International Journal of Scientific Research and Modern Education (IJSRME)
International Peer Reviewed - Refereed Research Journal, Website: www.rdmodernresearch.com
Impact Factor: 7.137, ISSN (Online): 2455 - 5630, Volume 9, Issue 2, July - December, 2024

ASSESSING ELECTRICITY GENERATION MIX ASSOCIATIONS WITH CARBON INTENSITY TO SUPPORT CARBON AWARE WORKLOAD SCHEDULING DECISIONS

Divya Vishwanath* & Paul Johnson Asamoah**
* Janardhan Rai Nagar Rajasthan Vidyapeeth, Udaipur, Rajasthan, India
** Power Life University College, Accra, Ghana, West Africa

Cite This Article: Divya Vishwanath & Paul Johnson Asamoah, "Assessing Electricity Generation Mix Associations with Carbon Intensity to Support Carbon Aware Workload Scheduling Decisions", International Journal of Scientific Research and Modern Education, Volume 9, Issue 2, July - December, Page Number 63-72, 2024.
Copy Right: © R&D Modern Research Publication, 2024 (All Rights Reserved). This is an Open Access Article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract:

This analysis examines whether electricity generation composition provides consistent signals of decarbonization across national and territorial systems. A balanced panel of 2,828 annual observations from 202 countries and territories during 2010-2023 supports descriptive distributional analysis, generation share comparisons and paired endpoint assessments. Equal location average lifecycle carbon intensity declines by 9.14 percent across the interval. Renewable expansion accompanies declining intensity in 125 of 136 locations, while coal reductions coincide with improvement in 47 of 51 locations. Gas expansion and contraction both accompany intensity reductions, revealing heterogeneous composition patterns. These relationships reflect joint portfolio changes rather than independently identified substitution effects. Relaxing complete period coverage changes pooled mean intensity by 1.72 grams of carbon dioxide equivalent per kilowatt hour. The contribution distinguishes differences across electricity systems from changes within locations, refining a compositional interpretation of electricity transitions. Because reported intensity partly incorporates generation composition, the evidence characterizes accounting relationships rather than causal impacts. The findings support evaluating transition targets through observed emissions performance and the surrounding electricity portfolio rather than isolated technology shares.

Key Words: Carbon Intensity; Comparative Analysis; Electricity Generation; Energy Transition; Panel Data

1. Introduction:

Electricity decarbonization requires more than identifying which generation sources expand or contract. The carbon implications of a changing generation share depend on the wider electricity mix and the emissions associated with producing electricity. A declining coal share, an increasing gas share and an expanding renewable share therefore convey different information about system performance. Treating these changes as interchangeable indicators can obscure whether electricity supply is becoming less emissions intensive. This distinction matters for policymakers assessing electricity transitions and organizations comparing the carbon characteristics of electricity supply. Industrial ecology provides a basis for examining these relationships through the material and energy processes connecting electricity technologies to greenhouse gas emissions. Lifecycle carbon intensity captures emissions per unit of electricity generated, including burdens beyond direct plant operation. Hertwich et al., 2015, demonstrate that changing electricity supply technologies affects environmental burdens throughout interconnected production systems. Pehl et al., 2017, further show that lifecycle emissions differ across technologies and depend on assumptions concerning production, operation and upstream processes. These findings establish why generation composition matters, while also showing that technology shares cannot be interpreted independently of their accounting boundaries.

The relationship is inherently compositional. A larger share for one source requires a smaller combined share for other sources. Coal expansion would generally raise intensity when it replaces generation with lower lifecycle emissions, but the observed relationship also reflects the technologies operating alongside coal. Whitaker et al., 2012, identify variation in coal electricity lifecycle emissions across generation technologies and assessment assumptions. Gas introduces a competing substitution logic because its expansion can accompany either lower or higher system intensity depending on the generation displaced. O'Donoughue et al., 2014, distinguish lifecycle performance across gas conversion technologies and fuel supply conditions. Consequently, a source share measures participation in electricity production rather than a uniform emissions effect. Renewable generation requires a similarly explicit interpretation. Wind and solar electricity avoid fuel combustion during generation, but equipment production, construction and other lifecycle processes still produce emissions. Dolan and Heath, 2012, and Hsu et al., 2012, show that consistent technical assumptions improve comparisons of wind and photovoltaic lifecycle estimates. At the electricity system level, an aggregate renewable share combines technologies with different characteristics. Its relationship with carbon intensity therefore represents the combined composition of renewable generation and the sources displaced, rather than an identical emissions contribution from every renewable technology.

The closest research addresses complementary analytical questions. Technology harmonization studies compare lifecycle emissions under more consistent assumptions, while integrated assessments examine electricity supply scenarios. Other work evaluates emissions associated with changes in electricity demand or traces emissions through interconnected grids. Hawkes, 2010, distinguishes marginal emissions relevant to electricity interventions from average grid emissions. De Chalendar et al., 2019, track emissions through electricity production, exchanges and consumption, demonstrating the importance of geographical and temporal accounting boundaries. These approaches establish that technology performance, system averages and intervention specific emissions are distinct quantities. They do not, by themselves, establish how consistently pooled generation share differences correspond to changes within the same broad set of national and territorial electricity systems. This analysis addresses that comparative question by examining coal, gas and renewable generation shares alongside annual lifecycle electricity carbon intensity. The purpose is to determine whether the patterns observed across electricity systems are also visible when locations are compared with their own earlier composition. This distinction is consequential because an average difference between systems need not describe the trajectory of an individual system. The contribution is therefore bounded: the analysis places pooled

composition comparisons and paired location changes within one consistent descriptive framework, rather than estimating an independent causal effect of each generation source.

This study aims to characterize how electricity generation composition corresponds to lifecycle carbon intensity across countries and territories during 2010-2023. The central research question asks how these relationships differ across generation sources and between pooled observations and changes within locations. Three objectives follow directly from the measured framework. The analysis examines carbon intensity across coal generation shares and coal share changes, evaluates the corresponding patterns for gas generation, and assesses whether renewable expansion accompanies declining intensity consistently across locations. The empirical setting comprises 2,828 annual observations from 202 countries and territories, with complete coverage for each location over 14 years. The four numerical measures come from the consolidated Our World in Data energy dataset. The analytical approach combines distributional summaries, equal location annual averages, exploratory generation share bands and paired comparisons between 2010 and 2023. A sensitivity calculation relaxes the complete period requirement while retaining complete annual records. Using the same locations throughout the main comparisons prevents changes in geographical sample composition from determining the reported annual averages.

