Uncovering CO2 Drivers with Machine Learning in High- and Upper-Middle-Income Countries

energies

MDPI

Article

Uncovering CO_2 Drivers with Machine Learning in High- and Upper-Middle-Income Countries

Cosimo Magazzino [1,2,*], Umberto Monarca [3], Ernesto Cassetta [4], Alberto Costantiello [5] and Tulia Gattone [6,7]

[1] Economic Research Center, Western Caspian University, AZ1001 Baku, Azerbaijan
[2] ARUCAD Research Centre, Arkin University of Creative Arts and Design, Kyrenia, Northern Cyprus, Mersin 10, Turkey
[3] Department of Law, University of Foggia, 71121 Foggia, Italy; umberto.monarca@unifg.it
[4] Department of Economics and Statistics, University of Udine, 33100 Udine, Italy; ernesto.cassetta@uniud.it
[5] Department of Management, Technology and Finance, LUM University "Giuseppe Degennaro", 70010 Bari, Italy; costantiello@lum.it
[6] Department of Economics, John Cabot University, 00165 Rome, Italy; tulia.gattone@johncabot.edu
[7] Syracuse University in Florence, 50132 Florence, Italy
* Correspondence: cosimomagazzino@hotmail.com

Abstract

Rapid decarbonization relies on knowing which structural and energy factors affect national carbon dioxide emissions. Much of the literature leans on linear and additive assumptions, which may gloss over curvature and interactions in this energy–emissions link. Unlike previous studies, we take a different approach. Using a panel of 80 high- and upper-middle-income countries from 2011 to 2020, we model emissions as a function of fossil fuel energy consumption, methane, the food production index, renewable electricity output, gross domestic product (GDP), and trade measured as trade over GDP. Our contribution is twofold. First, we evaluate how different modeling strategies, from a traditional Generalized Linear Model to more flexible approaches such as Support Vector Machine regression and Random Forest (RF), influence the identification of emission drivers. Second, we use Double Machine Learning (DML) to estimate the incremental effect of fossil fuel consumption while controlling for other variables, offering a more careful interpretation of its likely causal role. Across models, a clear pattern emerges: GDP dominates; fossil fuel energy consumption and methane follow. Renewable electricity output and trade contribute, but to a moderate degree. The food production index adds little in this aggregate, cross-country setting. To probe the mechanism rather than the prediction, we estimate the incremental role of fossil fuel energy consumption using DML with RF nuisance functions. The partial effect remains positive after conditioning on the other covariates. Taken together, the results suggest that economic scale and the fuel mix are the primary levers for near-term emissions profiles, while renewables and trade matter, just less than is often assumed and in ways that may depend on context.

Keywords: CO_2 emissions; fossil fuels; support vector machine; random forest; double machine learning

check for updates

Academic Editor: Piotr Kosowski

Received: 22 September 2025
Revised: 12 October 2025
Accepted: 15 October 2025
Published: 22 October 2025

Citation: Magazzino, C.; Monarca, U.; Cassetta, E.; Costantiello, A.; Gattone, T. Uncovering CO_2 Drivers with Machine Learning in High- and Upper-Middle-Income Countries. *Energies* **2025**, *18*, 5552. https://doi.org/10.3390/en18215552

1. Introduction

Fossil fuels still dominate the global energy mix, largely because rising fossil-fueled energy demand has offset efficiency gains and clean-energy rollouts; this "treadmill" dynamic helps explain why mitigation remains difficult even as wind and solar expand

(Murphy, 2024 [1]). Within advanced economies, the energy mix is heavily fossil-based and, as a result, contributes materially to global emissions. In contrast, renewable energy is associated with lower emissions, supporting long-term planning to reduce the fossil share (Hou et al., 2023 [2]). Long-range, target-oriented scenarios have indicated that a different path is technically feasible: under resilient efficiency measures and coordinated supply strategies, renewables could provide about half of global energy by 2050, with Organization for Economic Co-operation and Development (OECD) country emissions falling sharply and developing-country emissions stabilizing (Krewitt et al., 2007 [3]).

Grid operation, however, faces a number of concrete technical hurdles as variable renewables scale. Recent studies increasingly stress the need for power-system flexibility, through planning, operations, and new resources, to manage variability and uncertainty at high renewable penetration (Rahman et al., 2024 [4]). Country-specific analyses point to seasonal storage as a binding challenge and emphasize how full electrification could be most efficient and how photovoltaic and hydropower sources could be the cheapest without storage, while future nuclear options might deliver continuous, low-impact generation (Züttel et al., 2024 [5]).

Meanwhile, deployment trends are significant but still rather uneven. Recent analyses report energy-related carbon dioxide (CO_2) rising to about 37.4 billion tons in 2023, underscoring rapid cost declines in solar, wind, and storage that enable large-scale rollout (Ashraf & Sagheer, 2025 [6]). At the same time, financial development tends to lower emissions, and industrialization has heterogeneous effects (Haloui et al., 2025 [7]). Complementing these results, other studies document both the growth in renewable capacity since 2010 and multi-dimensional economic, technical, social, environmental, and institutional barriers, alongside a discussion of how renewables contribute to multiple sustainable development goals (Olabi et al., 2023 [8]). Workforce constraints also matter. Updated employment factors and occupational shares reveal volatility in projected jobs for large-scale renewable build-outs and point to the need for integrated workforce planning (Rutovitz et al., 2025 [9]). Regarding demand, a survey of European evidence catalogs wide-ranging co-benefits from energy demand reduction and proposes steps to standardize and better quantify them for policymaking (Finn & Brockway, 2023 [10]). Assessments of the transition's overall scale suggest that entirely replacing fossil fuels by 2050 may be possible only if multiple pathways, such as renewable build-out, efficiency, conservation, carbon pricing and trading, carbon capture, carbon utilization, carbon storage, nuclear, and lifestyle changes, are pursued aggressively, with renewable output increasing severalfold under most demand scenarios (Holechek et al., 2022 [11]).

Country studies usefully add further texture to this framework. For example, in Pakistan, "green industrial transformation" is associated with lower carbon intensity. At the same time, foreign direct investment (FDI) inflows, technological innovation, and research and development (R&D) are linked to higher emissions, consistent with a pollution haven mechanism (Mehmood et al., 2024 [12]). At the European Union regional level, there exist sizable differences in the pace and profile of energy, with a small group advancing most quickly and several countries lagging (Brodny et al., 2025 [13]). Looking at per capita outcomes, global trends show potential persistence and the shifting of the "center of gravity" of per capita CO_2 since the early 2000s, showing how diverging regional trajectories shape aggregate patterns (Yang et al., 2023 [14]). Firm-level and sectoral perspectives also identify adoption barriers to low-carbon operations in emerging-economy manufacturing as especially influential (Kumar et al., 2023 [15]).

Methodologically, much recent work still relies on econometric panels or prediction-oriented Machine Learning (ML). In environmental applications, ML has been used effectively for forecasting and threshold analysis, such as projecting atmospheric CO_2 levels and identifying reduction targets, with recommended shifts toward renewables and carbon neutrality as policy goals (see, for example, Bhatt et al., 2023 [16]), and for modeling spatiotemporal greenhouse gas fluxes where sequence models outperform other approaches (e.g., Hamrani et al., 2020 [17]). There is also work combining econometric and ML methods in global panels to study how fossil energy, methane, and renewable energy correlate with CO_2, with findings of strong associations and policy implications drawn for expanding renewables and improving resource management (Gattone et al., 2025 [18]). However, in contrast to studies with a global perspective, this study focuses on high- and upper-middle-income economies and refines the variable set to examine how economic scale and the fuel mix relate to CO_2 outcomes, while aiming to balance predictive performance with transparent interpretation.

Despite these advances, several gaps remain. Much of the current evidence comes from global or mixed-income samples, which may obscure the specific dynamics of higher-income economies where energy transitions are already underway. The interplay between economic scale, fossil fuel intensity, and nonlinear emission responses also remains insufficiently explored. Many studies focus on prediction or single-model estimation rather than comparing methods with different structural assumptions. These gaps motivate our research question: "Which structural and energy factors most strongly drive national CO_2 emissions once nonlinearity and interaction effects are considered?".

In this work, we construct a panel of 80 high- and upper-middle-income countries from 2011 to 2020 and model emissions as a function of fossil fuel energy consumption, methane, the food production index, renewable electricity output, GDP, and trade measured as trade over GDP. We compare a Generalized Linear Model (GLM) with Gaussian identity, a Support Vector Machine (SVM) for regression with a radial kernel, and Random Forest (RF) then compute test-set permutation importance to deliver a standard ranking of contributors across learners. To interpret the incremental role of fossil fuel energy consumption and methane after flexibly conditioning on the other covariates, we implement Double Machine Learning (DML) with RF.

Our contribution is, in essence, threefold. First, we offer a like-for-like assessment of scale, fossil intensity, non-CO_2 greenhouse gases, renewables, trade openness, and food-system activity in a harmonized panel of higher-income economies during the 2010s. Second, we provide model-agnostic importance based on held-out data so that relative rankings do not hinge on a single functional form. Third, as robustness checks, we estimate the partial effect of fossil fuel energy consumption using DML to separate association from the conditional contribution of the treatment variable under potentially nonlinear controls.

The remainder of this article unfolds as follows. Section 2 reviews the literature, while Section 3 introduces the conceptual framework. Section 4 sets out the data and empirical design, and Section 5 presents the main results, followed by robustness checks in Section 6. Section 7 concludes, discussing the policy lessons and giving directions for further research.

2. Literature Review

Research on cross-country emissions consistently shows how structural and energy factors determine whether renewable expansion actually translates into mitigation. For example, in emerging economies, renewable energy investment seems to lower CO_2 only when supported by effective governance and green finance (Yadav et al., 2024 [19]). Likewise, evidence from China places finance at the center of the transition. Furthermore, cointegration and causality tests reveal that green finance and private-sector participation are positively linked to renewable-related economic activity and innovation, with investment, trade, and human development acting as essential moderators (Li & Umair, 2023 [20]).

