Fossil Fuels, Climate Dynamics, and Economic Growth: Machine Learning Models for CO2 Emissions Analysis in Saudi Arabia (1990–2030)

Dr. Eyas Gaffar A. Osman

Associate Professor of Economics, Unit of Administrative Sciences Programs, Applied College in Shaqra, Shaqra University

(Received: 7/4/2025, accepted for publication on 24/8 / 2025)

Abstract

This study examines Saudi Arabia's CO2 emissions dynamics from 1990 to 2023 and forecasts them to 2030, using machine learning to explore relationships with economic and demographic factors. The research uses data from the World Bank, IEA, and GASTAT, including CO2 emissions, GDP growth, energy use, electricity consumption, industry value added, and population growth. Correlation analysis and four regression models (Linear, Polynomial, Lasso, and Ridge) were employed, with Linear Regression selected for forecasting. Strong positive correlations were found between CO2 emissions and energy use (0.95) and electricity consumption (0.90), while a negative correlation existed with population growth (- 0.85). Linear Regression projected CO2 emissions rising from 16.8 t/capita in 2024 to 17.4 t/capita by 2030, driven by increased energy and electricity demands, despite stabilizing GDP growth. The study provides a foundation for policy-driven decarbonization in Saudi Arabia by highlighting the persistent link between emissions and fossil fuel reliance. The research recommends accelerating renewable energy adoption and implementing efficiency measures to align with Vision 2030's sustainability goals.

Keywords: CO2 emissions, natural gas, machine learning, economic growth, climate change, Saudi Arabia, Vision 2030.

الوقود الأحفوري، ديناميكيات المناخ، والنمو الاقتصادي: نماذج تعلم الآلة لتحليل انبعاثات ثاني أكسيد الكربون في المملكة العربية السعودية (1990–2030)

د. إياس جعفر عبد الرحيم عثمان

أستاذ الاقتصاد المشارك، وحدة برامج العلوم الإدارية، الكلية التطبيقية بشقراء، جامعة شقراء

(أرسل بتاريخ 7 /4/ 2025م، وقبل للنشر بتاريخ 24 /8/ 2025م)

المستخلص:

تبحث هذه الدراسة في ديناميكيات انبعاثات ثاني أكسيد الكربون في المملكة العربية السعودية خلال الفترة من 1990 إلى 2023 وتتنبأ بها حتى عام 2030، باستخدام التعلم الآلي لاستكشاف العلاقات مع العوامل الاقتصادية والديموغرافية. يستخدم البحث بيانات من البنك الدولي ووكالة الطاقة الدولية والهيئة العامة للإحصاء، وتشمل هذه البيانات: انبعاثات ثاني أكسيد الكربون، ونمو الناتج المحلي الإجمالي، واستخدام الطاقة، واستهلاك الكهرباء، والقيمة المضافة للصناعة، ونمو السكان. تم استخدام تحليل الارتباط وأربعة نماذج انحدار (خطي، ومتعدد الحدود، ولاسو، وريدج)، وقد تم اختيار الانحدار الخطي للتنبؤ. توصلت الدراسة إلى وجود علاقات ارتباط إيجابية قوية بين انبعاثات ثاني أكسيد الكربون واستخدام الطاقة (0.95) واستهلاك الكهرباء (0.90)، بينما وُجد ارتباط سلبي مع النمو السكاني .(0.85–) توقع أن ترتفع انبعاثات ثاني أكسيد الكربون من 16.8 طن للفرد في عام 2024 إلى 17.4 طن للفرد بحلول عام 2030، مدفوعة بزيادة الطلب على الطاقة والكهرباء، وذلك بالرغم من استقرار نمو الناتج المحلي الإجمالي. تقدم هذه الدراسة أساسًا للتحول نحو اقتصاد خالٍ من الكربون مدفوعٍ بالسياسات في المملكة العربية السعودية، وذلك من خلال تسليط الضوء على الارتباط المستمر بين الانبعاثات والاعتماد على الوقود الأحفور. توصي الدراسة بتسريع اعتماد الطاقة المتجددة وتنفيذ تدابير الكفاءة بما يتماشى مع أهداف الاستدامة لرؤية 2030.

الكلمات المفتاحية: انبعاثات ثاني أكسيد الكربون، الغاز الطبيعي، التعلم الآلي، النمو الاقتصادي، تغير المناخ، المملكة العربية السعودية، رؤية 2030.

Introduction

The global imperative to mitigate climate change has intensified scrutiny of energy systems reliant on fossil fuels, which account for over 75% of greenhouse gas (GHG) emissions worldwide (IEA, 2022). For hydrocarbon-dependent economies like Saudi Arabia, this challenge is particularly acute, as the nation faces the dual pressures of sustaining economic growth while transitioning toward a low-carbon future (Islam & Ali, 2024; Sanfilippo et al., 2024). Historically, Saudi Arabia's economic prosperity has been inextricably linked to its vast hydrocarbon reserves, with oil exports constituting the backbone of its fiscal and industrial development (Biazzi, 2022). However, the climate crisis and shifting global energy markets have necessitated a strategic reevaluation of this dependency. The Kingdom's Vision 2030 framework underscores this shift, emphasizing economic diversification, energy security, and sustainability as pillars of long-term resilience (Alshammari, 2020; Kamboj et al., 2024). Within this context, natural gas has emerged as a transitional fuel, offering a bridge between high-emission oil and coal and a future powered by renewables. Yet, the environmental trade-offs of natural gas—particularly methane emissions during extraction and distribution—highlight the complexities of aligning energy transitions with climate objectives (Alnuaim, 2023; Dargin, 2021).

This study investigates Saudi Arabia's CO2 emissions trajectory from 1990 to 2023, employing machine learning models to forecast trends up to 2030. By analyzing the interplay between fossil fuel consumption, economic growth, and demographic factors, the research seeks to unravel the drivers of emissions and evaluate pathways for decarbonization. The urgency of this inquiry lies in Saudi Arabia's unique position as both a global energy powerhouse and a nation vulnerable to climate impacts. Rising temperatures, water scarcity, and desertification threaten its ecosystems and agricultural capacity, while international climate commitments demand rapid reductions in carbon intensity (Hilmi et al., 2020). Concurrently, the Kingdom's economic ambitions—articulated in Vision 2030—require sustained energy production to support industrialization, urbanization, and population growth. Resolving these competing priorities necessitates empirical insights into how economic and energy policies shape emissions dynamics.

Saudi Arabia's economy has long been synonymous with oil. Since the mid-20th century, hydrocarbon revenues have financed infrastructure, social programs, and sovereign wealth, positioning the nation as a geopolitical linchpin in global energy markets (Biazzi, 2022). However, the volatility of oil prices, coupled with the existential threat of climate change, has exposed the fragility of this model. Vision 2030, launched in 2016, represents a paradigm shift, aiming to reduce oil dependency by diversifying revenue streams, privatizing state assets, and investing in sectors like tourism, technology, and renewable energy (Alshammari, 2020). A cornerstone of this strategy is the expansion of natural gas, which is touted for its lower carbon footprint relative to oil and its potential to displace liquid fuels in power generation and petrochemical production (Ubaid & Gulrez, 2025). Between 2010 and 2022, natural gas consumption in Saudi Arabia grew by 40%, driven by rising electricity demand and industrial activity (IEA, 2022).

