

Renewable Electricity and Carbon Intensity in the UK

100428943

Introduction

The decarbonisation of electricity generation is a central component of the United Kingdom's transition towards a low-carbon economy. Over the past two decades, the UK has experienced a substantial expansion in renewable electricity generation alongside a marked decline in the carbon intensity of its power grid. Understanding the relationship between these developments is important for evaluating the role of renewable energy in reducing emissions from the electricity sector. This project examines whether increases in renewable electricity generation are associated with measurable reductions in the carbon intensity of electricity generation in the United Kingdom.

Using annual time-series data covering the period 2000–2024, the analysis combines information on the UK's electricity generation mix with data on electricity-sector carbon intensity obtained from Our World in Data (Ritchie, Roser and Rosado, 2024). The empirical approach proceeds in three stages. First, renewable electricity shares are constructed from disaggregated generation data by source. Second, descriptive analysis is used to document trends in renewable deployment and carbon intensity over time. Finally, regression analysis is employed to assess the relationship between renewable electricity shares and carbon intensity, alongside a set of robustness and diagnostic checks..

Data and Variable Construction

This analysis draws on two datasets obtained from Our World in Data (Ritchie, Roser and Rosado, 2024). The first dataset provides annual measures of the carbon intensity of electricity generation, expressed in grams of CO₂ per kilowatt-hour. The second dataset contains annual electricity generation in the United Kingdom disaggregated by source, including wind, solar, hydropower, fossil fuels, nuclear energy, and net electricity imports.

The electricity generation dataset is restricted to observations for the United Kingdom and is used to construct a measure of renewable electricity generation. Consistent with the categorisation used in Our World in Data, renewable generation is defined as the sum of electricity generated from wind, solar, hydropower, and other renewable sources (Ritchie, Roser and Rosado, 2024). Total electricity generation is calculated as the sum of renewable generation, fossil fuel generation, nuclear generation, and net electricity imports. The renewable share of electricity generation is then defined as renewable generation divided by total generation.

The carbon intensity dataset is similarly restricted to UK observations and merged with the electricity generation data by year. As carbon intensity data are available from 2000

onwards, the final analysis sample covers the period 2000–2024. Observations with missing values are removed prior to analysis.

```
analysis_data_clean <- electricity_uk %>%  
  dplyr::select(year, renewable_share) %>%  
  dplyr::left_join(carbon_uk, by = "year")
```

Descriptive Analysis

This section provides a descriptive overview of trends in renewable electricity generation and carbon intensity in the United Kingdom. Examining these trends over time helps to establish whether changes in the electricity generation mix coincide with changes in carbon intensity, prior to formal regression analysis.

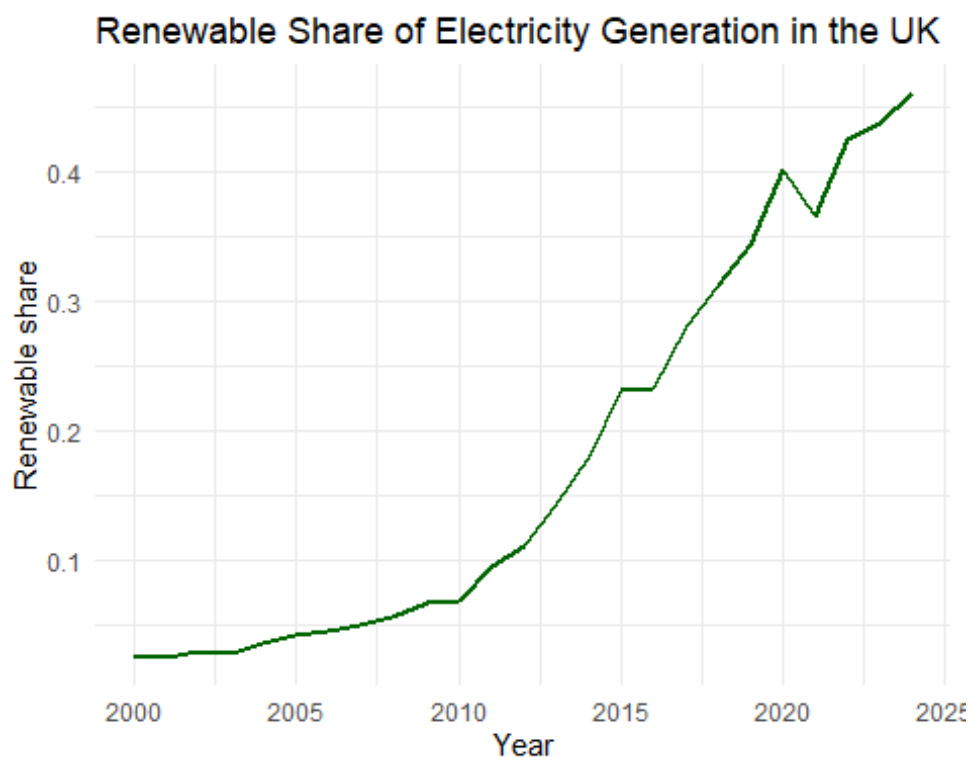


Figure 1 shows a pronounced upward trend in the share of electricity generated from renewable sources over the sample period. Renewable penetration remains very low in the early 2000s but increases rapidly from the mid-2000s onwards, reflecting the expansion of wind and solar generation in the UK. By the end of the sample period, renewable electricity accounts for a substantial proportion of total generation.