Global modeling reinforces the importance of this policy design. Removing fossil fuel subsidies and introducing carbon pricing, with part of the revenue recycled into renewables and efficiency, is projected to reduce emissions by around seven percent by 2030 and close to twenty percent by 2050 (Bassi et al., 2023 [21]). In parallel, the rapid expansion of artificial intelligence infrastructure may strain energy systems and climate goals unless energy sourcing strategies, storage, and policy frameworks are carefully aligned (Lal and You, 2025 [22]). Taken together, studies point to governance, finance, and pricing instruments as decisive channels that seem to shape the contribution of renewable deployment to mitigation.

In this framework, solar power provides a valuable case where results converge on emission reductions but also, as observed, highlight uneven outcomes and externalities. For instance, for the largest solar energy consumers, quantile-on-quantile estimation shows that higher solar consumption generally lowers CO_2 across the emission distribution. However, the strength of this effect varies across quantiles (Yu et al., 2022 [23]). Complementary assessments stress how environmental costs are not confined to operation. Life cycle evaluation of photovoltaic panels shows significant emissions and resource use in the stages of manufacturing, operation, and disposal, including high water consumption and hazardous waste (Lakhouit et al., 2025 [24]), making the overall balance of solar power seem more complex than often assumed. Hence, the literature debate here is less about the direction of the effect and more about its completeness, with solar being associated with lower operational emissions but a larger footprint.

Similarly, the relationship between income and emissions remains one of the most contested questions in environmental economics. Early work on trade liberalization and air quality reported inverted U-shaped relationships for some local pollutants across income levels, which helped to launch what became known as the Environmental Kuznets Curve (EKC) (Grossman & Krueger, 1991 [25]). More recent panel work at the city level adds sectoral nuance to how the composition of growth affects emissions. For 96 metropolitan areas, industrial growth is associated with higher concentrations of fine and coarse particulate matter, sulfur dioxide, and nitrogen dioxide. In contrast, growth in public administration and health services is associated with reductions (Leffel et al., 2021 [26]). Extensive surveys point out the limitations of the EKC. Some studies stress that only some pollutants display such a pattern and that turning points vary widely across contexts (see, for example, Dinda, 2004 [27]). Other critiques argue that many published results rest on weak econometrics and that once time dynamics and breaks are accounted for, emissions tend to rise monotonically with income, with reductions linked to policy and technological change rather than income per se (Stern, 2004 [28]). Long-horizon evidence complicates the picture further. For advanced economies, an inverted U emerges at the panel level only after the 1950s, with parametric models providing mixed results but with semi-parametric approaches recovering inverted U-shapes for most countries (Magazzino et al., 2023 [29]). A century-long study on Swedish sulfur emissions, instead, identifies distinct structural regimes, explicitly arguing that no stable curve should be presumed (Johansson & Kriström, 2007) [30]. Time lags also matter. For instance, in OECD countries, energy innovation seems to reduce emissions only after sufficient time has passed, which implies that public research and development policies must be sustained for their effects to materialize (Álvarez-Herránz et al., 2017 [31]). The debate over the EKC is thus to some extent still unresolved, with some studies recovering inverted U-shapes under particular conditions, while others stress methodological fragility and heterogeneity, reinforcing the case for methods that allow for nonlinear and interactive dynamics, which appear increasingly necessary given the diversity of empirical results.

Fostering the EKC debate, several studies highlight the role of trade, technology, and structural complexity. For instance, in the G7 countries, public investment in renewable research and development, broader technological innovation, and trade globalization are found to raise renewable energy supply and lower CO_2 in the long run (Ahmed et al., 2022 [32]). In the Middle East, North Africa (MENA), and Turkey, renewable energy consumption is linked to lower emissions and fossil energy to higher emissions, and in some models, economic complexity aligns with improved environmental quality (Al-Ayouty, 2024 [33]). Theoretical contributions specify conditions under which diversification away from hydrocarbons is possible, stressing the profitability of non-rent sectors, the size of the domestic economy, and the inclusivity of institutions (Kaya et al., 2019 [34]). These works shift attention from the scale of trade to its content and from the level of growth to the structure of production and capability, a shift that now feels increasingly relevant.

In this context, the energy portfolio debate extends beyond renewable variables. Several studies on hydrogen describe multiple production routes from renewable and non-renewable sources, assessing cost and storage challenges and underlining the need for policy to reduce investment risks (e.g., Amin et al., 2022 [35]; Aravindan & Kumar, 2023 [36]). Social acceptance also undeniably conditions feasibility. A longitudinal analysis of media narratives in France shows how nuclear energy is portrayed consistently positively, while wind energy faces persistent criticism; this framing fosters perceptions of nuclear energy as more viable and secure, influencing policy and public debate (Bally & Sebi, 2025 [37]).

Finally, work on measurement and indicators shows that how we track sustainability also shapes conclusions. A systematic review of more than two hundred definitions of the circular economy finds both consolidation and fragmentation (Kirchherr et al., 2023 [38]). Similarly, a review of sustainability indicators proposes a general framework to combine basic measures into aggregate indices for renewable energy systems, stressing that indicator selection and weighting involve fundamental uncertainties (Liu, 2014 [39]).

Across these threads, two issues remain unresolved for higher-income cross-country analysis. First, many studies discuss state dependence, mixed effects, lags, and feedbacks. However, estimation often relies on linear and additive forms. At the same time, reviews caution against universal income–emissions curves and highlight heterogeneity and specification sensitivity, which together imply that driver rankings based on linear averages may be misleading when relationships bend or interact (see Stern, 2004 [28]). Second, flexible and ML learners are increasingly used for prediction. Still, most applications do not provide an everyday, model-agnostic basis to compare the relative contribution of economic scale, fossil energy use, non-CO_2 gases, and renewables across algorithms, which leaves policy audiences without a stable ordering of levers, making comparative analysis particularly valuable. We aim to address these gaps by assembling a panel of high- and upper-middle-income countries across ten years and by computing test-set permutation importance to compare drivers on the same footing using GLM together with ML techniques to see which factors most consistently explain emissions across models.

3. Conceptual Framework

We treat national carbon emission outcomes as the product of interacting scale, fuel-mix, agrifood, and openness channels, with relationships that may often bend or interact rather than be added linearly. In line with this framework, Liu et al. (2022) [40] examine linear and nonlinear effects in Latin American countries and report a U-shaped association between natural resources and CO_2. CO_2 emission (CO2E) sits firmly inside energy and production systems (e.g., Yoro and Daramola, 2020 [41]), so the mapping from drivers to CO2E should be considered structural rather than purely statistical.

The fuel mix is, by design, our proximate lever; we include fossil fuel energy consumption (FFEC) to capture carbon intensity on the supply side. Specifically, FFEC produces CO_2, which accounts for about a quarter of greenhouse gases, and the renewable-and-non-renewable energy mix appears to show a quadratic effect on climate vulnerability in G7 countries, with a minimum renewable threshold needed before vulnerability falls (Dai et al., 2022 [42]). Signals of clean-energy penetration are largely policy-responsive, which motivates a distinct role for renewable energy consumption (REC) in the model. Precisely, Chang et al. (2025) [43] find threshold effects of renewable energy investment in a dynamic panel: below a critical level of investment, efficiency gains may not cut fossil use due to rebound, but above that threshold, efficiency contributes more to reducing FFEC, with more potent effects in developed countries and in systems with carbon trading. Given the fact that the electricity sector can be quite distinct from total energy use, we also include in our framework renewable electricity output (REO). In this regard, Galvin et al. (2021) [44] develop a framework to study rebound effects that arise from increases in REO, including sector coupling and the use of green hydrogen, and call for a transparent way to quantify such rebound in electricity systems. Portfolio evidence further seem to support modeling FFEC separately from REC and REO. For example, in five emerging Asian economies, long-run estimates show that REC reduces climate change, while non-renewable energy consumption increases it (Wang et al., 2023 [45]). However, financing conditions can sometimes tilt these channels. In China, expansions in green bonds promote renewable energy investment and reduce environmental pollution; contractions do the opposite (Feng et al., 2023 [46]).

Economic scale provides the baseline against which all other drivers operate, so we include GDP to capture size effects that may occasionally change sign across development stages. Grossman and Krueger (1991) [25] report that for two local pollutants, concentrations rise with per capita income at low income levels and fall at higher levels, framing scale, composition, and technique effects as key channels. Sectoral panel evidence for Canada confirms this pattern, with the greenhouse gas-and-growth relationship being nonlinear in a way consistent with the EKC (Hamit-Haggar, 2012 [47]).

Lastly, agrifood systems add production and land-use channels that standard macro-energy models may miss, so we include the food production index (FOOD) as a proxy for pressure on land and inputs with context-dependent consequences for emissions. Food processing accounts for about 40 percent of energy use in the global agrifood system, and livestock uses about 60 percent of energy inputs while delivering less than 20 percent of calories (Rasul et al., 2024 [48]). In this framework, near-term warming from methane is significant, so we add methane emission (ME) to our baseline. Recent studies report a positive correlation between ME concentrations and rising heatwave temperatures, further recommending waste-to-energy as a policy option (see Toha et al., 2025 [49]). Spatial allocation and institutional coordination clearly shape agricultural externalities from fertilizer and fossil inputs, which supports modeling FOOD and ME alongside energy variables rather than treating them as residuals. In particular, inefficient nitrogen, phosphorus, and potassium use degrades ecosystems and food quality, pointing out the need to align environmental sustainability with food security (Pandian et al., 2024 [50]). Sectoral evidence on conservation and sustainable practices also indicates that agrifood impacts can be moderated by policy. An analysis of Spanish agrifood firms highlights internal factors affecting eco-innovation, including training and green relationships (Chaparro-Banegas et al., 2024 [51]).

External openness transmits technology, finance, and demand, so we include trade over GDP (TRADE) as a scale-composition shifter whose sign can often vary across the outcome distribution and policy regimes. For instance, for emerging economies, panel quantile regressions show how globalization and renewable electricity generation curb CO_2 at all quantiles, pointing to the distribution-sensitive effects of openness and the electricity mix (Qin et al., 2021 [52]). Productive capabilities can also complicate this channel. Studies find that higher economic complexity generally may impede both energy efficiency and the transition to renewable energy (Adekoya et al., 2023 [53]), justifying flexible and nonlinear forms in specifications that make TRADE interact with GDP and with energy variables.