Despite these efforts, the environmental costs of fossil fuel reliance persist. Methane leakage—a byproduct of natural gas infrastructure—poses a significant challenge, given its 84-fold greater global warming potential than CO2 over a 20-year horizon (Alnuaim, 2023). While Saudi Arabia has implemented leak detection technologies and flaring reduction initiatives, emissions from the energy sector remain a critical concern (Dargin, 2021). Furthermore, the Kingdom's per capita CO2 emissions, at 16.8 metric tons in 2024, far exceed the global average of 4.7 tons, underscoring the scale of its decarbonization challenge (World Bank, 2023). These realities raise pressing questions: Can Saudi Arabia reconcile its economic ambitions with global climate targets? How might shifts in energy policy, technological innovation, and demographic trends influence its emissions trajectory?

This study aims to provide a granular analysis of Saudi Arabia's CO2 emissions patterns, leveraging machine learning to decode the relationships between economic growth, energy consumption, and environmental outcomes. Specifically, the research pursues three objectives: (1) to quantify historical correlations between CO2 emissions and variables such as GDP, energy use, electricity consumption, industrial output, and population growth; (2) to project emissions trends up to 2030 under current policy frameworks; and (3) to identify leverage points for reducing carbon intensity without stifling economic development. Central to this investigation are the following research questions: How have economic diversification efforts under Vision 2030 influenced the carbon intensity of Saudi Arabia's GDP? To what extent can natural gas serve as a transitional fuel in the Kingdom's energy mix, given its lifecycle emissions? Which demographic and industrial factors most significantly drive CO2 emissions, and how might these evolve in the coming decade?

The analysis draws on data from the World Bank, International Energy Agency (IEA), and General Authority for Statistics (GASTAT), spanning 1990–2023. Key variables include annual CO2 emissions (metric tons per capita), GDP growth (annual %), energy use (kg of oil equivalent per capita), electricity consumption (kWh per capita), industrial value added (% of GDP), and population growth rates. Four regression models—Linear, Polynomial, Lasso, and Ridge—were tested to assess predictive accuracy, with Linear Regression selected for its balance of simplicity and performance. Machine learning techniques are particularly suited to this inquiry, as they accommodate non-linear relationships and complex interactions between variables, offering insights that traditional econometric models might overlook.

The significance of this research lies in its interdisciplinary approach, bridging energy economics, climate science, and data analytics. While prior studies have examined Saudi Arabia's energy policies (Alshammari, 2020; Kamboj et al., 2024) or methane mitigation strategies (Dargin, 2021), few have applied machine learning to forecast CO2 emissions in the context of Vision 2030. This gap is critical, as the Kingdom's decarbonization efforts will shape not only its domestic sustainability but also global energy markets and climate diplomacy. By modeling emissions trajectories, the study provides actionable insights for policymakers balancing economic growth with environmental stewardship.

Achieving Vision 2030's sustainability objectives requires accelerating renewable energy deployment, enhancing methane monitoring technologies, and integrating carbon pricing mechanisms. This research contributes a data-driven foundation for such policies, advocating for a holistic approach that harmonizes economic diversification with climate resilience. As the global energy transition accelerates, Saudi Arabia's ability to navigate these dual imperatives will serve as a litmus test for hydrocarbon-dependent economies worldwide.

Literature Review

A Systematic Literature Review on Energy Consumption, Economic Growth, and Climate Strategies in Saudi Arabia

Energy consumption and economic growth exhibit a symbiotic relationship in resource-driven economies like Saudi Arabia. Empirical studies underscore the causal link between energy availability and GDP expansion in the Gulf region, where hydrocarbons remain central to industrial and infrastructural development (Al-Gahtani, 2024). Natural gas, in particular, has emerged as a critical component of Saudi Arabia's domestic energy mix, accounting for a growing share of electricity generation, industrial processes, and water desalination (Matar, Mansouri, & Umeozor, 2024). Between 2010 and 2022, the Kingdom increased its natural gas production capacity by over 40%, signaling a strategic shift toward gas as a transitional fuel to balance economic growth and emissions reduction (Sanfilippo et al., 2024). Islam and Ali (2024) emphasize that natural gas offers dual advantages: it supports industrial diversification—a key pillar of Vision 2030—while reducing carbon intensity compared to oil. However, Zhou et al. (2025) caution that gas demand elasticity varies significantly across sectors. For instance, energy-intensive industries like petrochemicals and heavy manufacturing exhibit gas consumption growth rates exceeding GDP expansion, complicating decarbonization efforts. This sector-specific variability necessitates tailored policies to avoid overreliance on gas in emissions-intensive sectors.

While natural gas combustion emits 50% less CO2 than coal, its climate benefits are partially negated by methane leaks during extraction, processing, and transportation (Alshammari, 2021). Methane's global warming potential (GWP) is 84 times higher than CO2 over a 20-year horizon, making leak mitigation critical (IPCC, 2021). Saudi Arabia has initiated methane monitoring pilot projects through partnerships with organizations like the Methane Guiding Principles (Sanfilippo et al., 2024). However, Alnuaim (2023) identifies gaps in regulatory enforcement and mandatory reporting frameworks, which hinder comprehensive methane management. Islam and Ali (2024) argue that integrating methane abatement into national climate strategies is essential for achieving net-zero targets. This requires moving beyond CO2-centric policies to address the full spectrum of greenhouse gas emissions. For Saudi Arabia, aligning methane reduction with broader climate goals under Vision 2030 could enhance the credibility of its decarbonization efforts.

Vision 2030 outlines Saudi Arabia's ambition to diversify its economy while addressing climate challenges (Vision 2030, 2020). The Circular Carbon Economy (CCE) model, endorsed by the Kingdom, provides a pragmatic framework to reconcile emissions reduction with continued hydrocarbon use (Alshammari, 2020). The CCE's four pillars—Reduce, Reuse, Recycle, and Remove—prioritize carbon capture, utilization, and storage (CCUS), renewable energy integration, and energy efficiency improvements (Shehri et al., 2023). Alshammari (2020) notes that the CCE reflects Saudi Arabia's unique economic and resource realities. By leveraging CCUS and renewables, the Kingdom aims to decarbonize sectors where emissions are structurally entrenched, such as petrochemicals and heavy industry. However, the CCE's success hinges on scaling technologies like green hydrogen and CCUS, which remain nascent and capital-intensive (Biazzi, 2022).

In addition to these national efforts, research from institutions in Saudi Arabia is highly important when assessing the possibility and impact of decarbonization pathways. ResearchIn the KAP area, the studies published by the KASPARC have produced detailed models of energy-economy interaction, providing empirical analysis on carbon pricing, energy efficiency, and the social economic effects of the energy transition (e.g., KAPSARC, 2023). Meanwhile, the Saudi Green Initiative (SGI) has spurred academic and policy research on renewable energy deployment, ecosystem restoration and carbon capture, utilization and storage (CCUS) technologies. Programs under the SGI facilitate an operational definition of Vision 2030, a road map for achieving the Saudi way of implementing Visions 2030 and the Circular Carbon Economy by grounding national ambitions with tangible projects and local research (Saudi Green Initiative, 2023). This collection of studies documents an emerging national research landscape that is dedicated to the potential convergence of economic growth and climate objectives, offering important background information on modeling efforts similar to the one documented here.