The findings distinguish the three generation dimensions. Higher renewable share bands exhibit progressively lower average intensity, and intensity declines in 125 of the 136 locations recording renewable expansion. Coal reductions commonly accompany declining intensity, whereas coal expansion produces mixed outcomes. Gas increases and decreases both accompany intensity reductions across substantial groups of locations. These results show why a common direction of change in a generation share does not imply a common system trajectory. They also qualify the interpretation of the observed correspondence because the source intensity measure incorporates generation composition within its emissions accounting framework. The practical contribution concerns comparative assessment of electricity supply. Generation shares and measured intensity provide complementary information, and neither should be interpreted without the surrounding composition. For electricity dependent activities, including computing, the results characterize annual supply conditions rather than operational emissions savings. Their international relevance rests on applying consistent definitions and comparisons across diverse electricity systems. The resulting evidence supports a differentiated account of electricity transitions while preserving the distinction between descriptive composition patterns and the emissions consequences of particular interventions.

2. Theory, Conceptual Framework and Research Questions Development:

This analysis draws on industrial ecology and lifecycle assessment to explain how electricity generation composition relates to the greenhouse gas intensity of electricity supply. Electricity technologies operate within interconnected fuel, material and production systems. Their environmental burdens therefore extend beyond emissions released during generation. The relevant mechanism combines the amount of electricity supplied by each technology with the emissions associated with its production chain. Hertwich et al., 2015, demonstrate that changing electricity supply technologies alters environmental burdens across these interconnected systems. This establishes a theoretical basis for examining generation composition alongside lifecycle carbon intensity. The framework distinguishes electricity generation shares from electricity carbon intensity. Coal, gas and renewable generation shares measure the percentage of annual electricity generated from each source. Lifecycle carbon intensity measures greenhouse gas emissions per kilowatt hour of electricity generated, expressed in grams of carbon dioxide equivalent. These are observed numerical variables rather than indicators of a latent construct. Their relationship is compositional: an increase in one generation share requires a decrease in at least one other share. The expected direction therefore depends on which electricity source loses share. This analysis examines these relationships descriptively across 202 countries and territories during 2010-2023, using 2,828 annual observations.

Coal generation provides the first dimension of the framework. Its principal emissions mechanism is fuel combustion, supplemented by emissions associated with extraction, processing and transport. A larger coal share would generally correspond to greater electricity carbon intensity when coal replaces sources with lower lifecycle emissions. However, generation efficiency and production conditions affect the intensity associated with a given coal share. Whitaker et al., 2012, identify substantial variation across coal technologies and show that consistent system boundaries and technical assumptions improve comparability. The theoretical expectation is therefore conditional rather than a universal relationship between coal share and national electricity intensity.

The available measures permit comparisons of coal generation shares and carbon intensity across locations and over time. They do not identify the specific technology displaced by each change in coal share. We therefore frame the first inquiry descriptively. RQ1: How does annual electricity carbon intensity vary across countries and territories with different coal generation shares, and how do these measures change together during 2010-2023?

Gas generation introduces a competing substitution mechanism. Replacing coal with gas can reduce the emissions associated with electricity generation, whereas replacing sources with lower lifecycle emissions can increase them. Gas intensity also depends on conversion efficiency and emissions along the fuel supply chain. O'Donoughue et al., 2014, distinguish the lifecycle performance of combustion turbine and combined cycle systems and identify methane leakage as an influential source of uncertainty. These findings support conditional expectations about gas substitution, but they do not establish a single directional relationship applicable to every electricity system.

The gas generation share combines technologies and supply conditions that the dataset does not distinguish. Its descriptive relationship with carbon intensity can therefore differ across electricity mixes. A higher gas share may accompany declining intensity in one location and increasing intensity in another without contradicting the underlying substitution mechanism. RQ2: How does annual electricity carbon intensity vary with gas generation share across countries and territories and within locations during 2010-2023?

Renewable generation provides the third dimension. Wind and solar generation avoid fuel combustion during electricity production, but their lifecycle emissions include equipment production, construction and maintenance. Dolan and Heath, 2012, show that wind electricity lifecycle estimates depend on technical performance and assessment assumptions. Hsu et al., 2012,

similarly demonstrate that photovoltaic estimates require consistent assumptions for meaningful comparison. These findings explain why renewable generation can have low lifecycle emissions without being emissions free.

The renewable share nevertheless combines technologies with different emissions characteristics. Pehl et al., 2017, identify differences between wind, solar, hydropower and bioenergy, including uncertainty surrounding hydropower and bioenergy emissions. Consequently, a larger renewable share is expected to accompany lower carbon intensity principally when it replaces generation with higher lifecycle emissions. An aggregate renewable share cannot identify the contribution of each renewable technology or establish a uniform substitution effect.

This analysis therefore compares renewable generation shares with carbon intensity while retaining variation across locations and years. RQ3: How does annual electricity carbon intensity vary with renewable generation share, and how consistently do increases in that share accompany declining intensity during 2010-2023?

The conceptual framework treats coal, gas and renewable generation shares as three measured dimensions of electricity generation composition and lifecycle carbon intensity as the measured outcome. The relationships represent descriptive associations. Because the source intensity measure incorporates generation composition, their correspondence partly reflects the emissions accounting structure. We therefore examine distributions, annual trajectories and joint changes rather than interpret correspondence as independent confirmation of a causal mechanism. The research questions permit positive, negative, weak or inconsistent patterns without imposing a predetermined result.

3. Data:

This section documents data provenance, sample construction, variable measurement and the descriptive analytical framework. It establishes a consistent basis for comparing electricity generation composition and lifecycle carbon intensity across locations and years.

3.1 Data Sources, Population, and Sample Construction:

The analysis uses the legacy Energy dataset compiled by Our World in Data. The electricity series draw on Ember's Yearly Electricity Data and the Energy Institute's Statistical Review of World Energy, as identified in the accompanying source codebook. The consolidated dataset provides the observations; the codebook supplies definitions, units and upstream attribution. Maintaining these distinctions supports the traceability of energy research emphasized by Pfenninger et al., 2017, and Wiese et al., 2019. Neither publication supplies observations to the analytical sample. The extraction date is recorded in the accompanying workbook, separately from the observation period. The retained source snapshot is identified by its SHA256 checksum rather than a repository commit. Distinguishing source documentation, processing decisions and analytical outputs follows the transparency principles discussed by Pfenninger et al., 2018, and Morrison, 2018. Our World in Data identifies its own contributions as available under a Creative Commons attribution licence while retaining the applicable terms of upstream providers. Public accessibility therefore does not establish unrestricted rights over every underlying contribution. Attribution and redistribution remain subject to the relevant source terms, consistent with the distinction between accessible and legally reusable electricity data examined by Hirth, 2020.