These considerations lead to the baseline representation of national emissions, where i represents countries and t indexes years:

$$CO2E_{it} = \alpha + \beta_1(FFEC)_{it} + \beta_2(ME)_{it} + \beta_3(REO)_{it} + \beta_4(REC)_{it} + \beta_5(GDP)_{it} + \beta_6(TRADE)_{it} + \beta_7(FOOD)_{it} + \in_{it} \quad (1)$$

where GDP is treated as the scale anchor with room for bends, FFEC as the primary emission driver, methane as a high-impact non-CO_2 control, REO and REC as distinct clean-energy channels, TRADE as an openness shifter with an ambiguous sign, and FOOD as a land-use intensity proxy whose effect likely depends on practices and inputs.

It is worth clarifying that our measure of REO follows the definition used by the World Bank's World Development Indicators, which is based on data from the International Energy Agency (IEA) Energy Statistics Data Browser. The IEA definition explicitly excludes hydropower and nuclear energy while focusing on electricity generated from geothermal, solar, tidal, wind, biomass, and biofuel sources. This distinction is important because both hydropower and nuclear energy involve quite different environmental trade-offs, cost structures, and scalability constraints compared to non-hydro-renewables such as solar and wind. We adopt this definition mainly to ensure consistency with international energy statistics and to avoid the aggregation bias that might result from grouping hydropower or nuclear energy with variable renewable technologies.

4. Data and Empirical Strategy

This study uses a panel of 80 high- and upper-middle-income countries observed annually from 2011 to 2020. Countries are identified using the World Bank income classification (available at https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups. Last accessed on 18 May 2025), and all time series are drawn from the World Development Indicators (WDIs) (available at https://databank.worldbank.org/source/world-development-indicators; last accessed in May–June 2025) and the Environmental, Social, and Governance (ESG) (available at: https://databank.worldbank.org/source/environment-social-and-governance; last accessed in May 2025) databases (see Table A1 for details). Figure 1 shows the broad geographic coverage. Countries' coverage spans the Americas, Europe, East Asia and the Pacific, and selected parts of South Asia and Southern Africa; significant gaps still remain in low-income Sub-Saharan Africa and among small island economies.

The resulting mix is largely suitable for our research question as it combines fossil exporters and importers, hydro-rich and thermal-heavy systems, and a range of trade openness and policy regimes, which should increase the identification of variation. It is important to note that although our empirical focus is on CO_2 and methane, this should not be taken to imply that other greenhouse gases are irrelevant. The choice reflects data availability and sectoral consistency rather than a deliberate conceptual exclusion. CO_2 and methane together account for most anthropogenic radiative forcing and are directly linked to the energy and agrifood channels central to our framework. Other gases, such as nitrous oxide and fluorinated compounds, also matter but tend to originate from more specialized industrial or chemical processes that are difficult to harmonize in cross-country panels.

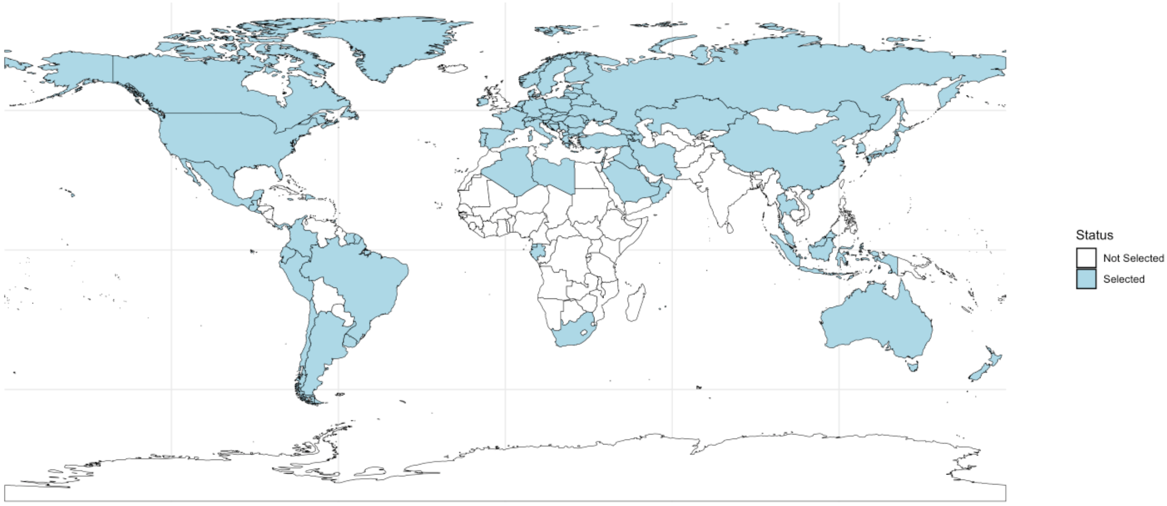

Figure 1. Political map of selected countries.

Figure 2 reports the empirical distributions for the selected variables.

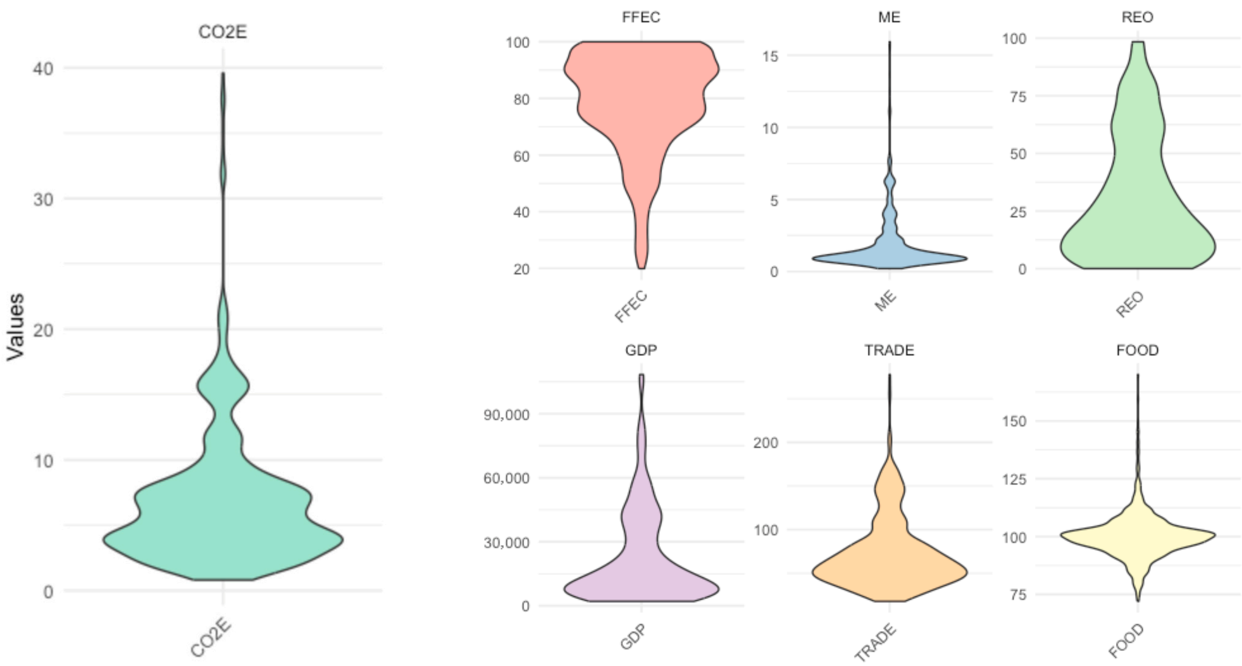

Figure 2. Violin plots for selected variables.

The dependent variable, CO2E, is notably right-skewed, with a dense mass at lower per capita emissions and a long upper tail reaching above 30–40 metric tons. The shape seems to hint at multiple regimes rather than a single location shift. While FFEC sits high for most observations, clustering near 80–100 percent and tapering toward lower reliance, the negative skew is consistent with heavy dependence on coal, oil, and gas in many country-years but still leaves room for diversified mixes. In contrast, ME is tightly bunched near the origin with occasional spikes, a pattern that likely reflects heterogeneity in livestock intensity, waste management, and rice cultivation. REO exhibits wide dispersion, with probability mass in both lower and mid-to-high shares, suggesting quite heterogeneous stages of renewable deployment across the panel. Surprisingly, GDP is highly skewed, reflecting the dispersion in income per capita within this income group; several observations lie far in the right tail. TRADE displays a long right tail above 150–200 percent of GDP, which may be typical of small open economies and entrepôt hubs. Unlike other variables, FOOD is centered close to its index base with thinner tails. Still, the violin is not entirely symmetric, indicating that a subset of country-years experienced unusually high or low food production intensity.

The summary statistics in Table A2 in the Appendix A align with these described shapes. Average CO2E is 6.95 metric tons per capita with substantial dispersion and pronounced right-tail risk, as might be expected. Mean FFEC is 78.74 percent, and its negative skew confirms that many observations sit near very high fossil shares. ME averages 1.92 t CO2E per capita but is highly leptokurtic, which cautions against Gaussian assumptions. Meanwhile, REO averages 30.43 percent with a widespread, again consistent with uneven transition progress. Finally, average GDP per capita is 22.51 thousand constant 2015 USD, and TRADE averages 73.17 percent of GDP, while FOOD centers near 100 with occasional extremes.

Figure 3 summarizes pairwise dependence with three facts clearly standing out.

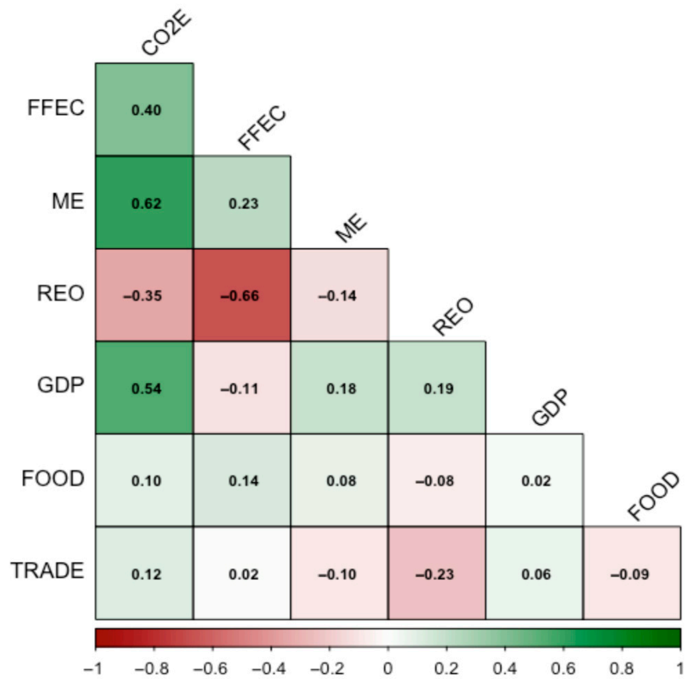

Figure 3. Correlation matrix.