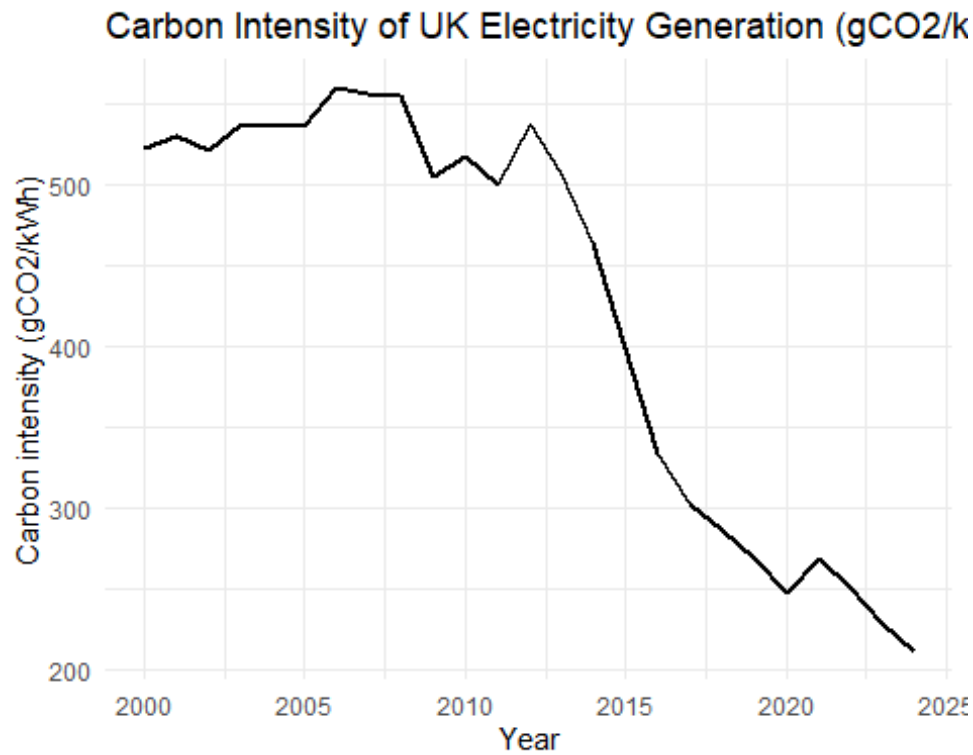


Figure 2 plots the carbon intensity of UK electricity generation from 2000–2024. Carbon intensity trends downward over the sample, indicating a sustained reduction in emissions per unit of electricity generated. This pattern is consistent with the UK’s shift away from coal and towards lower-carbon sources, and provides a visual benchmark for the regression relationship tested later.

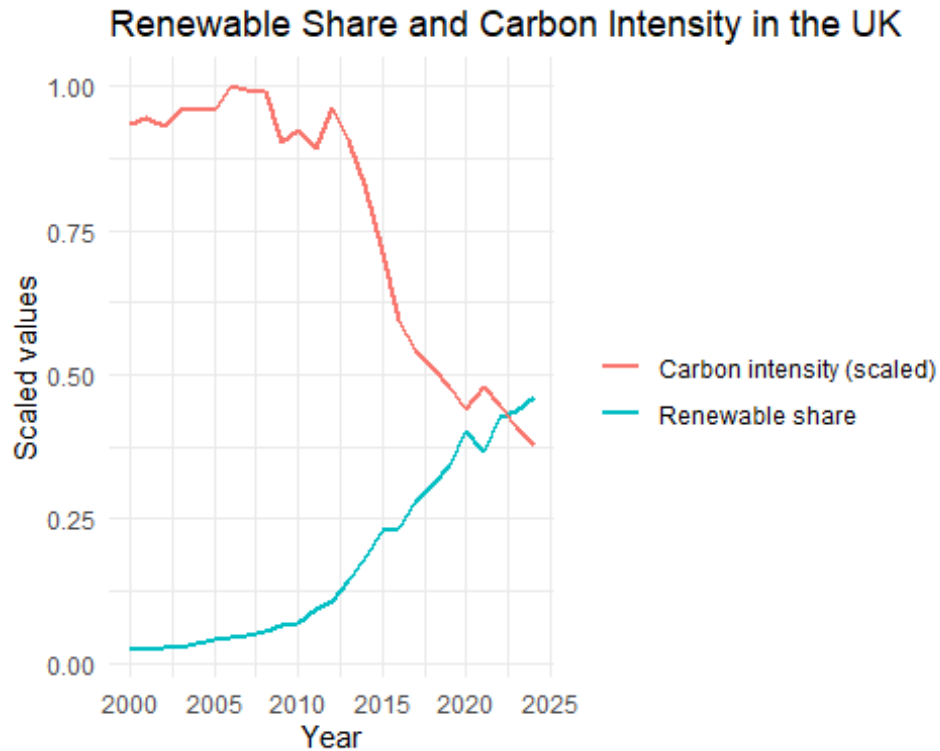


Figure 3 plots the UK's renewable share of electricity generation alongside carbon intensity over time. Since the two series are measured in different units, carbon intensity is linearly rescaled to lie between 0 and 1, allowing both variables to be visualised on the same axis without affecting their underlying time-series patterns.

The figure reveals a clear inverse relationship. From the early 2000s onward, the renewable share increases steadily, accelerating markedly after 2010. Over the same period, carbon intensity declines sharply, particularly during the mid-2010s, coinciding with rapid expansion of renewable generation and the phase-out of coal-fired electricity. While both series exhibit strong time trends, their opposing movements suggest a negative association between renewable penetration and the carbon intensity of electricity generation.

This visual evidence provides preliminary support for the hypothesis that increases in renewable electricity generation are associated with reductions in carbon intensity. However, as both variables trend over time, this descriptive analysis alone is insufficient to establish the strength or robustness of the relationship. The next section therefore formalises this association using regression analysis.

Relationship Analysis

Baseline: longrun association (levels)

I begin with a baseline levels regression that describes the long-run association between the UK's renewable electricity share (measured as a proportion) and the carbon intensity

of electricity generation (gCO₂/kWh). This specification captures how carbon intensity co-moves with the electricity generation mix over the sample period, but on its own it does not distinguish this relationship from broader underlying time trends.

```
levels_model <- lm(
  carbon_intensity ~ renewable_share,
  data = analysis_data_clean
)
```

Baseline levels regression: carbon intensity on renewable share

term	estimate	std.error	statistic	p.value
(Intercept)	572.701	8.783	65.209	0
renewable_share	-813.155	37.388	-21.749	0

Model fit statistics (baseline levels regression)

nobs	r.squared	adj.r.squared
25	0.954	0.952

In the baseline levels model, the coefficient on `renewable_share` is negative and statistically significant ($\beta = -813.155$, $p < 0.001$), indicating that higher renewable shares are associated with lower carbon intensity. Because `renewable_share` is measured as a proportion, a 1 percentage point increase (0.01) is associated with an average reduction of approximately 8.13 gCO₂/kWh (0.01×813.155). The model fit is high ($R^2 = 0.954$), but this should be interpreted cautiously in a trending time-series context, as part of the strong fit may reflect shared time trends rather than a stable structural relationship.