The target population comprises national and territorial electricity generation systems during 2010-2023. The accessible sampling frame contains 220 locations with three letter location identifiers and at least one recorded year within that interval. Regional aggregates are excluded. The empirical unit is the location and calendar year, and the sectoral boundary is electricity generation. The resulting data form an annual panel. Generation shares describe the composition of electricity produced during each year, while carbon intensity expresses emissions per unit of annual generation. They are neither electricity capacity stocks nor values measured at year end. Selection required all four study variables to be available in every year from 2010 through 2023. This rule was implemented during dataset construction and retained during the descriptive audit. It is a coverage-based selection rule rather than probability sampling. No preregistration record accompanies the extraction. Table 1 distinguishes exclusions of entire locations from deletion of incomplete annual records.

Table 1: Sample construction

Stage	Locations	Observations	Reason or Rule
Accessible frame	220	3,078	Three letter location identifiers and recorded years within 2010-2023
Exclusion of locations without any complete annual record	15 removed	210 removed	At least one study variable missing in every recorded year
Removal of incomplete records among remaining locations	0 removed	2 removed	One incomplete record each for Lesotho and Ukraine
Complete annual records before balancing	205	2,866	All four study variables observed
Exclusion for incomplete period coverage	3 removed	38 removed	Lesotho and Ukraine each contribute 13 complete years; South Sudan contributes 12
Final descriptive sample	202	2,828	All four variables observed in each of 14 years

The final panel is balanced, with 202 observations in every year and 14 observations per location. South Sudan has no source rows for the first two years of the interval; these absent records are distinct from blank values in existing rows. The descriptive calculations use the same retained sample throughout.

3.2 Data Integration, Cleaning, and Quality Assurance:

All four measures were extracted from one consolidated source table. No external dataset was merged into the analytical sample. The location identifier and year jointly define the record key, and source location names were preserved. The audit found no duplicated keys in the accessible frame. Every retained value matched the source extraction, including recorded zeros. The workbook stores one observation per row and one variable per column, consistent with the data organization principles described by Broman and Woo, 2018, and Wickham, 2014. Temporal alignment uses the annual periods already supplied by the source. No

frequency conversion, interpolation or averaging of monthly observations was performed. The analysis uses revised historical observations from a single source snapshot rather than information vintages available at each historical date. Consequently, the data describe historical electricity systems as represented in that snapshot.

Across the 3,078 recorded observations in the accessible frame, coal share has 157 missing values, gas share has 212, renewable share has 114 and carbon intensity has 114. These counts overlap: 212 records lack at least one study variable. Incomplete records number 15 annually during 2010-2022 and 17 in 2023. The coverage worksheet identifies missingness by location. Source blanks were treated as unavailable observations because the extraction does not establish whether each blank represents nonreporting or structural inapplicability. No missing value was replaced by zero, imputed or interpolated. This preserves the distinction between absence of information and a recorded absence of generation. It also avoids imposing an unverified missingness mechanism. The need to make such assumptions explicit is established in the National Research Council, 2010, recommendations on missing data.

The complete period requirement removes 18 locations and 250 recorded observations from the accessible frame. Of those observations, 212 are incomplete and 38 contain all four study measures but belong to locations without full period coverage. Complete records from excluded locations have a mean carbon intensity of 341.19 grams of carbon dioxide equivalent per kilowatt hour, compared with 470.63 in the retained panel. This comparison indicates a compositional difference between retained and excluded complete records. It does not establish the characteristics of observations whose intensity is unavailable. The interpretation follows the distinction between documented selection and assumptions about missingness discussed by White and Carlin, 2010, and Hughes et al., 2019. Quality checks confirmed that retained generation shares lie between 0 and 100 percent and carbon intensity is nonnegative. The three included generation shares never exceed 100 percent by more than 0.01 percentage points, allowing for source rounding. No observations were trimmed, winsorized or removed because of their magnitude. Twelve source records report zero carbon intensity. These values were retained as reported and should not be interpreted as independent verification of physically emission free electricity production. No changes to eligibility or numerical values were introduced after inspecting the descriptive results.

3.3 Variable Measurement, Construction, and Validation:

Electricity generation composition is represented by three separate observed measures: coal share, gas share and renewable share. Each expresses the percentage of annual electricity generation attributed to the specified source. They remain separate variables and are not combined into a composite index. Lifecycle carbon intensity is the outcome measure and expresses greenhouse gas emissions per kilowatt hour of generated electricity. Table 2 provides the operational definitions.

Table 2: Variable dictionary

Variable or Indicator	Analytical Role	Definition and Construction	Source and Identifier	Unit and Frequency	Transformation and Direction	Missing Value Rule	Coverage and Limitation
Coal generation share	Generation composition dimension	Source reported coal electricity generation as a percentage of total electricity generation	OWID legacy snapshot; coal share elec	Percent; annual	Untransformed; higher values indicate greater coal participation	Retain location only with complete coverage for all four variables	2,828 observations; does not distinguish coal technologies
Gas generation share	Generation composition dimension	Source reported gas electricity generation as a percentage of total electricity generation	OWID legacy snapshot; gas share elec	Percent; annual	Untransformed; higher values indicate greater gas participation	Same complete coverage rule	2,828 observations; does not distinguish conversion technology or fuel supply conditions
Renewable generation share	Generation composition dimension	Source reported renewable electricity generation as a percentage of total electricity generation	OWID legacy snapshot; renewable share elec	Percent; annual	Untransformed; higher values indicate greater renewable participation	Same complete coverage rule	2,828 observations; combines renewable technologies
Lifecycle carbon intensity	Descriptive outcome	Source reported lifecycle greenhouse gas emissions per unit of electricity generated	OWID legacy snapshot; carbon intensity elec	Grams of carbon dioxide equivalent per kilowatt hour; annual	Untransformed; higher values indicate greater emissions intensity	Same complete coverage rule	2,828 observations; modelled measure partly constructed from generation composition
Location identifier	Panel identifier	Source supplied three letter location code	OWID legacy snapshot; iso code	Categorical; annual records	Preserved without recoding	Required for sample eligibility	202 countries and territories
Year	Temporal identifier	Source supplied observation year	OWID legacy snapshot; year	Calendar year; annual	Restricted to 2010-2023	Required for sample eligibility	14 observations per retained location

The coal measure captures generation composition rather than plant efficiency or fuel quality. These distinctions matter because Whitaker et al., 2012, identify variation in lifecycle emissions across coal generation technologies and assessment assumptions. The gas measure likewise combines technologies with different efficiencies and supply chain emissions. O'Donoughue et al., 2014, show that these differences affect lifecycle comparisons. The analysis therefore interprets both shares as aggregate composition measures. The renewable measure aggregates technologies rather than imposing a common emissions coefficient. Dolan and Heath, 2012, establish the importance of technical assumptions in wind lifecycle estimates, while Hsu et al., 2012, demonstrate corresponding sensitivity in photovoltaic estimates. The aggregate share cannot isolate these technology specific differences.