First, CO2E is positively associated with ME (0.62), GDP (0.54), and FFEC (0.40), which is consistent with emissions rising with methane-intensive activity, economic scale, and fossil reliance. Second, REO is negatively related to CO2E (−0.35) and strongly negatively associated with FFEC (−0.66), a pattern suggestive of substitution on the supply side. Third, most other correlations are small in absolute value, including those involving TRADE and FOOD, which reduces concerns about broad confounding among the remaining regressors. Taken together, these descriptive statistics highlight considerable substantial heterogeneity, heavy tails, and at least one salient trade-off in the energy mix.

Our empirical strategy is therefore designed around these features. We estimate three learners that span rather different functional forms: a GLM with a Gaussian identity for transparency, an SVM with a radial kernel to accommodate curvature, and an RF to capture interactions. In the robustness checks, we also complemented ranking with a DML design that partialed out ME, REO, TRADE, and FOOD before estimating the incremental effect of FFEC and GDP; this orthogonalization limits functional-form bias and helps interpret the fuel-mix margin.

We estimate a GLM with Gaussian identity as the baseline, fitted under five-fold cross-validation repeated twice. This classical method identifies and estimates dispersion parameters for correlated data by extending univariate generalized linear results to higher dimensions (see the seminal work of McCullagh and Nelder, 1989 [54]). Thus, we treat the GLM as a transparent benchmark for signs and for average magnitudes. The broader econometric literature covering cross-section and panel methods (e.g., Wooldridge, 2010 [55]), including models with heteroskedasticity and clustering, supports this design choice given our data structures.

To allow smooth nonlinearity without prespecifying a functional form, we estimate an SVM model with a radial basis kernel. SVM is introduced as a regression technique and, in comparisons, is expected to have advantages in high-dimensional settings given the fact that its optimization does not depend on the dimensionality of the input space (Drucker et al., 1997 [56]; Smola & Schölkopf, 2004 [57]). Learning theory provides principles for controlling generalization and for empirical risk minimization that underpin support vector methods (Vapnik, 2000 [58]).

We also estimate an RF with 500 trees. RF combines several trees built on bootstrap samples with random feature selection at splits, and their generalization error converges as the number of trees grows, with internal estimates available for error, variable importance, and performance that compares favorably to boosting while being robust to noise (Breiman, 2001 [59]). Standard RF variable importance can be misleading when predictors differ in measurement scale or number of categories, and an alternative implementation with unbiased selection and subsampling without replacement has been proposed to address this problem (Strobl et al., 2007 [60]). Hence, for model comparison, we computed permutation-based variable importance on a held-out test set. This approach, rooted in the permutation framework of Breiman (2001) [59] and further generalized by Fisher et al. (2019) [61], measures the average loss in predictive accuracy when the information carried by a specific feature is destroyed. It thus provides an interpretable and model-agnostic indicator of each variable's contribution to out-of-sample performance.

The empirical analysis was conducted in R (version 4.4.1) using RStudio (version 2024.12.1+563), as in standard practice. Models were estimated through the *caret*, *e1071*, and *ranger* packages. The GLM with Gaussian identity serves as the linear benchmark, providing interpretable coefficients for direction and relative magnitude. The SVM with a radial kernel captures nonlinear and interaction effects without prespecifying functional forms, while the RF aggregates multiple decorrelated trees to approximate complex, data-driven relationships. Out-of-sample performance was assessed using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Squared Error (MSE), and the coefficient of determination (R^2), among others, and permutation-based variable importance was computed to quantify the contribution of each predictor to the test set.

As a robustness test, we estimated the partial effects of FFEC and GDP while allowing the other covariates to enter flexibly and interact. Double or debiased ML uses Neyman-orthogonal scores and cross-fitting to remove regularization bias from high-dimensional nuisance estimation and delivers approximately unbiased, near-normal estimators suitable for valid confidence statements in settings like partially linear regression and treatment effect estimation (see the seminal work of Chernozhukov et al., 2018 [62]).

To summarize our empirical strategy, the GLM anchors interpretability, the SVR supplies smooth curvature with margin-based control, and the RF framework naturally lets interactions and threshold effects surface. At the same time, in the robustness checks, DML targets the fuel-mix margin under flexible conditioning. It is important to note that the combination is chosen not to chase in-sample fit but rather to keep a clear, comparable ordering of drivers and a treatment-style estimate for FFEC that remains credible when linear additivity is unlikely.

5. Main Results and Discussion

This section presents the main estimates from three complementary learners: a GLM with Gaussian identity, an SVM with a radial kernel, and an RF. (All variables were standardized first; then we created an 80/20 outcome-stratified train/test split. Model fitting and tuning were performed with caret (standardization, splitting, and repeated five-fold cross-validation). The linear benchmark is a fitted Gaussian GLM. The nonparametric learners are (i) a radial-kernel SVM (tuned over cost and gamma) and (ii) an RF from ranger (500 trees; grid over *mtry* and node size).) The GLM provides a straightforward linear reference point, where estimated coefficients directly indicate how each predictor relates to CO_2 emissions. Moving beyond linearity, the SVM with a radial kernel reveals interactions and curved relationships that emerge naturally from the data. At the same time, the RF models the joint structure of variables through ensemble learning. Across models, we evaluate predictive accuracy and internal consistency by jointly comparing RMSE, MAE, MSE, and R^2. Variable importance, derived from test-set permutation methods, highlights how much each factor contributes to prediction accuracy, providing a comparable and interpretable measure of influence across modeling strategies.

Table 1 reports the individual results obtained from each learner, presenting estimated coefficients for the GLM and test-set variable importance for the SVM and RF.

Table 1. Summary of individual model results.

		GLM		SVM	RF
Variable/Feature	**Estimate**	**Std. Error**	***p*-Value**	**Importance**	**Importance**
FFEC	0.202	0.029	0.000	0.713	0.285
ME	0.458	0.021	0.000	0.474	0.269
REO	−0.240	0.030	0.000	0.436	0.196
GDP	0.533	0.022	0.000	0.649	0.540
FOOD	0.001	0.021	0.944	0.094	0.009
TRADE	0.085	0.023	0.000	0.347	0.131
(Intercept)	0.013	0.021	0.539	-	-

In the GLM, GDP, ME and FFEC emerge with the most significant positive coefficients, indicating that economic scale and fossil energy use remain the dominant contributors to emissions once other factors are controlled for. The SVM and RF confirm this pattern. However, the ranking of variable importance slightly differs: GDP consistently retains the highest influence, followed by FFEC and ME, while REO exerts an adverse effect across models. TRADE appears relevant but secondary, and FOOD plays a minimal role. The close alignment in variable ordering across learners suggests, as a matter of fact, that the main relationships are not model-specific artifacts. Linear, kernel-based, and ensemble estimators converge in highlighting scale and fossil dependence as primary sources of variation, with renewable energy and trade channels operating mainly at the margin.

Figure 4 reports model-agnostic permutation importance computed on the test window and normalized to sum to one within the learner.

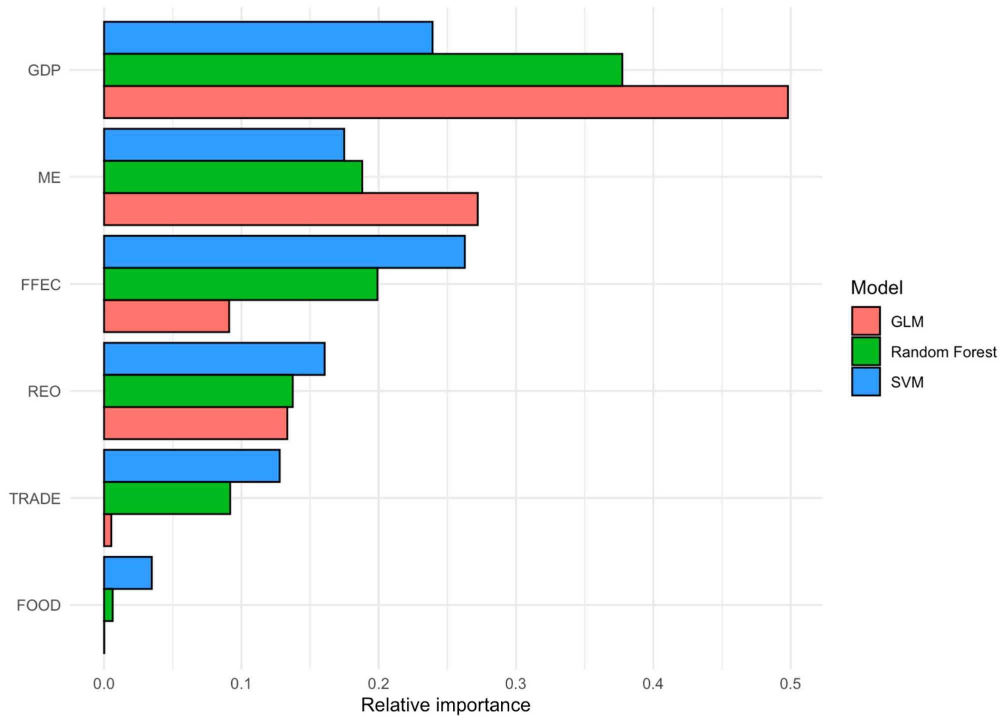

Figure 4. Model importance scores.