Robustness check 1: levels model with explicit time control

To reduce the risk that the baseline association is driven primarily by shared trends over time, I introduce an explicit time control using `year`. The purpose is not to interpret the year coefficient causally, but to absorb broad time varying influences such as policy changes, technological progress, and fuel-market conditions and to assess whether the renewables coefficient remains negative once common trends are accounted for.

```
levels_time_model <- lm(
  carbon_intensity ~ renewable_share + year,
  data = analysis_data_clean
)
```

Levels regression with time control (year)

term	estimate	std.error	statistic	p.value
(Intercept)	-11204.502	4807.652	-2.331	0.029
renewable_share	-1079.982	114.072	-9.468	0.000
year	5.877	2.399	2.450	0.023

Model fit statistics (levels + year)

nobs	r.squared	adj.r.squared
25	0.964	0.96

After controlling for year, the coefficient on renewable share remains negative and statistically significant ($\beta = -1079.982$, $p < 0.001$). Interpreted in percentage points, a 1pp increase in renewables is associated with an average reduction of approximately 10.80 gCO₂/kWh (0.01×1079.982). The change in magnitude relative to the baseline is consistent with both renewables and carbon intensity trending over time. Accordingly, year is treated as a trend control rather than given a structural interpretation; the key result is that the renewables carbon intensity relationship remains strongly negative after accounting for common time trends.

Robustness check 2: short-run dynamics (first differences)

Next, I estimate the relationship in first differences to focus on year to year changes rather than long run comovement in levels. Differencing reduces the influence of slow-moving trends and helps assess whether annual increases in renewables are associated with contemporaneous annual reductions in carbon intensity.

```
diff_data <- analysis_data_clean %>%
  arrange(year) %>%
  mutate(
    d_carbon = carbon_intensity - lag(carbon_intensity),
    d_renewable = renewable_share - lag(renewable_share)
  ) %>%
  filter(!is.na(d_carbon))

diff_model <- lm(d_carbon ~ d_renewable, data = diff_data)
```

First-differences regression: Δ carbon intensity on Δ renewable share

term	estimate	std.error	statistic	p.value
(Intercept)	-1.644	6.100	-0.270	0.790
d_renewable	-622.666	213.912	-2.911	0.008

Model fit statistics (first differences)

nobs	r.squared	adj.r.squared
24	0.278	0.245

In the differenced specification, changes in renewable share are negatively and statistically significantly associated with changes in carbon intensity ($\beta = -622.666$, $p = 0.008$). Because renewable_share is measured as a proportion, a 1pp increase (0.01) is associated with an average reduction of approximately 6.23 gCO₂/kWh (0.01×622.666). Explanatory power is lower ($R^2 = 0.278$), which is expected in first differences, as short-run annual movements are influenced by transitory factors such as weather conditions, demand variation, and fuel prices, in addition to longer-run structural change.

Robustness check 3: mechanism test (coal displacement)

Finally, I extend the levels specification by controlling for coal's share of total generation. This provides a mechanism consistent check: if renewables are associated with lower carbon intensity partly through the displacement of coal generation, then (i) coal_share should be strongly positively associated with carbon intensity, and (ii) the magnitude of the renewables coefficient should attenuate once coal share is included.

```
analysis_data_coal <- analysis_data_clean %>%
  left_join(
    electricity_uk %>% select(year, coal, total_generation),
    by = "year"
  ) %>%
  mutate(
    coal_share = coal / total_generation
  )

coal_model <- lm(
  carbon_intensity ~ renewable_share + coal_share,
  data = analysis_data_coal
)
```

Mechanism check: renewables and coal share (levels)

term	estimate	std.error	statistic	p.value
(Intercept)	375.648	18.925	19.849	0
renewable_share	-347.836	46.504	-7.480	0
coal_share	533.191	50.258	10.609	0

Model fit statistics (levels + coal share)

nobs	r.squared	adj.r.squared
25	0.992	0.992

Controlling for coal's share of electricity generation supports the coal-displacement mechanism. `coal_share` is strongly and positively associated with carbon intensity ($\beta = 533.191$, $p < 0.001$), meaning that a 10pp higher coal share is associated with roughly +53.3 gCO₂/kWh higher carbon intensity (0.10×533.191). The renewables coefficient remains negative and statistically significant ($\beta = -347.836$, $p < 0.001$), but its magnitude is smaller than in the simpler levels models; in percentage-point terms, a 1pp increase in renewables is associated with about -3.48 gCO₂/kWh (0.01×347.836) holding coal share constant. This pattern is consistent with the idea that part of the renewables-carbon relationship operates through contemporaneous fuel-mix changes, especially coal displacement.

Extension: heterogeneity via interaction (renewables × coal)

To test whether the relationship between renewable share and carbon intensity varies with how coal-intensive the electricity system is, I extend the coal mechanism model by including an interaction between `renewable_share` and `coal_share`, while also controlling for year. With an interaction, the marginal effect of renewable share on carbon intensity is given by the coefficient on `renewable_share` plus the interaction coefficient multiplied by `coal_share`. A negative interaction term therefore implies that the emissions benefit of renewables becomes larger in magnitude when coal's share is higher, which is consistent with the idea that renewables deliver greater reductions when they are more likely to displace coal generation.

```
interaction_model <- lm(
  carbon_intensity ~ renewable_share * coal_share + year,
  data = analysis_data_coal
)
```

Interaction model: renewables × coal share (with year control)

term	estimate	std.error	statistic	p.value
(Intercept)	-8564.942	2324.029	-3.685	0.001
renewable_share	-522.522	65.143	-8.021	0.000
coal_share	577.900	51.664	11.186	0.000
year	4.460	1.160	3.846	0.001
renewable_share:coal_share	-584.881	218.977	-2.671	0.015