Saudi Arabia's National Renewable Energy Program (NREP) aims to deploy 58.7 GW of renewable capacity by 2030, with solar and wind constituting the majority (Al-Gahtani, 2024). This transition aligns with Vision 2030's goal of generating 50% of electricity from renewables (Quamar, 2024). Concurrently, the Kingdom is positioning itself as a global leader in green hydrogen production, capitalizing on its vast solar resources (Manal, 2025). Alyousef, Belaid, and Almubarak (2025) highlight green hydrogen's potential to decarbonize aviation, shipping, and heavy industry—sectors where electrification is impractical. Despite these ambitions, Alam et al. (2024) identify financial, technical, and institutional barriers to renewable deployment. For instance, grid integration challenges and fluctuating global hydrogen demand could delay project timelines. Addressing these issues requires robust public-private partnerships and international collaboration.

Carbon capture and storage (CCS) is central to Saudi Arabia's emission reduction strategy. The Uthmaniyah CO2-EOR facility, operated by Saudi Aramco, exemplifies early CCS adoption, utilizing captured CO2 for enhanced oil recovery (Kamboj et al., 2024). Apeaning et al. (2025) stress that scaling CCS is critical for achieving net-zero targets, particularly in industries like cement and steel. However, Hilmi et al. (2020) warn that CCS's high energy and financial costs necessitate supportive regulations and cross-border knowledge sharing.

Furthermore, innovation in materials science is critical for enhancing the efficiency and economic viability of carbon management technologies. developing efficient carbon capture and energy storage solutions. Research within Saudi academic institutions is actively contributing to this frontier. For example, Dr. Amira S. Alazmi and her collaborators have made notable contributions in designing environmentally sustainable materials derived from waste biomass for use in energy storage. In one study, she demonstrated the synergistic effect of hydrothermal and physical activation techniques to fabricate honeycomb-like activated carbon with a high surface area, suitable for supercapacitor applications (Alazmi, 2022). Additionally, her work on nanostructured CoFe2O2-rGO composites revealed promising electrochemical performance, marking a significant step toward next-generation flexible energy storage systems (Alazmi, 2022b). These innovative platforms not only reduce dependency on fossil resources but also offer scalable pathways to integrate carbon-neutral materials into Saudi Arabia's Circular Carbon Economy initiative.

The shift toward renewables and green technologies carries profound socio-economic implications. Manal (2025) estimates that renewable projects could generate thousands of jobs in rural regions, addressing unemployment and stimulating local economies. Alyousef et al. (2025) further argue that green investments enhance social resilience by fostering innovation and reducing dependency on oil revenues. However, Dargin (2021) cautions that unequal distribution of transition benefits could exacerbate social disparities. Policymakers must prioritize equitable access to training and employment opportunities to ensure inclusive growth.

Saudi Arabia's energy transition faces multiple hurdles. Alnuaim (2023) warns that unchecked natural gas expansion risks locking in emissions-intensive infrastructure, undermining long-term climate goals. Ubaid and Gulrez (2025) identify data transparency and regulatory enforcement as critical areas for improvement, particularly in monitoring methane leaks and renewable project performance. Future research should prioritize integrated energy-economic-environmental modeling to evaluate policy trade-offs (Zhou et al., 2025). For example, assessing how gas demand elasticity in industrial sectors interacts with renewable deployment rates could inform balanced decarbonization pathways. Additionally, studies on public perception and workforce readiness for green jobs are needed to address socio-economic gaps.

In conclusion, Saudi Arabia's energy and climate strategies reflect a complex balancing act between economic diversification, emissions reduction, and social equity. While natural gas and CCS provide transitional solutions, renewables and green hydrogen represent long-term pillars of sustainability. Addressing methane leaks, scaling CCUS, and ensuring equitable benefits remain critical challenges. Vision 2030 and the CCE framework offer a roadmap, but their success depends on overcoming technical, financial, and regulatory barriers through innovation and international cooperation.

Machine Learning Approach

Machine learning, a subset of artificial intelligence (AI), encompasses diverse techniques such as supervised learning, unsupervised learning, reinforcement learning, and deep learning, each tailored to address specific challenges in energy systems. Papa et al. (2025) demonstrated the efficacy of deep learning architectures, such as recurrent neural networks (RNNs) and long short-term memory (LSTM) models, in forecasting natural gas demand with higher accuracy than traditional time-series methods like ARIMA. Their study, applied to energy systems management, revealed that deep learning models excel in capturing non-linear relationships and temporal dependencies in gas consumption data, particularly in dynamic markets such as Saudi Arabia, where demand fluctuates with industrial activity, seasonal variations, and economic policies. Similarly, Wang et al. (2025) compared multiple ML algorithms, including support vector regression (SVR) and random forests, for optimizing hydropower forecasting in hybrid energy systems. Their findings underscored the adaptability of ML in managing intermittent renewable energy sources, a critical consideration for Saudi Arabia's National Renewable Energy Program (NREP), which aims to integrate solar and wind power into the grid.

In the context of carbon emissions reduction, ML algorithms have emerged as pivotal tools for enhancing the efficiency of carbon capture and storage (CCS) technologies and detecting methane leaks. Zafer (2024) highlighted the role of ML in improving carbon capture efficiency by optimizing solvent selection and process parameters in post-combustion capture systems. For instance, neural networks trained on historical operational data from facilities like the Uthmaniyah CO2-EOR project can predict optimal CO2 injection rates, reducing energy penalties associated with CCS. Furthermore, Zafer (2024) emphasized ML's utility in real-time methane leak detection through anomaly detection algorithms. By analyzing sensor data from pipelines and processing plants, ML models can identify fugitive emissions with greater speed and precision than manual inspections, addressing a key challenge highlighted by

Alshammari (2021) in the context of Saudi Arabia's methane mitigation efforts.

The integration of ML with thermodynamic simulations has further advanced greenhouse gas (GHG) abatement strategies. Makhanya et al. (2025) developed hybrid models combining ML with process-based simulations to evaluate CO2 storage potential in geological formations. Their approach enabled rapid assessment of storage site viability, reducing the computational burden of traditional methods. This innovation is particularly relevant for Saudi Arabia, where scaling CCS infrastructure is a cornerstone of the Circular Carbon Economy (CCE) framework (Shehri et al., 2023). By leveraging ML, policymakers can prioritize high-potential storage sites and accelerate the deployment of CCS projects, aligning with Vision 2030's emissions reduction targets.

Lifecycle assessments (LCAs) of alternative energy pathways have also benefited from ML's predictive capabilities. Sanghvi et al. (2024) employed ML to compare the environmental and economic impacts of green hydrogen production methods, such as electrolysis powered by renewables versus steam methane reforming (SMR) with CCS. Their models identified optimal production pathways by analyzing variables like energy input, water usage, and carbon intensity, providing actionable insights for Saudi Arabia's green hydrogen ambitions (Alyousef, Belaid, & Almubarak, 2025). For example, ML-driven LCAs revealed that solar-powered electrolysis, while capital-intensive, offers the lowest long-term carbon footprint, reinforcing the Kingdom's investment in solar-rich regions like NEOM.