GDP represents the top contributor in every model, but the magnitude is largest in the GLM framework, followed by RF and then SVM. Like that in Table 1, the pattern is consistent overall across individual models, with strong scale effects even after conditioning on the energy mix, methane, trade, and food. Contrasts across learners help explain these allocations. The GLM concentrates explanatory weight on GDP and ME and assigns only modest roles to FFEC, REO, and TRADE. By comparison, both the SVM and RF reallocate some weight from GDP and ME to FFEC, REO, and TRADE. Two forces likely matter. First, the strong association between REO and FFEC in Figure 3 makes linear attribution sensitive to multicollinearity, whereas the flexible learners can still exploit their opposing signals out of sample. Second, trade appears to act through interactions with scale and the energy mix rather than through a large marginal slope, which is why its importance rises precisely in models that allow curvature and splits.

Figure 5 summarizes out-of-sample performance on the same test data.

The RF model attains the best scores on every error metric and the highest R^2; the SVM comes second, and the GLM lags. This ordering lines up with the data features we documented earlier in the previous section: heavy right tails, multimodality in several regressors, and plausible interactions between scale and the fuel mix. Kernel smoothing and tree ensembles can absorb structures that would otherwise be missed without prespecifying functional forms, while single linear coefficients, even on standardized inputs, are less forgiving in this sense.

Figure 6 presents two complementary diagnostics that help unpack the internal structure of the RF model used in the primary analysis (Figure 6 was generated using the *randomForestExplainer* package. The visualization is based on the same RF model trained in the primary analysis. The left panel displays the mean minimal depth of each predictor averaged across all trees, while the right panel plots the frequency with which each variable appears as a root split against its average depth).

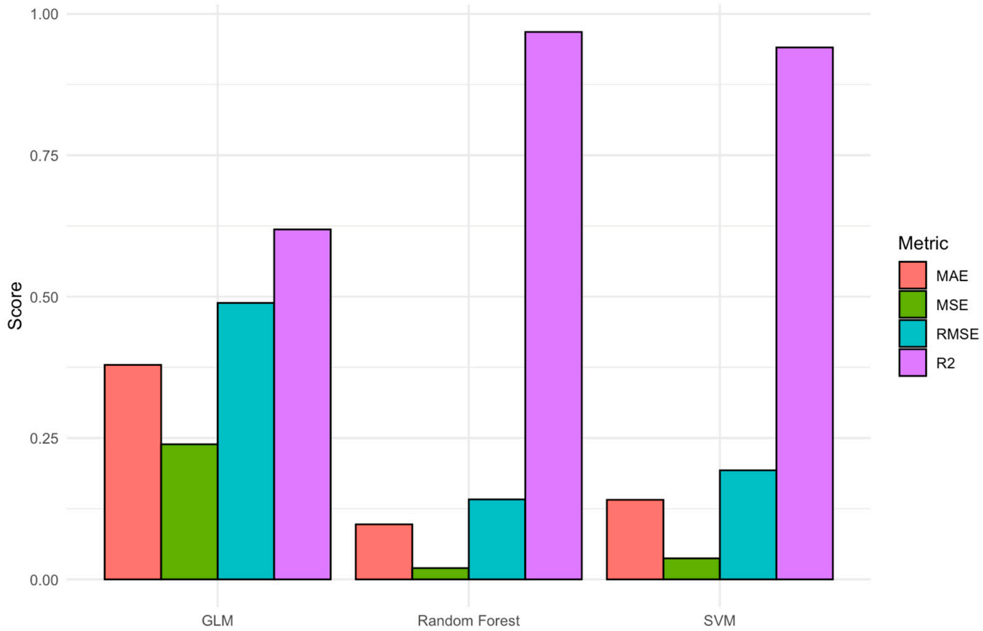

Figure 5. Model performance metrics.

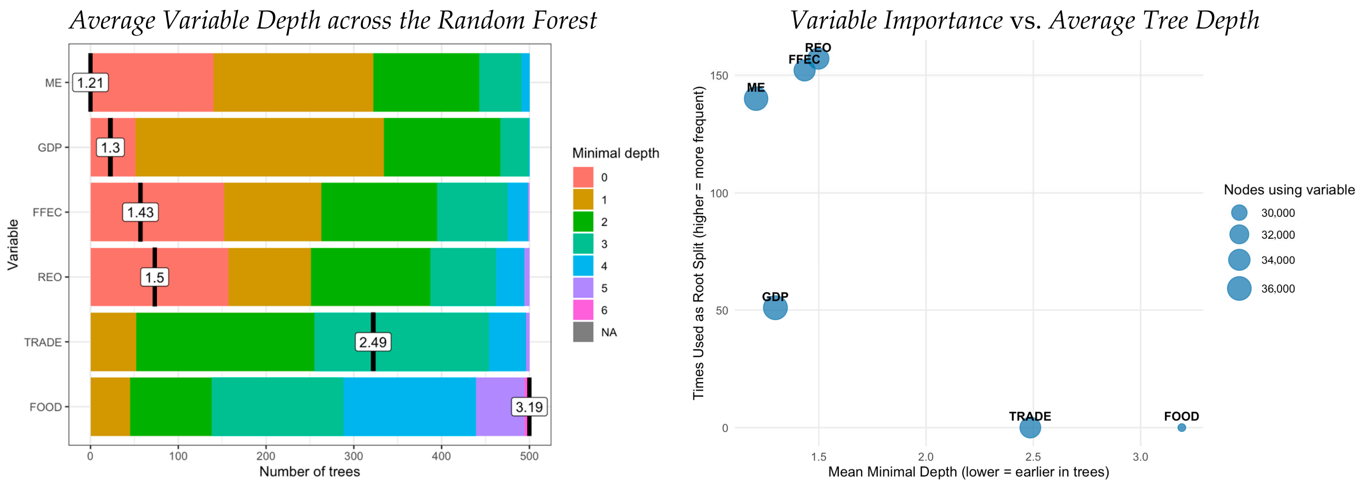

Figure 6. Variable depth and structural importance in the RF.

In the left panel, GDP, FFEC, and ME occupy the top ranks, appearing closer to the root across a large share of trees, while TRADE and FOOD enter later and more sporadically. The right panel links these depth patterns with the frequency of root-level usage. Again, GDP, FFEC, and ME emerge as structurally dominant, with REO showing moderate importance but higher depth, meaning that its predictive role is conditional on other factors rather than direct.

Re-reading Figure 3 with Figures 4 and 5 in mind helps to reconcile simple correlations with model-based contributions. Pairwise, CO2E rises with GDP, FFEC, and especially ME and falls with REO. REO and FFEC are strongly negatively related. The importance rankings respect these broad signals; GDP and ME are high, and FFEC and REO both matter, but they do so after conditioning on all covariates and evaluating out of sample. TRADE's modest pairwise link but nontrivial importance in flexible learners probably suggests that its effect comes through interactions rather than a significant marginal correlation with CO2E. FOOD's near-zero importance aligns with its weak pairwise associations and limited dispersion.

Two takeaways clearly follow. Economic scale dominates national carbon emission outcomes in high- and upper-middle-income countries, with methane and the fuel mix forming a consistent second tier. Electricity-sector renewables somewhat help at the margin after conditioning on GDP and FFEC, but they do not replace the central roles of scale and fossil intensity in the short run.

Our main patterns align with, but also qualify, several strands in the literature. The dominant role of GDP is consistent with studies that relate environmental outcomes to economic scale and document contexts where pollution rises at lower income levels and declines at higher income levels, while also noting that such turning points are not universal and that evidence varies by pollutant and setting (see, for example, Grossman and Krueger, 1991 [25]; Dinda, 2004 [27]; Stern, 2004 [28]). The positive contribution of FFEC is in line with panel evidence that non-renewable energy use generally raises CO2E in the long run and with sectoral results that link higher energy consumption to higher greenhouse gases (Al-Ayouty, 2024 [33]; Hamit-Haggar, 2012 [47]). The sizable role of ME is consistent with sector studies that identify landfill emissions as significant and correlated with heatwave intensity (see Toha et al., 2025 [49]), as well as with work highlighting how input practices in agrifood systems shape environmental pressures (Pandian et al., 2024 [50]). The moderate, not dominant, contribution of REO fits cross-country and review evidence that renewable deployment can reduce emissions, but that effects depend on enabling conditions and may vary by country and quantile (Qin et al., 2021 [52]). In addition, the relatively small pairwise role of trade but its nontrivial importance in flexible learners further aligns with work indicating that openness to trade can support mitigation mainly through channels of technology and policy (Ahmed et al., 2022 [32]). Finally, the limited explanatory power of FOOD is consistent with evidence that environmental impacts in agrifood systems often depend more on energy use, processing, and input management than on output volume alone (Rasul et al., 2024 [48]), which an annual aggregate index may not capture.

In conclusion, what is new in our findings is not the existence of these channels but rather the way we rank them on common ground and test their stability out of sample. Most cross-country studies lean on linear additivity or a single flexible learner; we provide a model-agnostic ranking that instead relies on the same test-set permutation logic for the GLM, SVM, and RF. In practical terms, this setup simply shows that economic scale remains the first-order driver in high- and upper-middle income samples, while methane and the fuel mix form a robust second tier, renewable electricity contributes but does not dominate without supportive conditions, and trade matters mainly through interactions that linear models tend to miss.

6. Robustness Checks

The robustness of the main patterns was therefore checked with DML-built RF nuisances. The choice for this setting was fairly simple: in the headline comparisons, RF delivered the best out-of-sample fit, and DML let us estimate treatment-style effects while still flexibly partialing out the remaining covariates through cross-fitted machine learners. Concretely, we treated FFEC and GDP as the treatments of interest, kept ME, REO, TRADE, and FOOD as controls, centered and scaled all variables, and used an outcome-stratified 80–20 train–test split. RF with 500 trees (the same setting as in the main empirical testing) was used for the nuisance steps, with five-fold cross-fitting to help curb overfitting. After fitting, we evaluated (1) how well the nuisance learners predict and (2) how much each treatment and control contribute to the final structural predictor using essentially the same permutation-importance routine as the main models in Figure 4.

Figure 7 reports the predictive diagnostics for the nuisance learners.