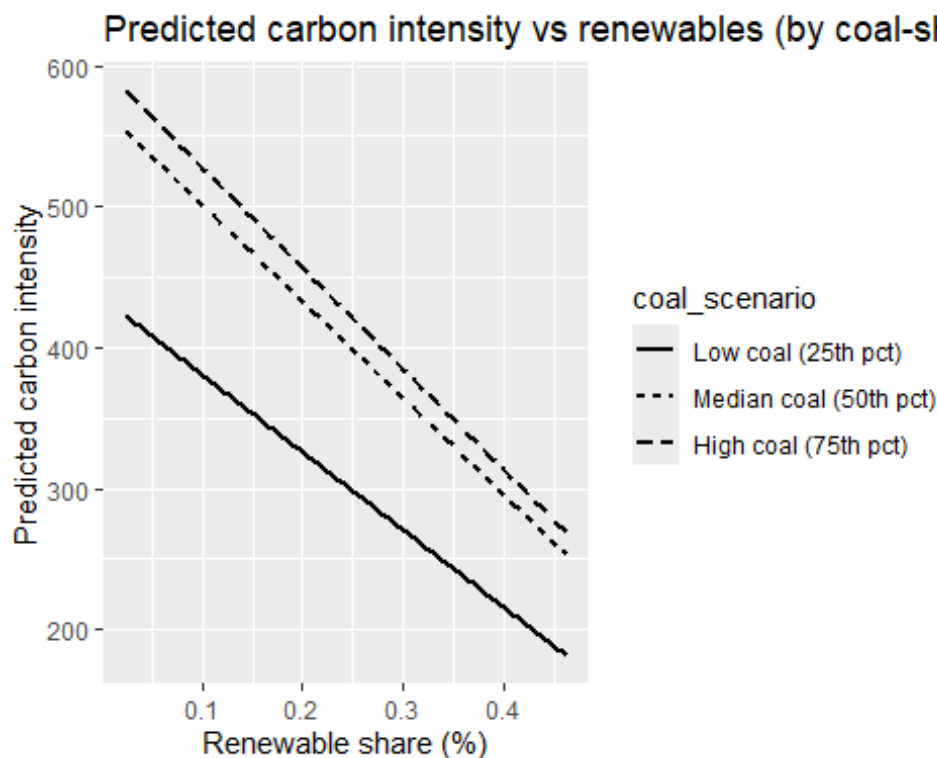
Model fit statistics (interaction model)

nobs	r.squared	adj.r.squared
25	0.996	0.995

The interaction term is negative and statistically significant ($\beta = -584.88$, $p = 0.015$), implying that the marginal association between renewables and carbon intensity becomes

more negative when coal's share is higher. Coal share remains strongly positively associated with carbon intensity ($\beta = 577.90$, $p < 0.001$). Consistent with this, Figure 4 shows steeper predicted declines in carbon intensity with renewables under higher coal-share scenarios, supporting the interpretation that renewable expansion delivers larger emissions benefits when the system is more coal-intensive (i.e., when coal displacement is more likely).

To make the interaction easier to interpret visually, I plot predicted carbon intensity against renewable share for three representative coal share scenarios (low / median / high coal share). This figure is descriptive: it translates the fitted linear model into implied relationships across plausible values.



When interpreting Figure 4, the key feature is the difference in slopes across coal-share scenarios. The steeper downward slope under higher coal shares indicates that expansions in renewable generation are associated with larger reductions in carbon intensity when the electricity system is more coal-intensive.

Extension: non-linear relationship (quadratic specification)

It is plausible that the relationship between renewable deployment and carbon intensity is non-linear. Early renewable expansion may displace particularly carbon intensive generation, while later expansion could yield diminishing marginal reductions as the grid becomes cleaner. To test this, I estimate a quadratic specification in renewable_share while retaining the year control to absorb common time trends.

```
nonlinear_model <- lm(
  carbon_intensity ~ renewable_share + I(renewable_share^2) + year,
  data = analysis_data_clean
)
```

In the quadratic model, the linear renewable term is negative and statistically significant ($\beta = -1620.913$, $p < 0.001$) while the squared term is positive and statistically significant ($\beta = 838.350$, $p = 0.041$), implying a downward-sloping relationship that flattens as renewable share rises (diminishing marginal reductions). The implied turning point is around $R^* \approx 0.97$ (97% renewables), which lies well beyond the observed renewable shares in this sample, so the fitted relationship remains decreasing over the data range. Figure 5 visualises this pattern: carbon intensity falls as renewables rise, but the curve becomes less steep at higher renewable penetration.

Quadratic model coefficients (non-linear renewables effect, with year control)

term	estimate	std.error	statistic	p.value
(Intercept)	-18572.899	5590.134	-3.322	0.003
renewable_share	-1620.913	270.228	-5.998	0.000
I(renewable_share^2)	838.350	385.579	2.174	0.041
year	9.565	2.793	3.425	0.003

To visualise the implied non-linear relationship, I plot the fitted quadratic curve over the observed data. This helps communicate the shape implied by the model.

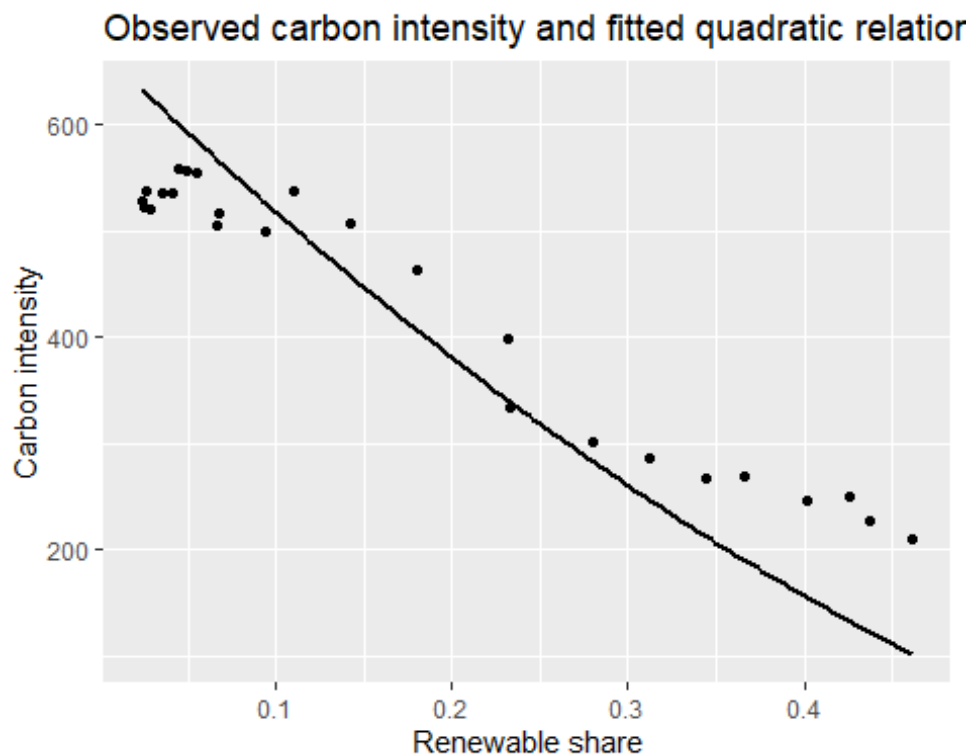


Figure 5 illustrates the fitted quadratic relationship between renewable share and carbon intensity. The downward sloping but flattening curve is consistent with diminishing marginal emissions reductions as renewable penetration rises, aligning with the positive squared term in the regression. While the linear specification captures the dominant trend, the quadratic model provides evidence that marginal effects weaken at higher renewable shares, serving as a robustness check on functional form rather than a substantive alternative.