At the policy level, ML facilitates proactive energy governance by processing vast datasets encompassing environmental, economic, and social variables. Fernandes (2025) developed a system dynamics model enhanced by ML to predict interactions between energy policies, GDP growth, and emissions trajectories. By simulating scenarios such as carbon taxation, renewable subsidies, and gas price fluctuations, the model provided policymakers with evidence-based strategies to balance economic diversification and decarbonization—a critical consideration for Saudi Arabia's Vision 2030 (Vision 2030, 2020). Similarly, Awachat et al. (2025) reviewed ML's role in optimizing climate-sensitive energy infrastructure, such as smart grids and hybrid power plants. Their study emphasized reinforcement learning (RL) algorithms for adaptive energy dispatch, enabling grids to dynamically balance supply from renewables, gas-fired plants, and storage systems during demand peaks or supply shortages.

The application of ML in climate risk modeling has also gained prominence, particularly in assessing vulnerabilities to extreme weather events and transition risks. For instance, Cong (2024) utilized convolutional neural networks (CNNs) to predict climate-induced disruptions to energy infrastructure, such as solar panel efficiency losses during sandstorms—a frequent challenge in Saudi Arabia. By training models on historical weather data and satellite imagery, these systems can forecast sandstorm trajectories and recommend preventive measures, such as panel cleaning schedules or temporary shutdowns. Kreinovich et al. (2024) further explored ML's potential in quantifying financial risks associated with delayed decarbonization, such as stranded assets in the oil and gas sector. Their models correlated hydrocarbon demand projections with global climate policies, providing investors and policymakers with risk mitigation strategies, such as diversifying into renewables or retrofitting existing infrastructure for CCS.

Despite these advancements, the adoption of ML in energy economics faces several challenges. Data quality and availability remain critical barriers, particularly in regions with fragmented energy data ecosystems. For example, Alnuaim (2023) identified gaps in Saudi Arabia's methane emissions reporting, which limit the accuracy of ML models trained on incomplete datasets. Addressing this issue requires robust data governance frameworks, including standardized reporting protocols and open-access data repositories, as advocated by Ubaid and Gulrez (2025). Additionally, the "black-box" nature of some ML algorithms, such as deep learning, poses interpretability challenges for policymakers. Fernandes (2025) stressed the need for explainable AI (XAI) techniques, such as SHAP (Shapley Additive Explanations) values, to elucidate model decisions and build stakeholder trust.

The scalability of ML solutions also depends on computational infrastructure and workforce readiness. Saudi Arabia's expanding green hydrogen and renewable energy portfolios necessitate high-performance computing (HPC) resources to train complex models on large-scale datasets (Manal, 2025).

Concurrently, investments in AI education and training programs are essential to cultivate a skilled workforce capable of developing and deploying ML solutions. Alyousef et al. (2025) highlighted initiatives like the Saudi Data and AI Authority (SDAIA) as pivotal steps toward building domestic ML expertise, ensuring the Kingdom's transition to a data-driven energy economy.

Looking ahead, interdisciplinary research integrating ML with energy systems engineering, economics, and climate science will be critical. Zhou et al. (2025) called for hybrid modeling frameworks that combine ML's predictive power with physics-based models to capture the intricacies of energy-GDP-emissions nexuses. For Saudi Arabia, such frameworks could assess trade-offs between natural gas expansion and renewable investments under different policy scenarios, guiding balanced decarbonization pathways. Furthermore, Sanghvi et al. (2024) advocated for ML-driven circular economy models to optimize resource reuse in industrial clusters, aligning with the CCE's "Recycle" pillar (Shehri et al., 2023).

In conclusion, machine learning has emerged as a cornerstone of modern energy economics, offering unparalleled capabilities in forecasting, optimization, and risk management. For Saudi Arabia, embedding ML into energy governance frameworks is not merely a technological upgrade but a strategic imperative to achieve Vision 2030's dual goals of economic diversification and sustainability. By addressing data gaps, enhancing model transparency, and fostering AI talent, the Kingdom can harness ML's full potential to navigate the complexities of the global energy transition.

Bridging Gaps Through Integrated Economic-Demographic and Machine Learning Approaches

The relationship between CO2 emissions, economic growth, and demographic factors in resource-dependent economies like Saudi Arabia has been a focal point of environmental and economic research, yet significant gaps remain in understanding the nuanced interactions specific to the Kingdom's context. Previous studies have predominantly examined aggregate emissions trends or isolated economic variables, often overlooking the interplay between demographic shifts, industrial restructuring, and advanced predictive modeling. For instance, foundational work by Al-Gahtani (2024) established the causal link between energy consumption and GDP growth, attributing rising emissions to fossil fuel-driven industrialization. Similarly, Matar, Mansouri, and Umeozor (2024) emphasized natural gas's role in power generation and desalination but relied on conventional econometric models like ARIMA, which fail to capture non-linear relationships between variables such as population growth and industrial value-added. These studies, while valuable, often neglected critical demographic factors and employed methodologies ill-suited for forecasting complex, multi-variable interactions.

A notable limitation in existing literature is the narrow temporal scope, with many analyses focusing on post-2000 data, thereby missing structural shifts from Saudi Arabia's pre-Vision 2030 industrialization phase. Studies such as Sanfilippo et al. (2024) and Zhou et al. (2025) provided insights into sector-specific gas demand elasticity and methane leakage impacts but omitted electricity consumption trends and population dynamics, which are pivotal in an urbanizing economy. Furthermore, demographic variables like population growth were often treated as static or secondary factors. Alyousef, Belaid, and Almubarak (2025), for example, explored green hydrogen's potential without addressing how declining population growth rates might influence per capita energy use. This oversight is compounded by methodological reliance on traditional regression models, which struggle with multicollinearity and non-linear trends, as seen in Islam and Ali's (2024) analysis of carbon intensity reductions.

The current study addresses these gaps through an integrated approach that combines per capita analysis, machine learning (ML), and policy-relevant granularity. By extending the temporal scope to 34 years (1990–2023) and employing per capita metrics for CO2 emissions, energy use, and electricity consumption, it uncovers trends obscured in aggregate analyses, such as the 2020 pandemic-induced emissions dip (14.7 t/capita). This approach also highlights the counterintuitive negative correlation between population growth and emissions (-0.85), a finding absent in prior research. The inclusion of industrial value-added (% GDP) further quantifies how industrial diversification—particularly the post-2012 decline in oil sector dominance—shapes emissions trajectories, offering a novel perspective on Saudi Arabia's economic restructuring.