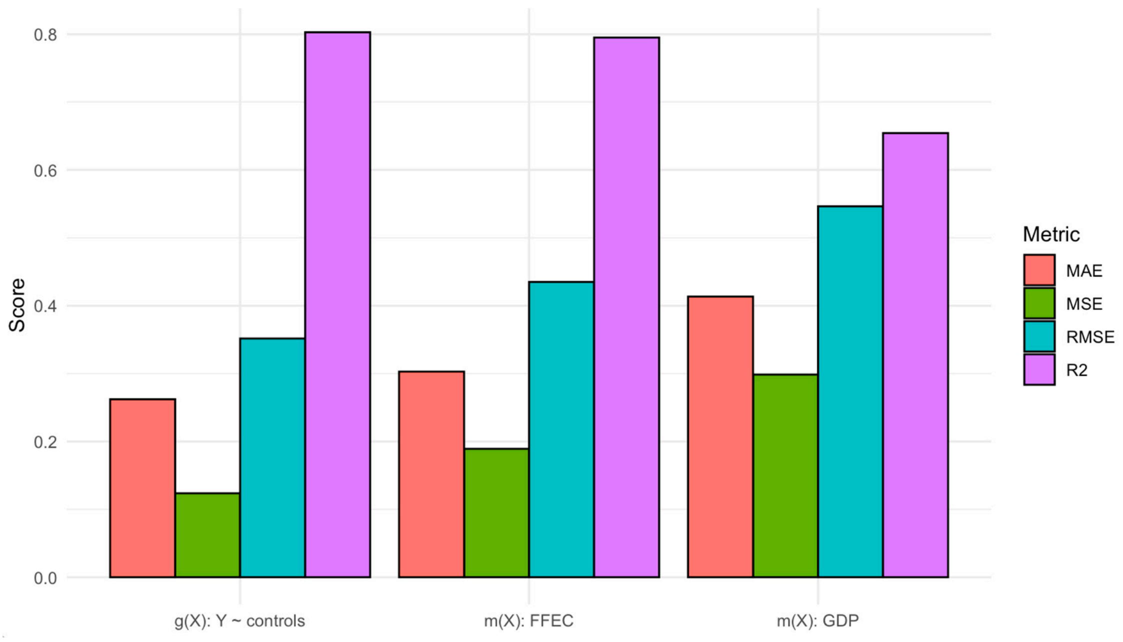

Figure 7. DML performance metrics.

The control-only model that predicts CO2E from ME, REO, TRADE, and FOOD performs quite strongly, with high explanatory power and relatively low errors. The model that predicts the fossil share from those same controls is also very accurate, indicating that the control set carries real information about the energy mix. By contrast, the GDP-from-controls model is somewhat weaker, suggesting that orthogonalization is especially effective for the fossil share, as a large share of its predictable variation is removed before we estimate its partial effect on emissions. It also hints that income contains considerable idiosyncratic variation that the controls do not explain well, which is consistent with the larger treatment effect we estimate for GDP.

Figure 8 turns from fit to composition.

Within the treatment block, income, proxied by GDP per capita, accounts for the larger share of the structural predictor, and FFEC takes the remainder. Within the control block, ME leads, REO follows, TRADE is minor but not clearly negligible, and FOOD is essentially absent. These bars are normalized within the group, so they rank importance relative to peers rather than in absolute terms. Even so, the ordering echoes what we saw in the predictive learners in the previous section: economic scale first, then methane and the fuel mix, and then renewables and trade, with food production close to zero.

Table 2 summarizes the DML treatment coefficients in standardized units. Both are positive and highly significant.

Table 2. DML results.

Treatment	Estimate	Std. Error	t Value	*p*-Value	CI Lower	CI Upper	Significance
FFEC	0.235	0.029	7.961	<0.0001	0.177	0.292	***
GDP	0.451	0.029	15.377	<0.0001	0.394	0.509	***

Note: *** $p < 0.001$.

Specifically, it appears that a 1-standard-deviation increase in GDP per capita is associated with about a 0.45-standard-deviation increase in CO2E after flexibly conditioning on the controls. Meanwhile, a 1-standard-deviation increase in FFEC is associated with about a 0.24-standard-deviation increase in emissions under the same conditioning. Standard errors are tight, *p*-values are below 0.0001, and the confidence intervals are narrow. Reading this together with Figures 6 and 7, these results show how economic scale represents the dominant margin in this country's sample. However, the fuel mix has a distinct and still economically meaningful incremental effect that does not decrease once methane, renewables, trade, and food-system activity are taken into account.

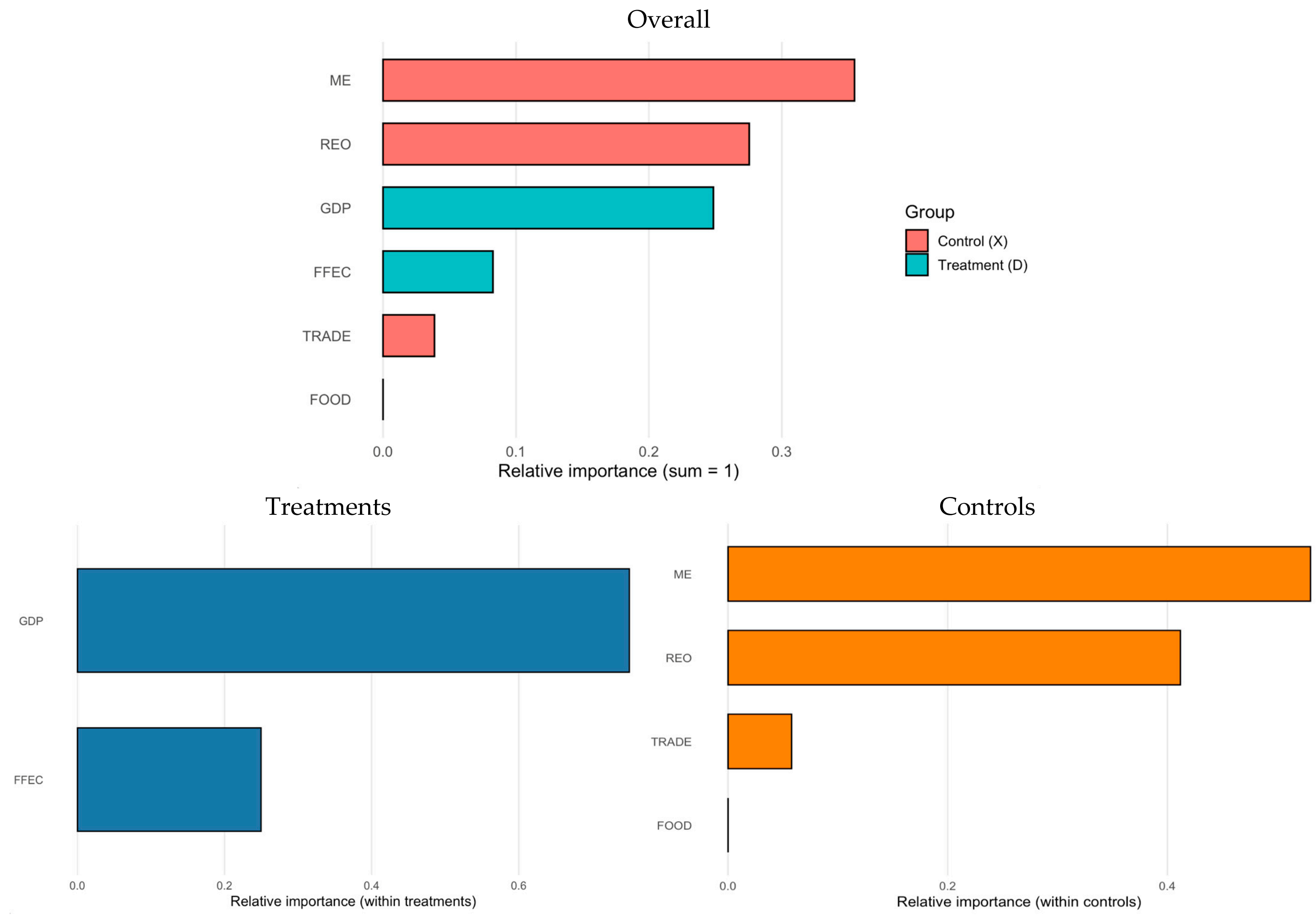

Figure 8. DML relative importance scores.

7. Concluding Remarks

CO_2 linked to fossil fuel use remains by far the primary source of global warming and environmental degradation, keeping mitigation a central concern for sustainable development. The relationship between emissions, economic growth, and energy intensity has been explored across a range of regional and methodological contexts. Prior studies have shown how emissions directly shape both environmental and social outcomes: in African economies, higher CO_2 and urbanization reduce life expectancy despite improvements in income and health spending (Beşer et al., 2025 [63]); and in Russia, the connection between energy use, GDP, and emissions reinforces itself, though efficiency gains can still be achieved without compromising growth (Magazzino et al., 2023 [64]). In MENA countries, patterns appear more heterogeneous, as GDP and energy consumption tend to increase emissions while trade and urbanization mitigate them to varying degrees (Magazzino & Cerulli, 2019 [65]). Meanwhile, long-run evidence also points to only partial convergence in per capita emissions, reflecting structural asymmetries that persist across countries despite regional integration (Magazzino, 2019 [66]).

Building on this body of work, our analysis broadens the focus to a larger set of high- and upper-middle-income economies and applies a combination of econometric and ML methods to uncover how GDP, FFEC, ME, and REO interact in shaping CO_2 outcomes. In the GLM framework, GDP emerges as the dominant driver of emissions, followed by FFEC and ME. When moving to more flexible learners such as SVM and RF, this general hierarchy largely remains, although REO and TRADE gain relevance once nonlinearities and threshold effects are accounted for. This shift might suggest that while scale effects continue to dominate, the contributions of the fuel mix and methane become more pronounced when models allow for curvature and interaction among drivers. The DML estimates further help refine this picture, confirming that GDP and FFEC retain positive and statistically significant effects after conditioning on all other covariates. REO, defined in line with the World Bank's WDIs based on IEA data, shows a modest positive contribution, likely constrained by financial, infrastructural, and regulatory factors that limit renewable penetration across high- and upper-middle-income economies.

The analysis adds value to the current literature in at least three main ways. First, it establishes a comparable ordering of emission drivers across linear and nonlinear learners. Second, it isolates treatment-style partial effects for GDP and FFEC through cross-fitted DML, distinguishing association from the incremental contribution of each treatment under flexible controls. Third, it applies this unified strategy to a harmonized panel of 80 high- and upper-middle-income countries from 2011 to 2020, representing a policy-relevant portion of the global economy.