Overall, allowing for non-linearity does not overturn the core finding: higher renewable shares remain strongly associated with lower carbon intensity. However, the quadratic fit (Figure 5) suggests that the marginal emissions benefit may weaken at higher renewable penetration, which is consistent with the system becoming cleaner over time. Together with the interaction evidence (Figure 4), this motivates checking how the estimated renewables coefficient changes as additional controls are introduced in the levels specifications below

Comparing coefficients across levels specifications

To make the model-nesting logic transparent, I compare the estimated renewable_share coefficient across three levels-based specifications: (i) the baseline levels model, (ii) the levels model with a time control (year), and (iii) the coal-augmented model that isolates the fuel-mix channel. This summary shows how the magnitude of the renewables coefficient changes as additional controls are introduced.

Coefficient comparison across levels specifications

model	term	estimate	std.error	statistic	p.value
Levels	renewable_share	-813.155	37.388	-21.749	0.000
Levels + year	renewable_share	-1079.982	114.072	-9.468	0.000
Levels + year	year	5.877	2.399	2.450	0.023
Levels + coal	renewable_share	-347.836	46.504	-7.480	0.000
Levels + coal	coal_share	533.191	50.258	10.609	0.000

The table shows that the renewables coefficient is negative in all three levels specifications, but its magnitude varies with controls. The baseline estimate (-813.155) corresponds to about -8.13 gCO₂/kWh per 1pp increase in renewables, while adding the time control yields a larger magnitude (-1079.982 ≈ -10.80 per 1pp), consistent with trending behaviour affecting the raw relationship. Once coal share is included, the renewables coefficient shrinks to -347.836 (≈ -3.48 per 1pp) and coal share enters strongly positively (533.191), consistent with coal displacement accounting for an important part of the renewables-carbon intensity relationship.

Overall, the results across specifications provide consistent evidence that higher renewable electricity shares are associated with lower UK electricity carbon intensity. In the baseline levels model, the estimated long-run association is large and highly

statistically significant ($\beta \hat{=} -813.2$, $p < 0.001$; $R^2 = 0.954$), but—because both series trend over time—this fit could partly reflect common time trends. Adding an explicit time control confirms the negative renewables relationship remains even after accounting for secular decarbonisation: the renewable coefficient stays negative and significant ($\beta \hat{=} -1080.0$, $p < 0.001$), alongside a positive year coefficient ($\beta \hat{=} 5.88$, $p = 0.023$; $R^2 = 0.964$). Focusing on short-run dynamics via first differences also supports the relationship: year-to-year changes in renewables are significantly associated with year-to-year changes in carbon intensity ($\beta \hat{=} -622.7$, $p = 0.008$; $R^2 = 0.278$), implying that a 1 percentage point increase in renewable share is associated with an average reduction of around 6.2 gCO₂/kWh.

Mechanism based models strengthen the interpretation that renewables reduce carbon intensity primarily through coal displacement. When coal share is included, coal share is strongly positively associated with carbon intensity ($\beta \hat{=} 533.2$, $p < 0.001$), while the renewable coefficient remains negative but shrinks substantially ($\beta \hat{=} -347.8$, $p < 0.001$; $R^2 = 0.992$), consistent with renewables operating partly by reducing coal generation. The interaction model reinforces this channel: the renewables \times coal share term is negative and statistically significant ($\beta \hat{=} -584.9$, $p = 0.015$), indicating that the marginal emissions benefit of renewables is larger when coal's role in the mix is higher, which is also visualised in Figure 4 through the steeper predicted decline under high-coal scenarios ($R^2 = 0.996$).

Allowing for non-linearity does not overturn the headline result, but suggests diminishing marginal benefits as renewable penetration increases. In the quadratic specification, the linear renewables term remains negative and significant ($\beta \hat{=} -1620.9$, $p < 0.001$) while the squared term is positive and significant ($\beta \hat{=} 838.4$, $p = 0.041$), implying a downward-sloping but flattening relationship, as shown in Figure 5 ($R^2 = 0.970$). Taken together, the consistent negative sign and statistical significance of renewables across levels, trend-controlled, differenced, mechanism, interaction, and quadratic models indicates a robust association between renewable expansion and lower carbon intensity, with the strongest evidence pointing to coal displacement as the primary transmission channel.

Predictions and Residuals

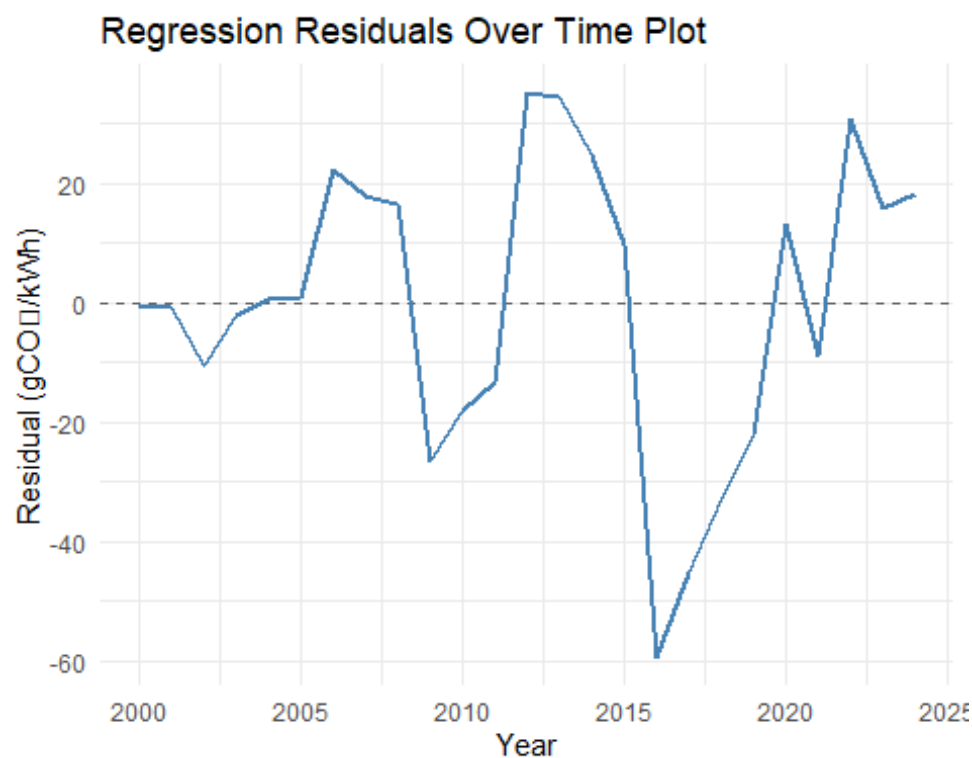
After estimating the relationship between renewable generation and carbon intensity, I next evaluate how well the preferred specification fits the data. This section uses predicted values and residuals to check whether the model captures the main movements in carbon intensity over time, and to identify any years where actual outcomes deviate substantially from what the model would predict. Examining residual patterns is also useful for diagnosing potential misspecification (e.g., non-linearity, changing variance, or omitted shocks) and for motivating any further robustness checks.