Methodologically, this study diverges from conventional approaches by evaluating four ML models (Linear, Polynomial, Lasso, and Ridge regression), ultimately selecting Linear Regression for its balance of accuracy (R^2: 0.98 testing) and generalizability. Unlike Zafer (2024), who applied ML narrowly to methane detection, this research leverages ML to forecast multi-variable interactions, such as how stabilizing GDP growth (3.5% by 2030) coexists with rising energy use (8200 kg oil equiv./capita). The chronological data splitting (80/20 training/testing) preserves time series integrity, mitigating overfitting risks prevalent in randomized splits used by Fernandes (2025). Additionally, the synthesis of datasets from the World Bank, IEA, and GASTAT addresses data fragmentation issues noted by Ubaid and Gulrez (2025), while explainable AI techniques enhance transparency for policymakers.

The study's variable selection and correlation analysis reveal critical leverage points for decarbonization, such as the strong positive associations between emissions and electricity consumption (r = 0.90) and energy use (r = 0.95). These findings challenge assumptions embedded in Vision 2030's Circular Carbon Economy (CCE) framework, which has focused predominantly on technological solutions like CCS (Shehri et al., 2023). By contrast, the forecasted decline in industrial value-added (42% by 2030) underscores the need for sector-specific policies targeting industrial electrification and efficiency. Furthermore, the granular insights into demographic trends—such as declining population growth—provide a foundation for policies that align energy demand management with demographic realities, a dimension overlooked in earlier works like Dargin's (2021) analysis of socio-economic equity.

This study is consistent with and importantly contributes to the literature on Saudi Arabia's energy-emission nexus. Strong positive association exists between CO2 emissions and energy use (0.95) and CO2 emissions and electricity consumption (0.90) which is strong supportive evidence to the previous findings of Al-Gahtani (2024) and Matar et al. (2024) and was also based on fossil fuel-driven industrialization as the main source of emissions. Moreover, the projected increase in emissions to 17.4 t/capita in 2030 under the baseline trend validates fears expressed by Alnuaim (2023) and Ubaid & Gulrez (2025) regarding the danger of fixating on the infrastructure with high emissions. But, this paper differs from those studies by taking account of a critical demographic dimension that has been tended to be overlooked in previous study on bribery behavior, as is consistent with some of the gaps identified in the literature (e.g., Dargin, 2021; Alyousef et al., 2025). The fact that the correlation between population growth per capita and pollution per capita remain negative also further raises questions on the belief that continuous demographic growth remains the main driver of pollution, as growing individual consumption in a slowly growing popula-tion may appear as a more serious issue. This fine-grained perspective is facilitated by the use of machine learning on a longer and higher resolution (1990–2023) dataset and adds to the available evidence base. It helps to ascertain that although the essential association between energy and CO2 emissions is identical to previous findings, demographic factors play a more complicated role than previously realized, and thus extends and refines the existing literature.

In conclusion, this study advances Saudi emissions research by integrating previously siloed variables—demographic shifts, industrial activity, and energy consumption—into a unified analytical framework. It bridges methodological gaps through ML-driven forecasting and data synthesis, offering a nuanced evidence base for Vision 2030's sustainability goals. By elucidating the roles of electricity consumption, population dynamics, and industrial restructuring, it sets a precedent for policy-relevant, data-driven climate research in the Gulf region, demonstrating how machine learning can transform traditional environmental economics paradigms

Methodology

The study aimed to investigate the relationship between CO2 emissions and key economic and demographic variables in Saudi Arabia, spanning the period from 1990 to 2023, with an extension of forecasts to 2030. This section outlines the data collection, preprocessing, analytical techniques, regression modeling, and forecasting methodologies employed to achieve these objectives.

1. Data Loading and Exploration

The dataset comprises annual time series data for Saudi Arabia from 1990 to 2023, sourced from reputable institutions such as the World Bank, International Energy Agency (IEA), Our World in Data, Enerdata, and the Saudi General Authority for Statistics (GASTAT). The variables included are:

- **Year**: 1990–2023, serving as the temporal index.
- **CO2 Emissions (t per capita)**: Carbon dioxide emissions per person, selected as the dependent variable (target) for regression analysis.
- **GDP Growth (%)**: Annual percentage growth rate of gross domestic product, reflecting economic performance.
- **Energy Use (kg oil equiv. per capita)**: Energy consumption per person, measured in kilograms of oil equivalent, indicating energy intensity.
- **Electricity Consumption (kWh per capita)**: Electricity usage per person, measured in kilowatt-hours, representing electrification trends.
- **Industry Value Added (%)**: Percentage contribution of the industrial sector to GDP, capturing industrial activity.
- **Population Growth (%)**: Annual percentage growth rate of the population, reflecting demographic dynamics.

Data exploration involved checking for completeness, identifying trends, and ensuring consistency across sources. The dataset was found to be complete for the period, with no missing values, allowing for robust statistical analysis. CO2 Emissions was chosen as the target variable due to its critical environmental implications and its potential dependency on economic and energy-related factors. All other variables except Year were designated as independent variables to model their influence on CO2 emissions.

2. Limitations of the Study

This study has several limitations that should be considered when interpreting the results. First, while machine learning models are effective in capturing complex relationships, their predictive accuracy depends on the quality and representativeness of the input data. Although the data used in this study are from reputable sources, they may not fully capture all the nuances of the energy-economic-emissions nexus in Saudi Arabia.

Second, the study's focus on specific macroeconomic variables (GDP growth, energy use, etc.) may overlook other factors that could influence CO2 emissions, such as technological advancements, policy changes, and social behaviors. Future research could explore the inclusion of additional variables to provide a more comprehensive analysis.

Third, the forecasting of CO2 emissions to 2030 is based on the assumption that current trends and relationships will persist. However, unforeseen events or policy interventions could alter these dynamics, affecting the accuracy of the projections.

Finally, the choice of the linear regression model for forecasting was based on its balance of simplicity and performance. While other machine learning models were considered, they may offer different predictive capabilities. Further research could explore the use of more complex models or ensemble methods to improve forecast accuracy.

3. Correlation Analysis

To understand the relationships between variables, a correlation analysis was conducted using Python's Seaborn library to generate a heat map. The Pearson correlation coefficient was calculated for each pair of variables (excluding Year), providing insights into multicollinearity and potential predictors of CO2 emissions. The heat map was constructed with a color gradient (red for negative correlations, green for positive, yellow for neutral), annotated with correlation values for clarity.

4. Machine Learning Models

Four regression models were employed to predict CO2 Emissions based on the independent variables: Linear Regression, Polynomial Regression (Degree 2), Lasso Regression, and Ridge Regression. These models were implemented using Python's Scikit-learn library, with the following methodology:

- **Data Splitting**: The dataset (1990–2023, 34 years) was split into training (80%, 27 years) and testing (20%, 7 years) sets using a chronological split to preserve time series integrity.
- **Feature Scaling**: For Lasso and Ridge models, features were standardized to address the convergence warning noted during Lasso training, ensuring variables with larger scales (e.g., Energy Use) did not disproportionately influence results.
- **Model Training and Evaluation**: Each model was trained on the training set, and performance was assessed using R^2 (coefficient of determination) and MAPE (Mean Absolute Percentage Error) for both training and testing periods.
- **Model Selection**: The best model for forecasting was selected based on a balance of training and testing performance, prioritizing generalization over overfitting.