The policy implications are fairly straightforward. If GDP remains the primary force behind emissions, mitigation strategies that rely only on compositional change are unlikely to deliver sufficient reductions. Measures that decouple scale from emissions, through efficiency, structural transformation, and pricing that internalizes externalities, appear especially indispensable. The persistent FFEC effect points to the urgency of fossil substitution both in power generation and end-use sectors. Similarly, targeted ME controls can yield notable near-term gains where agriculture and waste management play central roles. REO contributes at the margin but, within this income group, does not yet offset the influence of GDP and FFEC without supportive financial, infrastructural, and policy environments.

Several limitations must be noted. The variables used are broad aggregates and cannot fully reflect policy intensity, energy trade, or grid constraints. Annual frequency may also obscure short-run and lagged effects. The sample covers only 2011–2020 and includes high- and upper-middle-income economies, limiting extrapolation to lower-income settings or earlier decades. Although DML mitigates functional-form misspecification, it does not visibly eliminate endogeneity from unobserved shocks that jointly affect GDP and emissions. Coefficients are standardized, which aids comparison but abstracts from level effects and sectoral details. The REO definition, based on WDI and IEA data, excludes hydro- and nuclear sources. While this ensures consistency across countries, it inevitably narrows renewable coverage and is acknowledged as a methodological limitation.

Future research should further extend this framework by integrating explicit policy and price variables (carbon pricing, subsidies, grid interconnection, green finance), adopting dynamic or panel–DML approaches to capture lagged responses, and exploring heterogeneity across regions and income groups via interaction-rich or quantile designs. Further work could usefully incorporate spatial spillovers and electricity trade or link national data to sectoral emissions to identify which segments of the economy drive the GDP and FFEC margins. Finally, pairing DML with credible instruments for exogenous shocks to the fuel mix, income structure, or ME management would greatly strengthen causal interpretation.

Author Contributions: Conceptualization, C.M., U.M., E.C., A.C. and T.G.; Methodology, C.M. and T.G.; Software, T.G.; Validation, C.M.; Formal analysis, C.M. and T.G.; Investigation, C.M. and T.G.; Resources, C.M.; Data curation, T.G.; Writing—original draft, U.M., E.C., A.C. and T.G.; Writing—

review & editing, C.M.; Visualization, C.M.; Supervision, C.M.; Project administration, C.M.; Funding acquisition, C.M. All authors have read and agreed to the published version of the manuscript.

Funding: The authors received no specific funding for this work.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the corresponding author upon reasonable request.

Acknowledgments: The views expressed in this paper are those of the authors and do not necessarily represent those of the institutions with which the authors are affiliated. All errors and opinions are ours. All standard disclaimers apply.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Variables' description.

Variable	Definition	Unit	Source
CO2E	Measures total CO_2 released from energy use and industrial processes, serving as an indicator of national contributions to global emissions.	Metric tons per capita	ESG Data, World Bank
FFEC	Share of energy derived from coal, oil, and natural gas, indicating reliance on fossil fuels.	% of total energy consumption	WDI, World Bank
ME	Estimates methane output in CO_2-equivalent terms, reflecting its contribution to climate change.	Metric tons CO_2 equivalent per capita	ESG Data, World Bank
REO	Proportion of electricity generated from renewable sources relative to total electricity output.	% of total electricity generation	ESG Data, World Bank
GDP	Total economic output per person, adjusted for inflation to constant 2015 prices.	Constant 2015 USD	WDI, World Bank
TRADE	Measured as the sum of exports and imports of goods and services, expressed as a share of GDP.	% of GDP	WDI, World Bank
FOOD	Measures total edible crop and livestock production, adjusted for seed and feed use, expressed as an index relative to the 2014–2016 base period.	Index (2014–2016 = 100)	WDI, World Bank

Table A2. Descriptive statistics.

Variable	Mean	Stand. Dev.	Min.	Max.	Skewness	Kurtosis
CO2E	6.953	5.361	0.811	39.582	2.378	8.494
FFEC	78.743	17.659	19.910	100.000	−1.015	0.690
ME	1.920	2.049	0.204	15.928	2.999	12.008
REO	30.427	26.405	0.002	98.430	0.790	−0.482
GDP	22,511.541	21,700.281	2094.474	108,351.452	1.543	2.172
TRADE	73.171	40.638	17.944	277.521	1.417	2.310
FOOD	100.181	9.864	72.010	170.030	1.486	7.489

References

1. Murphy, R. What is undermining climate change mitigation? How fossil-fuelled practices challenge low-carbon transitions. *Energy Res. Soc. Sci.* **2024**, *108*, 103390. [CrossRef]
2. Hou, H.; Lu, W.; Liu, B.; Hassanein, Z.; Mahmood, H.; Khalid, S. Exploring the role of fossil fuels and renewable energy in determining environmental sustainability: Evidence from OECD countries. *Sustainability* **2023**, *15*, 2048. [CrossRef]
3. Krewitt, W.; Simon, S.; Graus, W.; Teske, S.; Zervos, A.; Schafer, O. The 2-degrees C scenario: A sustainable world energy perspective. *Energy Policy* **2007**, *35*, 4969–4980. [CrossRef]

4. Rahman, M.M.; Dadon, S.H.; He, M.; Giesselmann, M.; Hasan, M.M. An overview of power system flexibility: High renewable energy penetration scenarios. *Energies* **2024**, *17*, 6393. [CrossRef]
5. Züttel, A.; Nützenadel, C.; Schlapbach, L.; Gilgen, P.W. Power plant units for CO_2 neutral energy security in Switzerland. *Front. Energy Res.* **2024**, *12*, 1336016. [CrossRef]
6. Ashraf, A.; Sagheer, M. Renewable energy capacity and technological innovations: A review of global trends and future directions. *Environ. Prog. Sustain. Energy* **2025**, e70071. [CrossRef]
7. Haloui, I.; Amzil, H.; Yang, G.; Fourati, I.; Li, Y. The impact of renewable energy use, financial development, and industrialization on CO_2 emissions in middle-income economies—A GMM-PVAR analysis. *Sustainability* **2025**, *17*, 8178. [CrossRef]
8. Olabi, A.G.; Elsaid, K.; Obaideen, K.; Abdelkareem, M.A.; Rezk, H.; Wilberforce, T.; Maghrabie, H.M.; Sayed, E.T. Renewable energy systems: Comparisons, challenges and barriers, sustainability indicators, and the contribution to UN sustainable development goals. *Int. J. Thermofluids* **2023**, *20*, 100498. [CrossRef]
9. Rutovitz, J.; Langdon, R.; Briggs, C.; Mey, F.; Dominish, E.; Nagrath, K. Updated employment factors and occupational shares for the energy transition. *Renew. Sustain. Energy Rev.* **2025**, *212*, 115339. [CrossRef]
10. Finn, O.; Brockway, P.E. Much broader than health: Surveying the diverse co-benefits of energy demand reduction in Europe. *Energy Res. Soc. Sci.* **2023**, *95*, 102890. [CrossRef]
11. Holechek, J.L.; Geli, H.M.E.; Sawalhah, M.N.; Valdez, R. A global assessment: Can renewable energy replace fossil fuels by 2050? *Sustainability* **2022**, *14*, 4792. [CrossRef]
12. Mehmood, S.; Zaman, K.; Khan, S.; Ali, Z.; Khan, H.U.R. The role of green industrial transformation in mitigating carbon emissions: Exploring the channels of technological innovation and environmental regulation. *Energy Built Environ.* **2024**, *5*, 464–479. [CrossRef]
13. Brodny, J.; Tutak, M.; Grebski, W.W. Empirical evaluation of the energy transition efficiency in the EU-27 countries over a decade—A non-obvious perspective. *Energies* **2025**, *18*, 3367. [CrossRef]
14. Yang, S.; Wang, X.; Ge, Z.; Dong, G.; Ma, M.; Han, X. Global per capita CO_2 emission trends. *Atmosphere* **2023**, *14*, 1797. [CrossRef]
15. Kumar, A.; Luthra, S.; Mangla, S.K.; Garza-Reyes, J.A.; Kazancoglu, Y. Analysing the adoption barriers of low-carbon operations: A step forward for achieving net-zero emissions. *Resour. Policy* **2023**, *80*, 103256. [CrossRef]
16. Bhatt, H.; Davawala, M.; Joshi, T.; Shah, M.; Unnarkat, A. Forecasting and mitigation of global environmental carbon dioxide emission using machine learning techniques. *Clean. Chem. Eng.* **2023**, *5*, 100095. [CrossRef]
17. Hamrani, A.; Akbarzadeh, A.; Madramootoo, C.A. Machine learning for predicting greenhouse gas emissions from agricultural soils. *Sci. Total Environ.* **2020**, *741*, 140338. [CrossRef]
18. Gattone, T.; Costantiello, A.; Leogrande, A.; Quarto, A. Modelling CO_2 emissions through econometric and machine learning approaches: The role of renewable energy and resource use in climate action. *Geol. J.* **2025**, 1–23. [CrossRef]
19. Yadav, A.; Gyamfi, B.A.; Asongu, S.A.; Behera, D.K. The role of green finance and governance effectiveness in the impact of renewable energy investment on CO_2 emissions in BRICS economies. *J. Environ. Manag.* **2024**, *358*, 120906. [CrossRef]
20. Li, C.Z.; Umair, M. Does green finance development goals affects renewable energy in China. *Renew. Energy* **2023**, *203*, 898–905. [CrossRef]
21. Bassi, A.M.; Pallaske, G.; Bridle, R.; Bajaj, K. Emission reduction via fossil fuel subsidy removal and carbon pricing, creating synergies with revenue recycling. *World* **2023**, *4*, 225–240. [CrossRef]
22. Lal, A.; You, F. Advances and challenges in energy and climate alignment of AI infrastructure expansion. *Adv. Appl. Energy* **2025**, *20*, 100243. [CrossRef]
23. Yu, J.; Tang, Y.M.; Chau, K.Y.; Nazar, R.; Ali, S.; Iqbal, W. Role of solar-based renewable energy in mitigating CO_2 emissions: Evidence from quantile-on-quantile estimation. *Renew. Energy* **2022**, *182*, 216–226. [CrossRef]
24. Lakhouit, A.; Alhathlaul, N.; El Mokhi, C.; Hachimi, H. Assessing the environmental impact of PV emissions and sustainability challenges. *Sustainability* **2025**, *17*, 2842. [CrossRef]
25. Grossman, G.M.; Krueger, A.B. Environmental impacts of the North American Free Trade Agreement. *Natl. Bur. Econ. Res. Work. Pap.* **1991**, 3914.
26. Leffel, B.; Tavasoli, N.; Liddle, B.; Henderson, K.; Kiernan, S. Metropolitan air pollution abatement and industrial growth: Global urban panel analysis of PM10, PM2.5, NO2 and SO2. *Environ. Sociol.* **2021**, *8*, 94–107. [CrossRef]
27. Dinda, S. Environmental Kuznets curve hypothesis: A survey. *Ecol. Econ.* **2004**, *49*, 431–455. [CrossRef]
28. Stern, D.I. The rise and fall of the environmental Kuznets curve. *World Dev.* **2004**, *32*, 1419–1439. [CrossRef]
29. Magazzino, C.; Gallegati, M.; Giri, F. The environmental Kuznets curve in a long-term perspective: Parametric vs semi-parametric models. *Environ. Impact Assess. Rev.* **2023**, *98*, 106973. [CrossRef]
30. Johansson, P.O.; Kriström, B. On a clear day you might see an environmental Kuznets curve. *Environ. Resour. Econ.* **2007**, *37*, 77–90. [CrossRef]
31. Álvarez-Herránz, A.; Balsalobre, D.; Cantos, J.M.; Shahbaz, M. Energy innovations–GHG emissions nexus: Fresh empirical evidence from OECD countries. *Energy Policy* **2017**, *101*, 90–100. [CrossRef]