Carbon intensity is retained directly from the source rather than reconstructed from the three included shares. Its lifecycle boundary incorporates emissions beyond electricity generation alone. This interpretation is consistent with the integrated electricity assessment developed by Hertwich et al., 2015. Because generation composition enters the source accounting framework, correspondence between shares and intensity is partly structural rather than an independent validation of an emissions mechanism. All numerical measures remain in their original units. A change from 20 to 21 percent represents one percentage

point, not one percent growth. Recorded zeros remain zeros, and no logarithmic transformation is applied. Validation comprises source matching, identifier checks, coverage checks and numerical consistency checks. The measures are observed source series rather than latent constructs.

Table 3: Descriptive statistics for the retained panel

Variable	Valid Observations	Mean	Standard Deviation	Median	Minimum	Maximum
Coal generation share, percent	2,828	10.94	20.97	0.00	0.00	100.00
Gas generation share, percent	2,828	20.90	29.87	2.30	0.00	100.00
Renewable generation share, percent	2,828	32.00	32.30	20.24	0.00	100.00
Lifecycle carbon intensity, grams of carbon dioxide equivalent per kilowatt hour	2,828	470.63	246.38	522.72	0.00	1,306.72

The statistics give equal weight to each location and year. They do not estimate generation weighted global electricity composition or global emissions intensity. Coal's zero median and gas's median of 2.30 percent indicate that their pooled distributions differ substantially from their means. Renewable generation also varies widely, while the intensity distribution spans electricity systems with markedly different reported emissions characteristics.

3.4 Analytical Framework, Identification, and Reproducibility:

The analysis is descriptive. Its quantities of interest are the distributions of generation shares and carbon intensity, their annual location averages and their changes within a consistent geographical sample. This purpose is distinct from causal explanation and prediction, as emphasized by Shmueli, 2010. The analytical calculations comprise valid counts, arithmetic means, sample standard deviations, medians, minima, maxima and annual arithmetic averages across the 202 retained locations. The annual averages show a renewable generation share of 29.55 percent in 2010 and 36.38 percent in 2023. Over the same interval, average coal share changes from 11.12 to 9.40 percent, gas share from 20.00 to 21.72 percent and carbon intensity from 483.31 to 439.13 grams of carbon dioxide equivalent per kilowatt hour. These are changes in equal location averages within the balanced sample. Their joint movement does not identify the contribution of an individual generation source. Generation shares carry relative information. An increase in one source necessarily changes the composition represented by the remaining sources. Greenacre, 2021, explains why this constraint matters for interpreting percentage data. Lovell et al., 2015, further demonstrate that conventional correlations between relative components can be misleading. The analysis consequently describes the measured composition without interpreting isolated share changes as independent interventions.

An exploratory sensitivity calculation relaxes the requirement for complete coverage in all 14 years while retaining complete annual records. This alternative sample contains 205 locations and 2,866 observations. Its pooled mean carbon intensity is 468.92, compared with 470.63 in the balanced panel. Mean coal, gas and renewable shares are 10.95, 20.65 and 32.06 percent, respectively. This comparison assesses sensitivity of pooled means to the balancing rule. It changes sample composition and does not resolve potential bias from unavailable observations. Reproducibility rests on preserving the source snapshot identifier, series names, eligibility rules, coverage record and executable extraction logic. The accompanying code reconstructs the panel, checks the source checksum, verifies record uniqueness and numerical bounds, and reproduces the descriptive summaries. This separation of source data, processing and output follows the computational reproducibility principles articulated by Peng, 2011, and Sandve et al., 2013. The four study measures share the same annual frequency, geographical identifiers and retained observation set. Their units and missing value rules remain consistent throughout. Interpretation is restricted to national and territorial electricity generation composition and source reported lifecycle intensity during 2010-2023.

4. Methods:

4.1 Research Design and Analytical Scope:

This analysis uses a descriptive longitudinal design to examine how electricity generation composition corresponds to lifecycle carbon intensity across countries and territories during 2010-2023. The analytical unit is the location and calendar year. The research questions concern differences across electricity systems and changes within the same systems. The design therefore combines pooled distributional comparisons, annual averages and paired endpoint comparisons. These procedures describe observed relationships rather than estimate causal effects or predictive performance, maintaining the distinction between analytical purposes emphasized by Shmueli, 2010. The main sample is the balanced panel documented in Table 1, comprising 202 locations and 2,828 annual observations. Every retained location contributes observations for all 14 years. Applying a constant geographical sample ensures that changes in annual averages reflect changes among retained locations rather than entry into or exit from the sample. Table 2 supplies the measurement definitions and source identifiers used throughout the calculations.

The analysis treats generation shares as components of electricity composition. Their interpretation depends on the remaining generation portfolio because an increase in one component requires a decrease in others. Greenacre, 2021, explains why this relative structure affects the interpretation of percentage data. Accordingly, comparisons describe the electricity systems represented by each share category without treating category membership as an isolated intervention. The source carbon intensity measure also incorporates generation composition within its accounting framework, which constrains interpretation of their correspondence.

4.2 Distributional Summaries and Temporal Comparisons:

The numerical series retain their source units and precision. Generation shares are expressed as percentages of annual electricity production, and lifecycle intensity is expressed in grams of carbon dioxide equivalent per kilowatt hour. No normalization, logarithmic transformation or composite index construction is applied. Recorded zeros remain numerical observations. For each measure, the analysis calculates the valid count, arithmetic mean, sample standard deviation, median, minimum and maximum. Equation 1 defines the pooled mean and sample standard deviation:

$$\bar{x}_v = \frac{1}{n}\sum_{i=1}^{N}\sum_{t=2010}^{2023} x_{vit}, \qquad s_v = \sqrt{\frac{\sum_{i=1}^{N}\sum_{t=2010}^{2023}[x_{vit} - \bar{x}_v]^2}{n-1}}.$$

Here, x_{vit} denotes the observed value of measure vfor location iin year t; v indexes the three generation shares and carbon intensity; $N = 202$is the number of locations; $n = 2828$is the number of pooled observations; $\bar{x}_v$is the pooled arithmetic mean; and $s_v$is the sample standard deviation. The median describes the central ordered observation, while the minimum and maximum retain the observed range. Annual averages are calculated over the same locations using Equation 2:

$$\bar{x}_{vt} = \frac{1}{N}\sum_{i=1}^{N} x_{vit}.$$

Here, $\bar{x}_{vt}$ denotes the arithmetic mean of measure v in year t. Every location receives equal weight. The averages therefore characterize the sampled electricity systems rather than generation weighted global electricity production. Temporal comparisons examine the complete annual sequence and the difference between the averages for 2023 and 2010. Changes in generation shares are expressed in percentage points. Changes in intensity retain its original emissions unit. The relative change in average intensity is calculated by dividing its endpoint difference by its 2010 average and multiplying by 100. Calculations use unrounded values; rounding occurs only when presenting results.