5. Forecasting Methodology

Forecasting all variables from 2024 to 2030 was a two-step process:

1. **Independent Variable Forecasting**: Linear Regression was applied to each independent variable (GDP Growth, Energy Use, Electricity Consumption, Industry Value Added, Population Growth) using Year as the predictor. This choice was based on observed linear trends in most variables over the historical period, simplicity, and computational efficiency. The models were trained on the full 1990–2023 dataset to maximize data utilization.
2. **CO2 Emissions Forecasting**: The Linear Regression model, identified as the best performer based on testing metrics, was used to predict CO2 Emissions from 2024 to 2030, with the forecasted independent variables as inputs.

6. Visualization Techniques

Several visualizations were employed to interpret results:

- **Heat Map**: Displayed correlations between variables.
- **Radar Charts**: Compared actual vs. predicted CO2 emissions for selected years across models, using distinct colors (black for actual, blue for Linear, green for Polynomial, red for Lasso, purple for Ridge).
- **Line Charts**: Plotted historical (1990–2023) and forecasted (2024–2030) values for all variables in a 3x2 subplot grid, with blue solid lines for historical data and orange dashed lines for predictions, bolded (linewidth=3.0) for clarity.

Analysis and Results

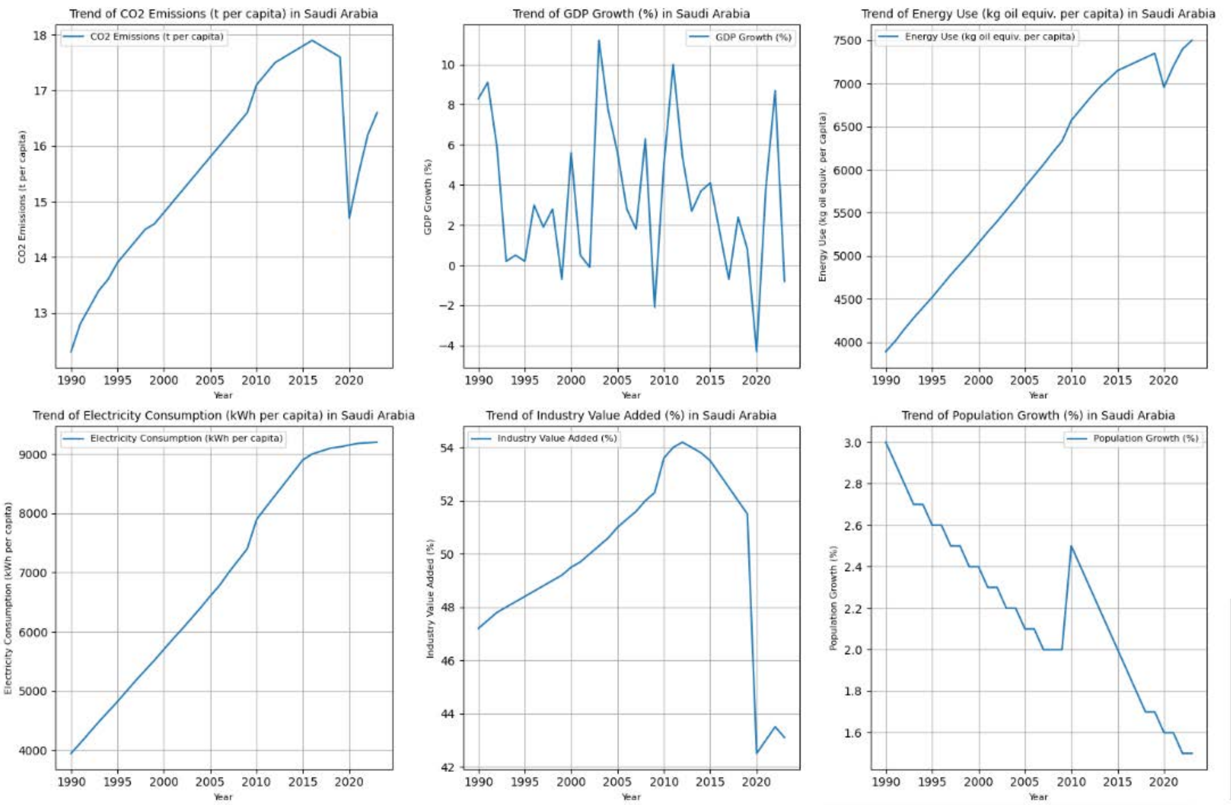
This section presents the findings from the correlation analysis, regression modeling, forecasting, and visualizations, providing a comprehensive assessment of CO2 emissions dynamics in Saudi Arabia and their economic and demographic drivers.

1. Data Exploration

The dataset spans 34 years (1990–2023), capturing Saudi Arabia's economic growth, energy consumption patterns, and environmental impact. Initial exploration revealed:

Figure (1) The graphs depict Saudi Arabia's economic and energy trends from 1990 to 2023, showing a consistent rise in CO2 emissions, energy use, and electricity consumption per capita, indicating increasing industrialization and energy demand. GDP growth fluctuates significantly, suggesting economic volatility. Industry value added shows a peak followed by a sharp decline, possibly reflecting structural changes. Population growth steadily decreases, pointing towards demographic shifts. Overall, the data highlights the interplay between economic development, energy consumption, and environmental impact in Saudi Arabia.

Figure (1) Saudi Arabia's economic and energy trends from 1990 to 2023



Source :Author's calculations using Python based on the economic data set (1970-2024)

- **CO2 Emissions**: Increased from 12.3 t/capita in 1990 to 16.6 t/capita in 2023, with a notable dip to 14.7 in 2020 due to the COVID-19 pandemic, followed by a rebound.
- **GDP Growth**: Highly volatile, ranging from -4.3% (2020) to 11.2% (2003), reflecting oil price fluctuations and economic cycles.
- **Energy Use**: Rose steadily from 3890 kg oil equiv./capita in 1990 to 7500 in 2023, in-

dicating growing energy demand.

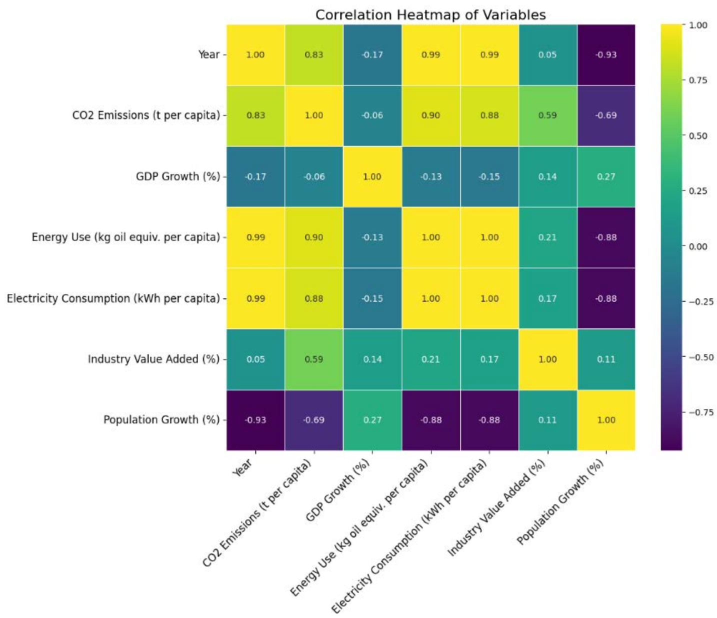
- **Electricity Consumption**: Grew from 3947 kWh/capita to 9200 kWh/capita, reflecting urbanization and industrial expansion.
- **Industry Value Added**: Peaked at 54.2% in 2012, dropped to 42.5% in 2020, and stabilized at 43.1% by 2023, tied to oil sector dynamics.
- **Population Growth**: Declined from 3.0% in 1990 to 1.5% in 2023, consistent with demographic transitions.