32. Ahmed, Z.; Ahmad, M.; Murshed, M.; Shah, M.I.; Mahmood, H.; Abbas, S. How do green energy technology investments, technological innovation, and trade globalization enhance green energy supply and stimulate environmental sustainability in the G7 countries? *Gondwana Res.* **2022**, *112*, 105–115. [CrossRef]
33. Al-Ayouty, I. Economic complexity and renewable energy effects on carbon dioxide emissions: A panel data analysis of Middle East and North Africa countries. *J. Knowl. Econ.* **2024**, *15*, 12006–12025. [CrossRef]
34. Kaya, A.; Tok, E.; Koc, M.; Mezher, T.; Tsai, I.-T. Economic diversification potential in the rentier states towards a sustainable development: A theoretical model. *Sustainability* **2019**, *11*, 911. [CrossRef]
35. Amin, M.; Shah, H.H.; Fareed, A.G.; Khan, W.U.; Chung, E.; Zia, A.; Lee, C. Hydrogen production through renewable and non-renewable energy processes and their impact on climate change. *Int. J. Hydrogen Energy* **2022**, *47*, 33112–33134. [CrossRef]
36. Aravindan, M.; Kumar, G.P. Hydrogen towards sustainable transition: A review of production, economic, environmental impact, and scaling factors. *Results Eng.* **2023**, *20*, 101456. [CrossRef]
37. Bally, F.; Sebi, C. Shifting public perception: A longitudinal analysis of national and local media narratives on nuclear and wind energy in France. *Technol. Forecast. Soc. Chang.* **2025**, *215*, 124111. [CrossRef]
38. Kirchherr, J.; Yang, N.-H.N.; Schulze-Spüntrup, F.; Heerink, M.J.; Hartley, K. Conceptualizing the circular economy (revisited): An analysis of 221 definitions. *Resour. Conserv. Recycl.* **2023**, *194*, 107001. [CrossRef]
39. Liu, G. Development of a general sustainability indicator for renewable energy systems: A review. *Renew. Sustain. Energy Rev.* **2014**, *31*, 611–621. [CrossRef]
40. Liu, H.; Alharthi, M.; Atil, A.; Zafar, M.W.; Khan, I. A non-linear analysis of the impacts of natural resources and education on environmental quality: Green energy and its role in the future. *Resour. Policy* **2022**, *79*, 102940. [CrossRef]
41. Yoro, K.O.; Daramola, M.O. CO_2 emission sources, greenhouse gases, and the global warming effect. In *Advances in Zcarbon Capture*; Woodhead Publishing: Cambridge, UK, 2020; pp. 3–28.
42. Dai, H.; Mamkhezri, J.; Arshed, N.; Javaid, A.; Salem, S.; Khan, Y.A. Role of energy mix in determining climate change vulnerability in G7 countries. *Sustainability* **2022**, *14*, 2161. [CrossRef]
43. Chang, Q.; Fan, X.; Zou, S. Threshold effects of renewable energy investment on the energy efficiency–fossil fuel consumption nexus: Evidence from 71 countries. *Energies* **2025**, *18*, 2078. [CrossRef]
44. Galvin, R.; Dütschke, E.; Weiß, J. A conceptual framework for understanding rebound effects with renewable electricity: A new challenge for decarbonizing the electricity sector. *Renew. Energy* **2021**, *176*, 423–432. [CrossRef]
45. Wang, L.; Ali, A.; Ji, H.; Chen, J.; Ni, G. Links between renewable and non-renewable energy consumption, economic growth, and climate change: Evidence from five emerging Asian countries. *Environ. Sci. Pollut. Res.* **2023**, *30*, 83687–83701. [CrossRef]
46. Feng, Y.; Xiao, Z.; Zhou, J.; Ni, G. Asymmetrically examining the impact of green finance and renewable energy consumption on environmental degradation and renewable energy investment: The impact of the COVID-19 outbreak on the Chinese economy. *Energy Rep.* **2023**, *9*, 5458–5472. [CrossRef]
47. Hamit-Haggar, M. Greenhouse gas emissions, energy consumption and economic growth: A panel cointegration analysis from Canadian industrial sector perspective. *Energy Econ.* **2012**, *34*, 358–364. [CrossRef]
48. Rasul, K.; Bruckner, M.; Mempel, F.; Trsek, S.; Hertwich, E.G. Energy input and food output: The energy imbalance across regional agrifood systems. *PNAS Nexus* **2024**, *3*, 524. [CrossRef] [PubMed]
49. Toha, M.; Ariyan, T.N.; Alam, M. Methane hotspots in landfills: A systematic review on transboundary dispersion and impacts on heatwaves in South Asia. *Clean. Waste Syst.* **2025**, *10*, 100235. [CrossRef]
50. Pandian, K.; Mustaffa, M.R.A.F.; Mahalingam, G.; Paramasivam, A.; Prince, A.J.; Gajendiren, M.; Mohammad, A.R.R.; Varanasi, S.T. Synergistic conservation approaches for nurturing soil, food security and human health towards sustainable development goals. *J. Hazard. Mater. Adv.* **2024**, *16*, 100479. [CrossRef]
51. Chaparro-Banegas, N.; Sánchez-Garcia, M.; Calafat-Marzal, C.; Roig-Tierno, N. Transforming the agri-food sector through eco-innovation: A path to sustainability and technological progress. *Bus. Strategy Environ.* **2024**, *33*, 9075–9097. [CrossRef]
52. Qin, L.; Raheem, S.; Murshed, M.; Miao, X.; Khan, Z.; Kirikkaleli, D. Does financial inclusion limit carbon dioxide emissions? Analyzing the role of globalization and renewable electricity output. *Sustain. Dev.* **2021**, *29*, 1138–1154. [CrossRef]
53. Adekoya, O.B.; Kenku, O.T.; Oliyide, J.A.; Al-Faryan, M.A.S.; Ogunjemilua, O.D. Does economic complexity drive energy efficiency and renewable energy transition? *Energy* **2023**, *278*, 127712. [CrossRef]
54. McCullagh, P.; Nelder, J.A. *Generalized Linear Models*, 2nd ed.; Chapman and Hall: London, UK; New York, NY, USA, 1989.
55. Wooldridge, J.M. *Econometric Analysis of Cross Section and Panel Data*, 2nd ed.; MIT Press: Cambridge, MA, USA, 2010.
56. Drucker, H.; Burges, C.J.C.; Kaufman, L.; Smola, A.; Vapnik, V. Support vector regression machines. In *Advances in Neural Information Processing Systems 9*; A Bradford Book; The MIT Press: Cambridge, MA, USA; London, UK, 1997; pp. 155–161.
57. Smola, A.J.; Schölkopf, B. A tutorial on support vector regression. *Stat. Comput.* **2004**, *14*, 199–222. [CrossRef]
58. Vapnik, V.N. *The Nature of Statistical Learning Theory*; Springer: New York, NY, USA, 2000.
59. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

60. Strobl, C.; Boulesteix, A.-L.; Zeileis, A.; Hothorn, T. Bias in random forest variable importance measures: Illustrations, sources, and a solution. *BMC Bioinform.* **2007**, *8*, 25. [CrossRef] [PubMed]
61. Fisher, A.; Rudin, C.; Dominici, F. All models are wrong, but many are useful: Variable importance for black-box, proprietary, or inaccessible models. *J. Mach. Learn. Res.* **2019**, *20*, 1–81.
62. Chernozhukov, V.; Chetverikov, D.; Demirer, M.; Duflo, E.; Hansen, C.; Newey, W.; Robins, J. Double/debiased machine learning for treatment and structural parameters. *Econom. J.* **2018**, *21*, C1–C68. [CrossRef]
63. Beşer, N.Ö.; Çabaş, M.; Uğurlu, S.; Magazzino, C. The determinants of life quality in African countries: Moderating role of CO_2 emissions and health expenditures. *Green Technol. Sustain.* **2025**, *3*, 100165. [CrossRef]
64. Magazzino, C.; Mele, M.; Drago, C.; Kuşkaya, S.; Pozzi, C.; Monarca, U. The trilemma among CO_2 emissions, energy use, and economic growth in Russia. *Sci. Rep.* **2023**, *13*, 10225. [CrossRef]
65. Magazzino, C.; Cerulli, G. The determinants of CO_2 emissions in MENA countries: A responsiveness scores approach. *Int. J. Sustain. Dev. World Ecol.* **2019**, *26*, 522–534. [CrossRef]
66. Magazzino, C. Testing the stationarity and convergence of CO_2 emissions series in MENA countries. *Int. J. Energy Sect. Manag.* **2019**, *13*, 977–990. [CrossRef]