4.3 Generation Share Bands and Paired Location Changes:

Exploratory band comparisons examine how average carbon intensity differs across generation compositions. Each generation dimension is classified separately into five mutually exclusive categories: zero, above zero through 25 percent, above 25 through 50 percent, above 50 through 75 percent, and above 75 through 100 percent. A separate zero category preserves the substantive distinction between no reported generation from a source and positive participation. The remaining intervals provide fixed descriptive boundaries rather than estimated thresholds. Equation 3 defines average intensity within each band:

$$\bar{c}_{db} = \frac{\sum_{i=1}^{N}\sum_{t=2010}^{2023} I_{dbit}\, c_{it}}{n_{db}}, \qquad n_{db} = \sum_{i=1}^{N}\sum_{t=2010}^{2023} I_{dbit}.$$

Here, c_{it} denotes carbon intensity for location i in year t; d identifies the generation dimension; b identifies its share band; I_{dbit} equals one when the observation belongs to that band and zero otherwise; n_{db} is the number of observations in the band; and $\bar{c}_{db}$is its arithmetic mean intensity. Each dimension partitions all 2,828 observations. Categories across different dimensions can contain the same observations. Paired comparisons then examine changes within each location between 2010 and 2023. Equation 4 defines the endpoint differences:

$$\Delta x_{di} = x_{di,2023} - x_{di,2010}, \qquad \Delta c_i = c_{i,2023} - c_{i,2010}.$$

Here, Δx_{di} denotes the percentage point change in generation dimension d for location i, and Δc_idenotes the corresponding intensity change. Locations are classified as recording an increase, decrease or unchanged share according to whether $\Delta x_{di}$is positive, negative or zero. Classification uses source precision without an additional tolerance. Within each share movement group, the analysis counts locations with declining, increasing and unchanged intensity. It also calculates mean intensity change using Equation 5:

$$\overline{\Delta c}_{dg} = \frac{1}{N_{dg}}\sum_{i \in G_{dg}} \Delta c_i.$$

Here, G_{dg} is the set of locations assigned to movement category gfor dimension d; $N_{dg}$is the number of locations in that set; and $\overline{\Delta c}_{dg}$ is their mean intensity change. The corresponding median is calculated from the ordered location changes. Reported proportions divide the relevant intensity movement count by N_{dg}. The endpoint comparisons complement the pooled bands by distinguishing differences across observed compositions from changes within locations. They do not reconstruct the intervening sequence or identify the electricity source displaced. Both the band comparisons and endpoint classifications are exploratory extensions of the descriptive audit.

4.4 Quality Assurance, Sensitivity and Reproducibility:

Processing preserves one observation per row and one variable per column, following the spreadsheet organization principles described by Broman and Woo, 2018. The location identifier and year jointly define the record key. Checks assess key uniqueness, valid observation years, numerical completeness and agreement between the retained workbook values and the source extraction. Numerical checks require each generation share to lie between zero and 100 percent and carbon intensity to be nonnegative. The combined share of the three included generation dimensions is checked against 100 percent with a tolerance of 0.01 percentage points for source rounding. No upper intensity threshold is imposed. Observations are not trimmed or winsorized, and valid extreme values remain in the summaries.

Missing values are handled through the complete coverage rule documented in Table 1. No imputation, interpolation or replacement of missing values with zero is performed. Selection is therefore explicit, but complete coverage is not interpreted as evidence that missingness is random. Hughes et al., 2019, emphasize that the suitability of missing data procedures depends on their underlying assumptions rather than a universal preference for a particular technique. An exploratory sensitivity calculation relaxes the requirement for complete coverage in every year while retaining only annual records containing all four study measures. This produces an unbalanced comparison sample of 205 locations and 2,866 observations. Pooled means are recalculated and compared with the balanced panel means. Because this procedure changes the sample, its interpretation concerns sensitivity to the balancing restriction. It does not establish robustness of the band comparisons or endpoint classifications.

Calculations use Python and standard library functions for CSV processing, grouping, arithmetic summaries and checksum verification. The audit script checks the source file against its recorded SHA256 checksum and reconstructs the retained panel, coverage record, distributional statistics, annual averages and alternative sample means. The band and endpoint procedures

follow Equations 3-5. Source precision is preserved during computation, and displayed summaries are rounded to two decimal places. The documented workflow separates source identification, eligibility decisions, calculations and reported outputs, consistent with the reproducibility principles articulated by Sandve et al., 2013. Its analytical strength lies in applying consistent measurements and comparison rules across the same locations and periods. The resulting quantities answer the research questions as descriptive patterns in observed electricity composition and reported lifecycle intensity.

5. Findings:

This analysis examines how coal, gas and renewable generation shares correspond to lifecycle electricity carbon intensity during 2010-2023. The findings distinguish differences across electricity systems from changes within locations. Distributional summaries, generation share comparisons and sample sensitivity checks provide descriptive answers to the three research questions. The analytical sample contains 2,828 annual observations across 202 countries and territories, with 14 observations per location. All four study variables are complete within this balanced panel. Selection removed 18 locations from the accessible frame of 220 because they lacked complete coverage across the observation period. The results give equal weight to each location and year. They therefore describe the sampled electricity systems rather than generation weighted global electricity supply. Table 4 presents the distribution of the measured variables. Coal generation has a median share of zero, reflecting 1,703 observations without reported coal generation. Gas generation also has a concentrated distribution: its mean share is 20.90 percent, compared with a median of 2.30 percent, and 1,240 observations report zero gas generation. These differences establish that average fossil generation shares do not represent a typical electricity system uniformly.