Observed and predicted carbon intensity and residuals (first 6 years, levels + year model)

year	renewable_share	carbon_intensity	predicted_carbon	residual
2000	0.025	521.986	522.681	-0.695
2001	0.024	528.991	529.842	-0.851

year	renewable_share	carbon_intensity	predicted_carbon	residual
2002	0.028	520.891	531.438	-10.547
2003	0.027	536.777	538.833	-2.056
2004	0.035	536.022	535.503	0.519
2005	0.042	535.258	534.463	0.795

The table confirms how residuals are constructed: $\text{residual} = \text{actual carbon intensity} - \text{predicted carbon intensity}$. Positive residuals indicate under-prediction (actual > predicted), while negative residuals indicate over prediction. The remaining figures use these residuals to assess whether the model errors are mostly noise or display systematic structure.



The residual series, figure 6, (from the levels + year model) shows that even after controlling for renewable share and an explicit time trend there are periods where the model persistently over or under-predicts carbon intensity. This is consistent with omitted time-varying influences such as shifts in relative fuel prices, policy changes affecting dispatch, outages, and weather-driven demand. The residual patterns therefore reinforce that the regression is best interpreted as a descriptive association, and that year specific shocks can still matter even after accounting for renewables and a linear trend.

In the early 2000s, residuals are generally close to zero (with occasional modest negatives), suggesting the fitted relationship captures average outcomes reasonably well in that period. Larger deviations appear later in the sample. The largest under-prediction

occurs in 2012 (residual $\approx +35.2$ gCO₂/kWh), while the strongest over-prediction occurs in 2016 (residual ≈ -59.5 gCO₂/kWh), implying carbon intensity fell more sharply than the model predicts in 2016 given renewable share and the linear trend alone. These episodes are consistent with additional structural or shock-driven factors affecting the generation mix and emissions intensity beyond the included regressors

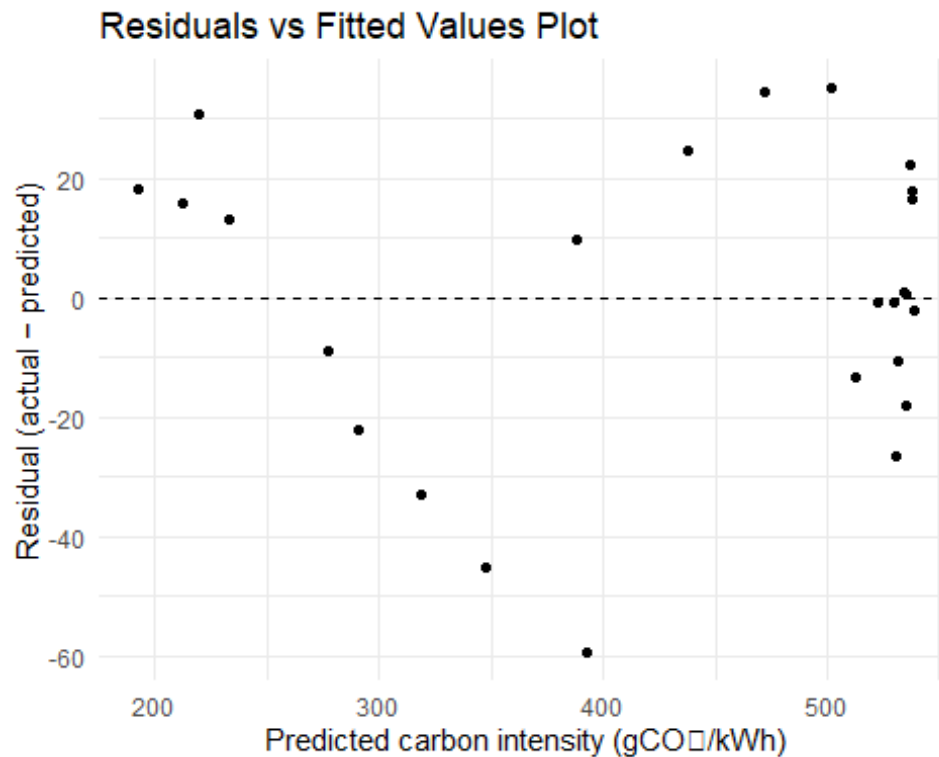


Figure 7 plots residuals against fitted values to check for functional form problems and changing variance. A clear curve would suggest non-linearity not captured by the model, while a funnel shape would indicate heteroskedasticity (error variance changing with the fitted level of carbon intensity). Any remaining structure should therefore be interpreted as evidence of potential misspecification and/or omitted time-varying influences (e.g., fuel-market conditions, outages, weather-driven demand, imports, or policy-related changes) beyond renewable share and a linear trend.

Additional diagnostic detail is provided in Appendix Figure 8, which visualises residuals as vertical segments between observed and fitted carbon intensity.

Discussion and Limitations

Discussion

The descriptive evidence (Figures 1–3) shows a pronounced rise in renewable electricity share alongside a substantial decline in carbon intensity over the sample. The baseline levels regression formalises this long-run relationship: renewable share is strongly negatively associated with carbon intensity ($\beta = -813.155$, $p < 0.001$), implying that a 1

percentage point (pp) increase in renewable share (0.01) corresponds to an average reduction of about 8.13 gCO₂/kWh. The model fit is high ($R^2 = 0.954$), which is consistent with a strong long-run co-movement between the two series.

