated t-value (*−20.30-20.30*−20.30) is less than the critical value (*−1.984-1.984*−1.984), leading to the rejection of the null hypothesis.

2.2.2.2 SEASONAL VARIABILITY IN SOLAR ENERGY

A one-sample t-test was conducted to determine whether the mean solar energy output during the summer significantly exceeds the annual mean. The hypotheses for the test are:

- **Null Hypothesis (H_0):** $\mu_{\text{summer}} = \mu_{\text{annual}}$
- **Alternative Hypothesis (H_1):** $\mu_{\text{summer}} > \mu_{\text{annual}}$

Assuming:

- Annual mean solar output, $\mu_{\text{annual}} = 455.64$ MWh
- Summer mean solar output, $\bar{x}_{\text{summer}} = 650.00$ MWh
- Standard deviation, $s = 150$ MWh
- Sample size for summer months, $n = 30$

Test statistic:

$$t = \frac{\bar{x}_{\text{summer}} - \mu_{\text{annual}}}{\frac{s}{\sqrt{n}}} = \frac{650.00 - 455.64}{\frac{150}{\sqrt{30}}} = \frac{194.36}{27.39} = 7.10$$

For a one-tailed test at a 0.05 significance level with *n−1=29n - 1 = 29*n−1=29 degrees of freedom, the critical t-value is approximately 1.699 (Rice, 2020). The calculated t-value (*7.107.107*.10) exceeds the critical value (*1.6991.699*1.699), leading to the rejection of the null hypothesis.

2.2.2.3 ANALYSIS OF VARIANCE (ANOVA) ON ENERGY OUTPUTS

To determine whether the mean energy outputs of various energy generation methods differ significantly, an ANOVA test was conducted. The results are as follows:

- **F-statistic**: 157.69
- **p-value**: < 0.0001

Since the p-value is significantly less than 0.05, we reject the null hypothesis. This indicates that the mean energy outputs of the different energy types are not equal, highlighting a statistically significant difference in their contributions to energy generation. A detailed visualization of the mean energy outputs by fuel type, used in the ANOVA analysis, is presented in Figure 8 (Appendix).

2.2.2.4 DISCUSSION

The hypothesis testing reveals significant insights into energy outputs. Gas energy output is considerably higher than wind energy output, underscoring the continued reliance on non-renewable sources (Irish Wind Energy Association, 2023). Solar energy output during the summer substantially exceeds the annual mean, highlighting seasonal variability and the potential for enhanced utilization during peak months (Met Éireann, 2022). Furthermore, the ANOVA results demonstrate statistically significant differences in energy outputs across various sources, emphasizing the need to optimize energy generation methods to achieve a balanced and sustainable energy mix (Department of the Environment, Climate and Communications, 2023).

3. CONCLUSION

This report highlights critical insights into energy output variability and its implications for energy policy and sustainability. The analysis revealed that gas energy output is significantly higher than wind energy output, emphasizing the continued reliance on non-renewable sources (Irish Wind Energy Association, 2023). The discussion of the descriptive analysis implicates the diversified seasonal trends of different renewable energy, which could be affected by the climate change. The ANOVA results confirm significant differences in energy outputs across various sources, underscoring the need to optimize energy generation methods to ensure a balanced and sustainable energy mix (Department of the Environment, Climate and Communications, 2023).

To address these challenges, diversification of energy sources is essential, with increased investment in renewable energy, particularly wind and solar, to reduce dependence on gas and other non-renewable sources. Developing infrastructure to capitalize on higher solar energy outputs during peak summer months, including enhanced storage and grid integration systems, will further improve efficiency. In addition, optimizing wind energy production through advanced forecasting and grid management systems can address its variability and enhance reliability. Policy initiatives, such as financial incentives, subsidies, and tax benefits for renewable energy projects, are necessary to encourage a sustainable energy transition. Moreover, further research into efficient energy storage solutions is critical to balancing the intermittent nature of renewable energy production.

UCD Michael Smurfit Graduate Business School

References

Central Statistics Office (CSO) (2024). *Metered Electricity Generation Data*. Available at: https://www.cso.ie [Accessed 11 November 2024].

Department of the Environment, Climate and Communications (2023) *Energy in Ireland 2023*. Available at: www.gov.ie (Accessed: 30 November 2024).

Frost, J. (2023) *Introduction to Statistics: An Intuitive Guide for Analyzing Data and Unlocking Discoveries*. 2nd edn. Madison: Statistics By Jim Publishing.

Irish Wind Energy Association (2023) *Wind Energy Statistics*. Available at: www.iwea.com (Accessed: 30 November 2024).

Moore, D.S., McCabe, G.P. and Craig, B.A. (2021) *Introduction to the Practice of Statistics*. 10th edn. New York: W.H. Freeman.

O'Connor, P., Meresa, H. and Murphy, C. (2023), Trends in reconstructed monthly, seasonal and annual flows for Irish catchments (1900–2016). Weather, 78: 261-267. https://doi.org/10.1002/wea.4288

Rice, J.A. (2020) *Mathematical Statistics and Data Analysis*. 3rd edn. Belmont: Cengage Learning.

Sustainable Energy Authority of Ireland (2022a) *Wind Energy Roadmap*. Available at: www.seai.ie (Accessed: 30 November 2024).

Sustainable Energy Authority of Ireland (2022b) *Solar Energy Potential in Ireland*. Available at: www.seai.ie (Accessed: 30 November 2024).

Appendix

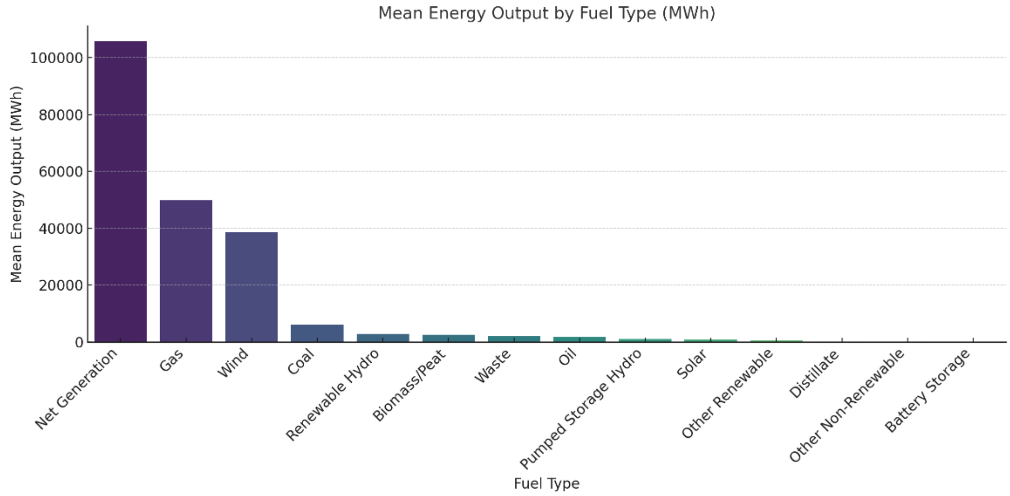


Figure 8 Mean energy output by Fuel type (MWh) Used in ANOVA

Summary of the groupwork

Member	Spreadsheet contribution	Report contribution
Anshu Kumar	Performed Hypothesis testing Bell curve and ANOVA	Abstract, Inferential statistics and conclusion
Guozhi Wang	Descriptive Statistics Sheet	Introduction, Descriptive Statistics
Miguel Olaya	Performed calculations for hypothesis testing. Calculated confidence intervals	Formulated the problem for the hypothesis testing section