Table 4: Descriptive statistics for the balanced panel

Variable	Observations	Mean	Standard Deviation	Median	Minimum	Maximum
Coal generation share, percent	2,828	10.94	20.97	0.00	0.00	100.00
Gas generation share, percent	2,828	20.90	29.87	2.30	0.00	100.00
Renewable generation share, percent	2,828	32.00	32.30	20.24	0.00	100.00
Lifecycle carbon intensity, grams of carbon dioxide equivalent per kilowatt hour	2,828	470.63	246.38	522.72	0.00	1,306.72

Renewable generation has a mean share of 32.00 percent and a median of 20.24 percent. Its standard deviation of 32.30 percentage points indicates substantial variation across the pooled observations. Carbon intensity averages 470.63 grams of carbon dioxide equivalent per kilowatt hour, with a standard deviation of 246.38. The reported range extends from zero to 1,306.72. The twelve zero intensity observations are retained source values and do not independently establish an absence of lifecycle emissions. The annual comparisons show a net movement toward greater renewable participation and lower average carbon intensity. Between 2010 and 2023, renewable generation share increases by 6.82 percentage points, while coal share decreases by 1.72 percentage points. Gas share increases by 1.72 percentage points. Average carbon intensity declines by 44.18 grams of carbon dioxide equivalent per kilowatt hour, equivalent to 9.14 percent of its initial value.

Table 5: Changes in equal location averages

Measure	2010	2023	Net Change
Coal generation share, percent	11.12	9.40	Decrease of 1.72 percentage points
Gas generation share, percent	20.00	21.72	Increase of 1.72 percentage points
Renewable generation share, percent	29.55	36.38	Increase of 6.82 percentage points
Carbon intensity, grams of carbon dioxide equivalent per kilowatt hour	483.31	439.13	Decrease of 44.18

The decline in average carbon intensity is not continuous throughout the interval. The annual average increases from 483.31 in 2010 to 485.18 in 2011 and rises again during 2014 and 2015. It subsequently declines each year from 2015 through 2023. The endpoint comparison therefore captures a net reduction alongside shorter periods of reversal. To examine the research questions directly, exploratory comparisons group observations into fixed generation share bands. Zero shares form a separate category, followed by four intervals of 25 percentage points. Table 6 reports observation counts and average carbon intensity within each band. These comparisons extend the audit summaries and describe pooled composition differences without adjusting for other generation sources.

Table 6: Carbon intensity across generation share bands

Generation Share Band, Percent	Coal Observations	Coal Group Intensity	Gas Observations	Gas Group Intensity	Renewable Observations	Renewable Group Intensity
Zero	1,703	478.51	1,240	459.39	399	717.42
Above zero through 25	623	340.03	740	392.15	1,162	605.86
Above 25 through 50	290	514.05	353	457.13	488	433.62
Above 50 through 75	144	693.06	221	549.86	349	262.30
Above 75 through 100	68	813.77	274	686.99	430	87.31

For RQ1, carbon intensity increases across the positive coal share bands. Observations with coal shares above zero through 25 percent average 340.03 grams of carbon dioxide equivalent per kilowatt hour, compared with 813.77 where coal exceeds 75 percent. However, the zero coal group averages 478.51, above the lowest positive coal group. Absence of coal therefore does not uniformly identify lower intensity systems because the remaining generation mix differs across observations. The positive coal band gradient is consistent with the lifecycle emissions burden documented by Whitaker et al., 2012, without establishing a coal substitution effect. Changes within locations reinforce the importance of distinguishing coal reductions from

coal increases. Among 51 locations with lower coal shares in 2023 than in 2010, 47 also record lower carbon intensity. Their average intensity reduction is 108.54 grams of carbon dioxide equivalent per kilowatt hour. Among 33 locations with higher coal shares, intensity declines in 17 and increases in 16. The latter group records an average intensity increase of 13.84 despite a median decrease of 6.88, indicating that the mean does not summarize a uniform directional pattern.

For RQ2, gas share does not provide a consistent descriptive classification of lower intensity electricity systems. Mean intensity increases across the positive gas share bands, from 392.15 where gas contributes above zero through 25 percent to 686.99 where it exceeds 75 percent. Yet the zero gas group averages 459.39, exceeding the lowest positive gas band. This pattern is compatible with gas having different implications depending on the sources alongside it. O'Donoughue et al., 2014, similarly distinguish lifecycle performance across gas technologies and supply conditions. The endpoint comparisons also show mixed gas patterns. Of 59 locations with increasing gas shares, 37 record declining intensity and 22 record increasing intensity. Their average intensity change is a decrease of 26.40 grams of carbon dioxide equivalent per kilowatt hour. Intensity also declines in 50 of the 63 locations with decreasing gas shares. Consequently, both gas expansion and gas contraction accompany declining intensity in substantial parts of the sample. Gas share alone does not resolve which generation substitutions underlie these changes.

For RQ3, renewable share produces the clearest ordered pattern across the pooled generation bands. Mean intensity decreases from 717.42 in observations with no renewable generation to 87.31 where renewable generation exceeds 75 percent. Every successive renewable share band has a lower average intensity. This ordering is consistent with the lifecycle evidence on wind and photovoltaic generation reported by Dolan and Heath, 2012, and Hsu et al., 2012. The aggregate renewable measure nevertheless combines technologies rather than estimating their separate contributions. Within locations, 136 record an increase in renewable share between 2010 and 2023. Carbon intensity declines in 125 of these locations, increases in ten and remains unchanged in one. The proportion recording declining intensity is 91.91 percent. Conversely, among 46 locations with decreasing renewable shares, 41 record increasing intensity. These paired movements provide descriptive consistency between renewable expansion and lower intensity while retaining the contradictory cases.

Table 7: Joint changes in generation shares and carbon intensity between 2010 and 2023

Generation Dimension	Share Movement	Locations	Intensity Declines	Intensity Increases	Intensity Unchanged	Mean Intensity Change
Coal	Increase	33	17	16	0	Increase of 13.84
Coal	Decrease	51	47	4	0	Decrease of 108.54
Coal	Unchanged	118	75	38	5	Decrease of 32.58
Gas	Increase	59	37	22	0	Decrease of 26.40
Gas	Decrease	63	50	13	0	Decrease of 75.65
Gas	Unchanged	80	52	23	5	Decrease of 32.51
Renewables	Increase	136	125	10	1	Decrease of 86.56
Renewables	Decrease	46	5	41	0	Increase of 62.71
Renewables	Unchanged	20	9	7	4	Decrease of 1.81

The renewable expansion group records an average intensity reduction of 86.56 grams of carbon dioxide equivalent per kilowatt hour, compared with an average increase of 62.71 in the renewable contraction group. This contrast is consistent with the broader lifecycle assessment of electricity supply transitions presented by Hertwich et al., 2015. Its interpretation remains descriptive because the outcome incorporates generation composition through the source emissions accounting framework. The sample sensitivity calculation relaxes the requirement for complete coverage in all 14 years while retaining complete annual records. The resulting sample contains 205 locations and 2,866 observations. Mean intensity is 468.92, compared with 470.63 in the balanced panel, a difference of 1.72 grams of carbon dioxide equivalent per kilowatt hour. Mean coal, gas and renewable shares are 10.95, 20.65 and 32.06 percent, respectively. These small differences indicate limited sensitivity of pooled means to the balancing restriction. They do not establish that the generation band comparisons or endpoint patterns are equally insensitive. The research questions receive distinct descriptive answers. Coal reductions commonly accompany declining intensity, but coal expansion produces mixed outcomes. Gas increases and decreases both accompany intensity reductions, making its aggregate relationship conditional on the wider electricity mix. Renewable expansion shows the strongest consistency across pooled share bands and paired location changes. Together, the findings establish differentiated composition patterns within the observed sample rather than independent causal effects of individual generation sources.