Introducing an explicit time control is an important robustness step given shared trends. In the levels + year model, the renewable coefficient remains strongly negative ($\beta = -1079.982$, $p < 0.001$), implying about 10.80 gCO₂/kWh lower carbon intensity for a 1pp increase in renewable share. The year coefficient is positive ($\beta \approx 5.88$, $p = 0.023$), reflecting collinearity between year and renewable share rather than a meaningful structural effect. Accordingly, year is treated purely as a trend control; the substantive result is that the renewables–carbon intensity relationship remains strongly negative after accounting for common time trends

To focus on short run comovement rather than long-run levels, the first-differences specification examines annual changes. The differenced model still shows a negative and statistically significant relationship ($\beta = -622.666$, $p = 0.008$), implying that a 1pp increase in renewables is associated with an average reduction of around 6.23 gCO₂/kWh in carbon intensity. As expected for year-to-year variation, explanatory power is lower ($R^2 = 0.278$), but the persistence of a negative association in changes reduces concern that the levels result is purely an artefact of common trending behaviour.

Mechanism analysis incorporating coal share is consistent with fuel-mix displacement playing an important role. Coal share is strongly positively associated with carbon intensity: a 10pp increase in coal share corresponds to roughly +53.3 gCO₂/kWh higher carbon intensity (0.10×533.191). Once coal share is included, the renewables coefficient shrinks markedly ($\beta = -347.836$), implying about –3.48 gCO₂/kWh for a 1pp increase (0.01×347.836), which is consistent with renewables operating partly through reduced coal generation (and potentially through broader system effects beyond coal displacement alone). The interaction extension reinforces this channel: coal share remains strongly positive ($\beta = 577.90$, $p < 0.001$) and Figure 4 shows a steeper predicted decline in carbon intensity with renewables when the system is more coal-intensive; overall fit is very high ($R^2 = 0.996$). The quadratic extension suggests a concave/flattening relationship at higher renewable shares (Figure 5; $R^2 = 0.970$), which is best interpreted as a functional-form diagnostic rather than a definitive structural claim given the short annual sample.

Finally, the residual diagnostics show that the model captures broad movements but leaves notable year-specific deviations. In Figure 6, the largest under-prediction occurs in 2012 (residual $\approx +35.2$ gCO₂/kWh) and the largest over-prediction occurs in 2016 (residual ≈ -59.5 gCO₂/kWh), indicating that some years deviate materially from what the model predicts based only on renewable share and a linear trend. Figure 7 and the residual “segment” diagnostic (Appendix Figure 8) reinforce that while residuals are centred around zero overall, a small number of years contribute disproportionately to model error. These patterns support interpreting the regressions as descriptive associations that capture average relationships, rather than a complete model of all drivers of emissions intensity.

Limitations

First, the analysis is descriptive and does not establish causality. Renewable share may be endogenous to conditions that also affect carbon intensity (e.g., weather, demand, fuel prices, outages, and policy), potentially biasing the estimated association. Without an identification strategy or richer covariates, the coefficients should therefore be interpreted as descriptive associations rather than causal effects

Second, the data are annual with a limited sample size. Annual aggregation can mask short-run operational dynamics and policy timing, limiting the ability to distinguish persistent structural change from transitory shocks. This also makes extensions (e.g., interactions and non-linearity) harder to estimate precisely and increases sensitivity to modelling choices.

Third, the time control is parsimonious (linear). If decarbonisation involves non-linear trends or structural breaks, a linear trend may be insufficient. Moreover, with annual time-series data, serial correlation is plausible, implying standard errors may be understated. The residual patterns and outlying years therefore suggest omitted time-varying drivers remain important; future work could incorporate additional controls and time-series-robust inference

Conclusion

Across specifications, renewable electricity share is consistently negatively associated with the carbon intensity of UK electricity generation. The baseline levels estimate implies about $-8.13 \text{ gCO}_2/\text{kWh}$ per 1pp increase in renewable share ($R^2 = 0.954$), the association remains strongly negative with a time trend (about $-10.80 \text{ gCO}_2/\text{kWh}$ per 1pp; $R^2 = 0.964$), and it persists in year-to-year changes (about $-6.23 \text{ gCO}_2/\text{kWh}$ per 1pp; $p = 0.008$). Mechanism results are consistent with coal displacement being an important channel (e.g., $+53.3 \text{ gCO}_2/\text{kWh}$ for a 10pp increase in coal share), while residual diagnostics highlight a small number of years—such as 2012 (+35.2) and 2016 (−59.5)—where additional unmodelled factors materially affect carbon intensity. Overall, the evidence supports renewable expansion as a strong correlate of UK electricity decarbonisation over the period studied, but the estimates should be treated as descriptive rather than causal.

References

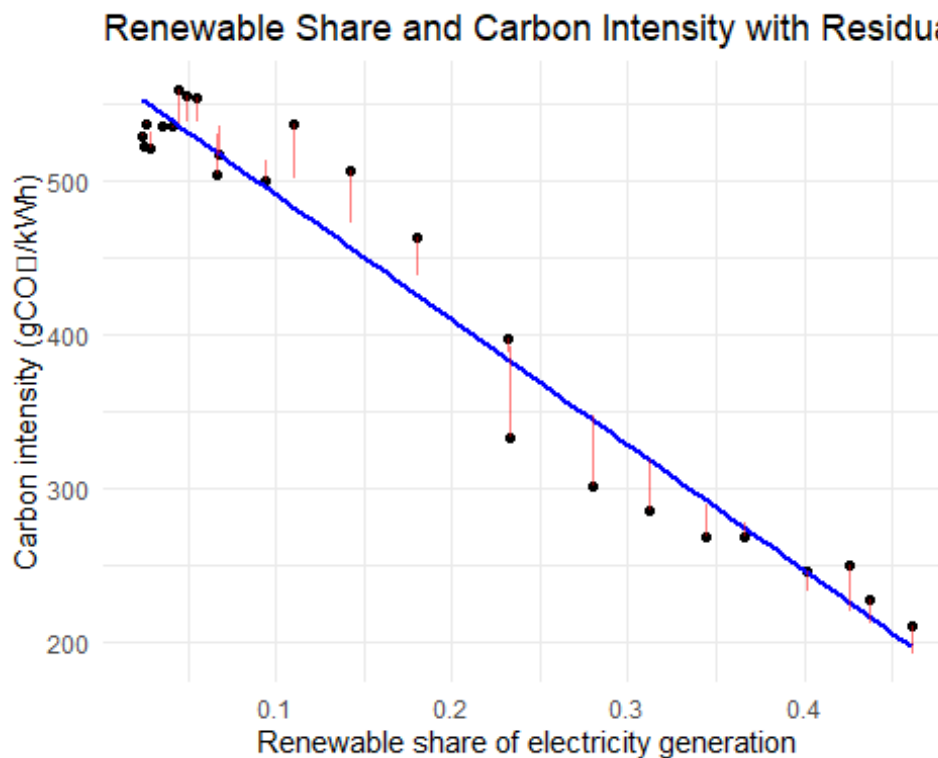
Ritchie, H., Roser, M. and Rosado, P. (2024) Electricity mix in the United Kingdom. Our World in Data. Available at: <https://ourworldindata.org/grapher/electricity-mix-uk> (Accessed: 11 January 2026).