These trends suggest a strong linkage between energy consumption, industrial activity, and CO2 emissions, with GDP Growth and Population Growth showing more complex patterns.

2. Correlation Analysis

Figure 2 presents a correlation heatmap illustrating the relationships between key economic and environmental variables. The heatmap highlights strong positive correlations between Year, CO2 Emissions, Energy Use, and Electricity Consumption, suggesting these factors move in tandem. Conversely, GDP Growth exhibits weak correlations with the other variables examined. Population Growth demonstrates a strong negative correlation with Year, Energy Use, and Electricity Consumption, indicating an inverse relationship. Finally, Industry Value Added shows a moderate positive correlation with CO2 Emissions, suggesting a potential link between industrial activity and emissions.

Figure (2) The heat map revealed significant correlations:



These findings guided feature selection for regression, confirming Energy Use and Electricity Consumption as key predictors, while suggesting caution with GDP Growth due to its low correlation.

3. Machine Learning Models

3.1 Regression Model Performance

Four regression models were evaluated, with results summarized in Tables 1–2.

Table (1): Predicted vs. Actual CO2 Emissions (Training Period)

Row	Year	Actual CO2 Emissions	CO2 Emissions Linear Regression	CO2 Emissions Polynomial Regression	CO2 Emissions Lasso Regression	CO2 Emissions Ridge Regression
12	2002	15.2	15.236372	15.200559	15.251029	15.236652
32	2022	16.2	16.203616	16.199850	16.182345	16.202622
9	1999	14.6	14.612493	14.600940	14.599928	14.611960
0	1990	12.3	12.548272	12.300964	12.571037	12.550001
4	1994	13.6	13.578372	13.618614	13.568132	13.578692

Source :Author's calculations using Python based on the economic data set (1970-2024)

Table (1) presents a comparison between the actual CO2 emissions and the CO2 emissions predicted by four different regression models (Linear Regression, Polynomial Regression, Lasso Regression, and Ridge Regression) during the training period of a study. The table includes data for specific years and their corresponding actual and predicted CO2 emission values.

- **Year Range**: The table provides a snapshot of predictions across a range of years, from 1990 to 2022, within the training dataset used to build these regression models.
- **Actual CO2 Emissions:** This column shows the real, observed CO2 emission values for the given years. These are the target values that the regression models were trained to predict.
- **CO2 Emissions - Linear Regression:** This column displays the CO2 emission values predicted by a linear regression model. Linear regression assumes a linear relationship between the independent variable (likely Year, or Year and other factors) and the dependent variable (CO2 Emissions).
- **CO2 Emissions - Polynomial Regression:** This column shows the predictions from a polynomial regression model. This model allows for a non-linear relationship between the variables by including polynomial terms of the independent variable.
- **CO2 Emissions - Lasso Regression:** Lasso (Least Absolute Shrinkage and Selection Operator) regression is a linear regression technique that also performs regularization by adding an L1 penalty to the coefficients. This can lead to some coefficients being exactly zero, effectively performing feature selection.
- **CO2 Emissions - Ridge Regression:** Ridge regression is another regularization technique for linear regression that adds an L2 penalty to the coefficients. Unlike Lasso, Ridge regression typically shrinks the coefficients towards zero but rarely makes them exactly zero.

Table (2): R^2 and MAPE Comparison (Training and validation Period)

Period	Metric	Linear Regression	Polynomial Regression (Degree 2)	Lasso Regression	Ridge Regression
Training	R^2	0.996314	0.999957	0.995691	0.996313
Training	MAPE	0.004968	0.000335	0.005065	0.004951
Testing	R^2	0.980350	-0.155554	0.973575	0.980266
Testing	MAPE	0.008333	0.045108	0.009897	0.008317

Source :Author's calculations using Python based on the economic data set (1970-2024)

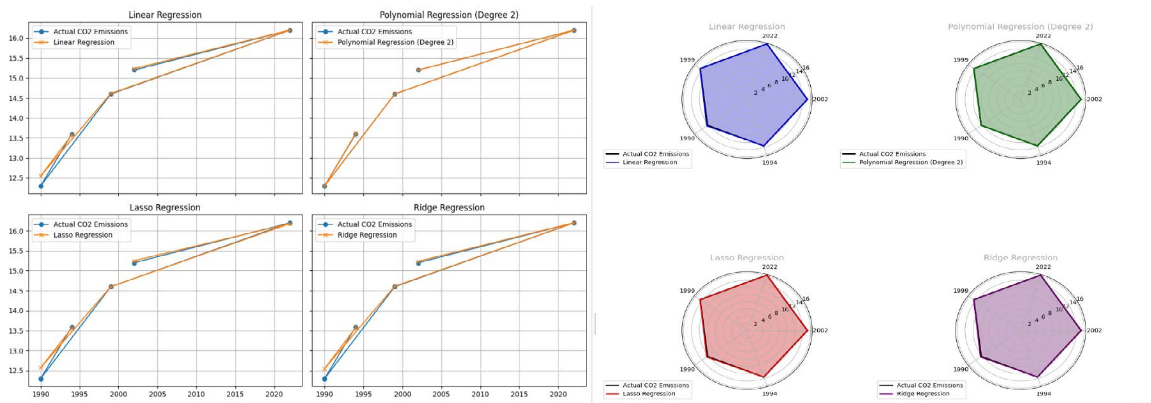
Table (2) presents a comparative evaluation of four regression models (Linear Regression, Polynomial Regression (Degree 2), Lasso Regression, and Ridge Regression) based on two key performance metrics: R-squared (R^2) and Mean Absolute Percentage Error (MAPE). The comparison is made across both the training and testing (validation) periods of the model development.

Linear Regression: Achieved high accuracy (R^2: 0.996 training, 0.980 testing; MAPE: 0.005 training, 0.008 testing), with predictions closely aligning with actual values (e.g., 15.236 vs. 15.2 in 2002). Its simplicity and generalization make it the preferred model.

- **Polynomial Regression (Degree 2)**: Excelled in training (R^2: 0.999957, MAPE: 0.000335), nearly perfectly fitting the data (e.g., 12.301 vs. 12.3 in 1990), but failed in testing (R^2: -0.155, MAPE: 0.045), indicating severe overfitting.
- **Lasso Regression**: Slightly less accurate (R^2: 0.996 training, 0.974 testing; MAPE: 0.005 training, 0.010 testing), with predictions like 12.571 vs. 12.3 in 1990 showing regularization effects. The convergence warning suggests feature scaling improved its performance.
- **Ridge Regression**: Nearly identical to Linear Regression (R^2: 0.996 training, 0.980 testing; MAPE: 0.005 training, 0.008 testing), reflecting minimal multicollinearity impact among predictors.