6. Discussion:

A change in electricity generation share does not carry a uniform carbon implication across electricity systems. Tables 6 and 7 reveal an asymmetry: renewable expansion accompanies lower intensity consistently, coal reductions commonly accompany improvement, and gas changes produce mixed patterns. The contribution of this analysis lies in distinguishing pooled composition differences from paired changes across the same 202 locations. These comparisons show why the direction of a generation share change cannot substitute for examination of the surrounding electricity mix. Hawkes, 2010, distinguishes average electricity emissions from the marginal emissions relevant to demand interventions. The present evidence adds a comparative descriptive perspective: even within average intensity accounting, the interpretation of a source share depends on the composition against which it changes. The findings therefore refine the characterization of electricity transitions without attributing independent causal effects to individual sources. The gas findings sharpen this compositional interpretation. Table 7 shows declining intensity among locations with both increasing and decreasing gas shares. Gas expansion therefore cannot be treated as a sufficient indicator of electricity decarbonization. The observed pattern is compatible with different substitutions, but the aggregate measures do not identify which generating technologies were displaced. Cullen, 2013, evaluates renewable electricity benefits through the generation it replaces, demonstrating why displacement matters for environmental interpretation. Here, the contribution is to

establish that opposite movements in gas share can accompany the same direction of intensity change across national and territorial systems. This shifts attention from fuel labels toward joint changes in electricity composition. It also prevents the descriptive evidence from being interpreted as a universal endorsement or rejection of gas expansion.

Renewable generation exhibits the strongest descriptive consistency across the two comparisons. Table 6 shows progressively lower mean intensity across successive renewable share bands, while Table 7 shows declining intensity in 125 of the 136 locations with renewable expansion. This correspondence is stronger than the mixed pattern associated with gas expansion, although it does not establish that renewable generation accounts for a particular proportion of the overall intensity reduction. Siler Evans et al., 2012, demonstrate that marginal emissions benefits vary across electricity regions because the generation responding to demand differs. The present results complement that evidence by showing a consistent aggregate renewable pattern alongside substantial variation in fossil generation relationships. The theoretical implication is that compositional coherence and causal attribution are separate: the former is visible in these data, while the latter requires identification of the generation displaced. The results also expose the limits of classifying electricity systems through the absence of a single fuel. In Table 6, observations without coal have higher average intensity than observations in the lowest positive coal share band. This does not indicate an environmental benefit from introducing coal. It indicates that the remaining generation portfolio determines the meaning of coal's absence. Table 4 further shows that pooled averages summarize highly heterogeneous systems rather than a representative national mix. De Chalendar et al., 2019, trace emissions through production, electricity exchanges and consumption within the United States, demonstrating that geographical accounting boundaries affect emissions interpretation. Across the broader sample examined here, production based annual comparisons provide a common descriptive basis, but they should not be interpreted as consumption footprints. The international contribution is a comparable account of electricity composition across diverse systems, without assuming that national boundaries capture every electricity related emissions relationship.

For decision makers, the evidence supports evaluating generation composition jointly with measured intensity rather than using renewable, coal or gas shares as interchangeable carbon indicators. Tables 5 and 7 distinguish an improving sample average from heterogeneous location trajectories, making both levels relevant to comparative assessment. For carbon aware computing, these findings provide background information about electricity supply conditions rather than evidence of workload scheduling savings. Masanet et al., 2020, show the importance of explicitly accounting for activity and efficiency when assessing data centre electricity use. Applied to the present results, this means that electricity intensity comparisons must remain distinct from claims about computational energy demand or operational emissions reductions. The theoretical extension is therefore precise: generation composition helps characterize the carbon environment of electricity dependent activities, while conclusions about particular operational decisions require matching that environment to the activity being evaluated.

7. Conclusion and Implications:

Electricity transitions cannot be judged by changes in a single generation source. This analysis shows that similar shifts can accompany different carbon trajectories, while comparable improvements can emerge from contrasting changes in supply composition. Its contribution is a comparative framework that distinguishes differences across electricity systems from changes within the same systems. The resulting asymmetry sharpens interpretation: some transition patterns correspond consistently with improvement, whereas others depend on the surrounding production portfolio. This refines a compositional understanding of electricity decarbonization without treating descriptive correspondence as evidence of independent causal effects. For managers, the findings support evaluating electricity supply through joint portfolio assessment rather than isolated technology labels. Resource allocation and environmental reporting should therefore connect changes in supply composition with observed emissions performance. Policymakers should assess transition progress against measured outcomes and recognize that identical directional targets can have different implications across electricity systems. Operationally, organizations can strengthen reporting routines by preserving consistent accounting boundaries, distinguishing annual comparisons from intervention assessments, and documenting the basis of environmental claims. These practices can improve the credibility of public information and investment decisions. Across diverse economies, the evidence supports a shared principle: assess electricity transitions through contextual performance rather than assuming uniform benefits from compositional change.

The annual panel provides opportunities for finer temporal and technological investigation. Hourly observations could distinguish persistent composition patterns from operational variation, while generation and electricity exchange records could clarify displacement and consumption boundaries. Because reported intensity partly incorporates generation composition, independent emissions measurements would strengthen external validation. Broader coverage could assess sensitivity to excluded locations, and generation weighted comparisons could distinguish typical system trajectories from aggregate electricity outcomes. Designs using identifiable policy or infrastructure changes could investigate causal pathways. These extensions would establish when the observed relationships persist across institutional settings and when technological or market conditions alter their interpretation.