Ritchie, H., Roser, M. and Rosado, P. (2024) Carbon intensity of electricity generation. Our World in Data. Available at: <https://ourworldindata.org/grapher/carbon-intensity-electricity> (Accessed: 11 January 2026).

Appendix

Appendix A: Reproducibility

The main document selectively displays code to conserve the word count. The full workflow remains reproducible from this R Markdown file together with the input datasets. Package versions are reported below.



Appendix Figure 8 visualises residuals as vertical segments between observed carbon intensity and the fitted regression line. Larger segments highlight years where outcomes deviate most strongly from the fitted relationship at a given renewable share. This figure is included as an intuitive visual complement to the formal residual diagnostics in Figures 6-7

Appendix B: Code listings (not executed)

The following chunks reproduce key data-processing and diagnostic steps. They are shown for transparency but are not re-run in the appendix.

```
carbon_raw <- read_csv("carbon-intensity-electricity.csv")
electricity_raw <- read_csv("electricity-mix-uk.csv")

carbon_uk <- carbon_raw %>%
  filter(Entity == "United Kingdom") %>%
  select(
    year = Year,
    carbon_intensity = `Carbon intensity of electricity - gCO2/kWh`
```

```

)

electricity_uk <- electricity_raw %>%
  filter(Entity == "United Kingdom") %>%
  select(
    year = Year,
    wind = `Electricity generation from wind`,
    solar = `Electricity generation from solar`,
    hydro = `Electricity generation from hydropower`,
    other_renewables = `Electricity generation from other renewables`,
    gas = `Electricity generation from gas`,
    coal = `Electricity generation from coal`,
    oil = `Electricity generation from oil`,
    nuclear = `Electricity generation from nuclear`,
    net_imports = `Net electricity imports`
  ) %>%
  mutate(
    renewable_generation = wind + solar + hydro + other_renewables,
    total_generation = renewable_generation + gas + coal + oil + nuclear +
net_imports,
    renewable_share = renewable_generation / total_generation
  )

analysis_data_clean <- left_join(electricity_uk, carbon_uk, by = "year") %>%
  filter(!is.na(carbon_intensity)) %>%
  select(year, renewable_share, carbon_intensity)

diff_data <- analysis_data_clean %>%
  arrange(year) %>%
  mutate(
    d_carbon = carbon_intensity - lag(carbon_intensity),
    d_renewable = renewable_share - lag(renewable_share)
  ) %>%
  filter(!is.na(d_carbon))

diff_model <- lm(d_carbon ~ d_renewable, data = diff_data)

analysis_data_clean <- analysis_data_clean %>%
  mutate(
    predicted_carbon = predict(levels_time_model),
    residual = carbon_intensity - predicted_carbon
  )

resid_summary <- analysis_data_clean %>%
  summarise(
    mean_resid = mean(residual, na.rm = TRUE),
    sd_resid = sd(residual, na.rm = TRUE),
    p10 = quantile(residual, 0.10, na.rm = TRUE),
    p50 = quantile(residual, 0.50, na.rm = TRUE),
    p90 = quantile(residual, 0.90, na.rm = TRUE)
  )

```

```
)

resid_outliers <- analysis_data_clean %>%
  mutate(abs_resid = abs(residual)) %>%
  arrange(desc(abs_resid)) %>%
  slice_head(n = 5) %>%
  select(year, renewable_share, carbon_intensity, predicted_carbon, residual)
```

Appendix C: Session information

```
## R version 4.5.2 (2025-10-31 ucrt)
## Platform: x86_64-w64-mingw32/x64
## Running under: Windows 11 x64 (build 26200)
##
## Matrix products: default
## LAPACK version 3.12.1
##
## locale:
## [1] LC_COLLATE=English_United Kingdom.utf8
## [2] LC_CTYPE=English_United Kingdom.utf8
## [3] LC_MONETARY=English_United Kingdom.utf8
## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United Kingdom.utf8
##
## time zone: Europe/London
## tzcode source: internal
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods   base
##
## other attached packages:
## [1] knitr_1.50      broom_1.0.10  lubridate_1.9.4 forcats_1.0.1
## [5] stringr_1.6.0  dplyr_1.1.4   purrr_1.2.0   readr_2.1.6
## [9] tidyr_1.3.1     tibble_3.3.0  ggplot2_4.0.1  tidyverse_2.0.0
##
## loaded via a namespace (and not attached):
## [1] generics_0.1.4  stringi_1.8.7  lattice_0.22-7  hms_1.1.4
## [5] digest_0.6.38   magrittr_2.0.4  evaluate_1.0.5  grid_4.5.2
## [9] timechange_0.3.0 RColorBrewer_1.1-3 fastmap_1.2.0   Matrix_1.7-4
## [13] backports_1.5.0 mgcv_1.9-3     scales_1.4.0    cli_3.6.5
## [17] rlang_1.1.6     crayon_1.5.3   bit64_4.6.0-1
## [21] splines_4.5.2
## [25] withr_3.0.2     yaml_2.3.10    tools_4.5.2
## [29] parallel_4.5.2
## [33] tzdb_0.5.0      vctrs_0.6.5    R6_2.6.1
## [37] lifecycle_1.0.4
## [41] bit_4.6.0       vroom_1.6.6    pkgconfig_2.0.3
## [45] pillar_1.11.1
## [49] gtable_0.3.6    glue_1.8.0     xfun_0.54
## [53] tidyselect_1.2.1
```

```
## [37] rstudioapi_0.17.1  farver_2.1.2      htmltools_0.5.8.1  nlme_3.1-168
## [41] rmarkdown_2.30     labeling_0.4.3    compiler_4.5.2     S7_0.2.1
```