References:

1. Broman, K. W., & Woo, K. H. 2018. Data organization in spreadsheets. The American Statistician, volume 72, issue 1, pages 2-10. https://doi.org/10.1080/00031305.2017.1375989
2. Cullen, J. 2013. Measuring the environmental benefits of wind generated electricity. American Economic Journal: Economic Policy, Volume 5, Issue 4, Pages 107-133. https://doi.org/10.1257/pol.5.4.107
3. De Chalendar, J. A., Taggart, J., & Benson, S. M. 2019. Tracking emissions in the US electricity system. Proceedings of the National Academy of Sciences, volume 116, issue 51, pages 25497-25502. https://doi.org/10.1073/pnas.1912950116
4. Dolan, S. L., & Heath, G. A. 2012. Life cycle greenhouse gas emissions of utility scale wind power: Systematic review and harmonization. Journal of Industrial Ecology, volume 16, supplement 1, pages S136-S154. https://doi.org/10.1111/j.1530-9290.2012.00464.x
5. Greenacre, M. 2021. Compositional data analysis. Annual Review of Statistics and Its Application, volume 8, pages 271-299. https://doi.org/10.1146/annurev-statistics-042720-124436

6. Hawkes, A. D. 2010. Estimating marginal CO_2 emissions rates for national electricity systems. Energy Policy, volume 38, issue 10, pages 5977-5987. https://doi.org/10.1016/j.enpol.2010.05.053
7. Hertwich, E. G., Gibon, T., Bouman, E. A., Arvesen, A., Suh, S., Heath, G. A., Bergesen, J. D., Ramirez, A., Vega, M. I., & Shi, L. 2015. Integrated lifecycle assessment of electricity supply scenarios confirms global environmental benefit of low carbon technologies. Proceedings of the National Academy of Sciences, volume 112, issue 20, pages 6277-6282. https://doi.org/10.1073/pnas.1312753111
8. Hirth, L. 2020. Open data for electricity modeling: Legal aspects. Energy Strategy Reviews, volume 27, Article 100433. https://doi.org/10.1016/j.esr.2019.100433
9. Hsu, D. D., O'Donoughue, P., Fthenakis, V., Heath, G. A., Kim, H. C., Sawyer, P., Choi, J. K., & Turney, D. E. 2012. Life cycle greenhouse gas emissions of crystalline silicon photovoltaic electricity generation: Systematic review and harmonization. Journal of Industrial Ecology, volume 16, supplement 1, pages S122-S135. https://doi.org/10.1111/j.1530-9290.2011.00439.x
10. Hughes, R. A., Heron, J., Sterne, J. A. C., & Tilling, K. 2019. Accounting for missing data in statistical analyses: Multiple imputation is not always the answer. International Journal of Epidemiology, volume 48, issue 4, pages 1294-1304. https://doi.org/10.1093/ije/dyz032
11. Lovell, D., Pawlowsky Glahn, V., Egozcue, J. J., Marguerat, S., &Bähler, J. 2015. Proportionality: A valid alternative to correlation for relative data. PLOS Computational Biology, volume 11, issue 3, Article e1004075. https://doi.org/10.1371/journal.pcbi.1004075
12. Masanet, E., Shehabi, A., Lei, N., Smith, S., &Koomey, J. 2020. Recalibrating global data center energy use estimates. Science, volume 367, issue 6481, pages 984-986. https://doi.org/10.1126/science.aba3758
13. Morrison, R. 2018. Energy system modeling: Public transparency, scientific reproducibility, and open development. Energy Strategy Reviews, volume 20, pages 49-63. https://doi.org/10.1016/j.esr.2017.12.010
14. National Research Council. 2010. The prevention and treatment of missing data in clinical trials. The National Academies Press. https://doi.org/10.17226/12955
15. O'Donoughue, P. R., Heath, G. A., Dolan, S. L., & Vorum, M. 2014. Life cycle greenhouse gas emissions of electricity generated from conventionally produced natural gas: Systematic review and harmonization. Journal of Industrial Ecology, volume 18, issue 1, pages 125-144. https://doi.org/10.1111/jiec.12084
16. Pehl, M., Arvesen, A., Humpenöder, F., Popp, A., Hertwich, E. G., &Luderer, G. 2017. Understanding future emissions from low carbon power systems by integration of lifecycle assessment and integrated energy modelling. Nature Energy, volume 2, issue 12, pages 939-945. https://doi.org/10.1038/s41560-017-0032-9
17. Peng, R. D. 2011. Reproducible research in computational science. Science, volume 334, issue 6060, pages 1226-1227. https://doi.org/10.1126/science.1213847
18. Pfenninger, S., DeCarolis, J., Hirth, L., Quoilin, S., &Staffell, I. 2017. The importance of open data and software: Is energy research lagging behind? Energy Policy, volume 101, pages 211-215. https://doi.org/10.1016/j.enpol.2016.11.046
19. Pfenninger, S., Hirth, L., Schlecht, I., Schmid, E., Wiese, F., Brown, T., Davis, C., Gidden, M., Heinrichs, H., Heuberger, C., Hilpert, S., Krien, U., Matke, C., Nebel, A., Morrison, R., Müller, B., Pleßmann, G., Reeg, M., Richstein, J. C., … Wingenbach, C. 2018. Opening the black box of energy modelling: Strategies and lessons learned. Energy Strategy Reviews, volume 19, pages 63-71. https://doi.org/10.1016/j.esr.2017.12.002
20. Sandve, G. K., Nekrutenko, A., Taylor, J., & Hovig, E. 2013. Ten simple rules for reproducible computational research. PLOS Computational Biology, volume 9, issue 10, Article e1003285. https://doi.org/10.1371/journal.pcbi.1003285
21. Shmueli, G. 2010. To explain or to predict? Statistical Science, volume 25, issue 3, pages 289-310. https://doi.org/10.1214/10-STS330
22. Siler Evans, K., Azevedo, I. L., & Morgan, M. G. 2012. Marginal emissions factors for the U.S. electricity system. Environmental Science & Technology, volume 46, issue 9, pages 4742-4748. https://doi.org/10.1021/es300145v
23. Whitaker, M., Heath, G. A., O'Donoughue, P., & Vorum, M. 2012. Life cycle greenhouse gas emissions of coal fired electricity generation: Systematic review and harmonization. Journal of Industrial Ecology, volume 16, supplement 1, pages S53-S72. https://doi.org/10.1111/j.1530-9290.2012.00465.x
24. White, I. R., & Carlin, J. B. 2010. Bias and efficiency of multiple imputation compared with complete case analysis for missing covariate values. Statistics in Medicine, volume 29, issue 28, pages 2920-2931. https://doi.org/10.1002/sim.3944
25. Wickham, H. 2014. Tidy data. Journal of Statistical Software, volume 59, issue 10, pages 1-23. https://doi.org/10.18637/jss.v059.i10
26. Wiese, F., Schlecht, I., Bunke, W. D., Gerbaulet, C., Hirth, L., Jahn, M., Kunz, F., Lorenz, C., Mühlenpfordt, J., Reimann, J., & Schill, W. P. 2019. Open Power System Data: Frictionless data for electricity system modelling. Applied Energy, volume 236, pages 401-409. https://doi.org/10.1016/j.apenergy.2018.11.097