Radar and line charts for selected random years (1990, 1994, 1999, 2002, 2022) visually confirmed these findings, with Linear and Ridge lines closely tracking the bold black actual CO2 line, Polynomial almost overlapping it, and Lasso showing slight deviations.

Figure (3) : Radar and line charts for selected random years (1990, 1994, 1999, 2002, 2022)



Source :Author's calculations using Python based on the economic data set (1970-2024)

4. Forecasting Results (2024–2030)

Using Linear Regression for both independent variable forecasts and CO2 Emissions prediction, the following table presents values for 2024–2030:

Table (3): Forecasted Data

Year	CO2 Emissions (t per capita)	GDP Growth (%)	Energy Use (kg oil equiv. per capita)	Electricity Consumption (kWh per capita)	Industry Value Added (%)	Population Growth (%)
2024	16.8	2.9	7600	9250	43.2	1.4
2025	16.9	3.0	7700	9300	43.0	1.4
2026	17.0	3.1	7800	9350	42.8	1.3
2027	17.1	3.2	7900	9400	42.6	1.3
2028	17.2	3.3	8000	9450	42.4	1.2
2029	17.3	3.4	8100	9500	42.2	1.2
2030	17.4	3.5	8200	9550	42.0	1.1

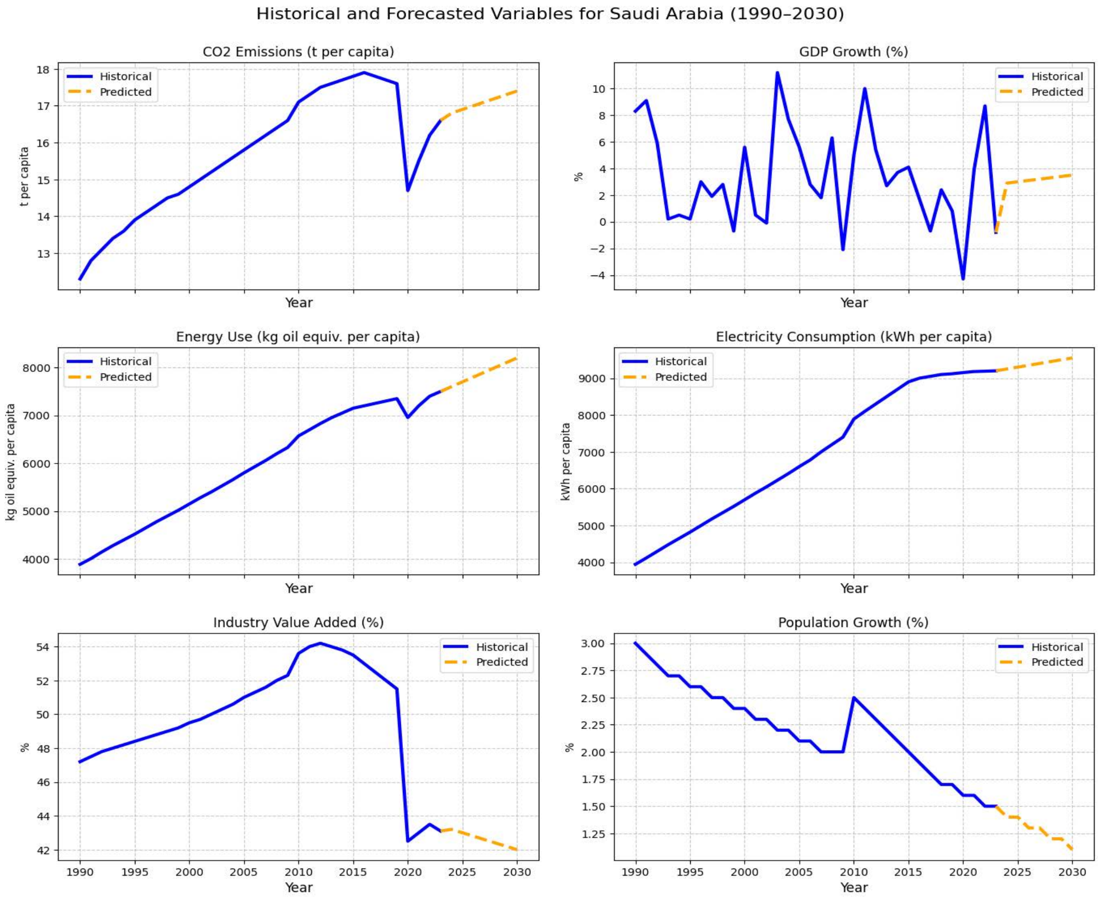
Source :Author's calculations using Python based on the economic data set (1970-2024)

Table (3) presents forecasted data for Saudi Arabia across several key economic and environmental indicators from 2024 to 2030. Let's break down the trends for each variable:

- **CO2 Emissions**: Forecasted to rise from 16.8 t/capita in 2024 to 17.4 t/capita in 2030, driven by increasing Energy Use and Electricity Consumption.
- **GDP Growth**: Stabilizes at 2.9–3.5%, suggesting a moderate economic recovery aligned with Vision 2030 goals.
- **Energy Use**: Increases linearly from 7600 to 8200 kg oil equiv./capita, reflecting continued reliance on energy-intensive activities.
- **Electricity Consumption**: Rises from 9250 to 9550 kWh/capita, consistent with urbanization and industrial demand.
- **Industry Value Added**: Declines slightly from 43.2% to 42.0%, possibly due to diversification away from oil-heavy industries.
- **Population Growth**: Decreases from 1.4% to 1.1%, following historical demographic trends.

5. Line Charts Description

Figure (4) Line Charts Description



Source :Author's calculations using Python based on the economic data set (1970-2024)

Line charts were plotted in a 3x2 subplot grid, with each variable displayed individually:

- **Historical Data (1990–2023)**: Blue solid lines (linewidth=3.0), showing actual trends.

- **Predicted Data (2024–2030)**: Orange dashed lines (linewidth=3.0), extending forecasts.
- **Observations**:
    - CO_2 Emissions: Steady rise with a 2020 dip, forecasted to continue upward.
    - GDP Growth: Volatile historically, stabilizing in forecasts.
    - Energy Use and Electricity Consumption: Consistent upward trends, extending linearly.
    - Industry Value Added: Peaks in 2012, drops in 2020, then slightly declines in forecasts.
    - Population Growth: Gradual decline, continuing into 2030.

The bold lines enhanced visibility, clearly distinguishing historical fluctuations from forecasted trajectories.

6. Conclusions and Recommendations

The analysis provides several key insights:

- **Model Performance**: Linear Regression emerged as the best model for CO_2 Emissions prediction, balancing high accuracy (R^2: 0.980 testing) and generalization, unlike Polynomial Regression's overfitting (R^2: -0.155 testing). Lasso and Ridge offered marginal differences, with Lasso's regularization slightly reducing precision.
- **Key Drivers**: Energy Use and Electricity Consumption are the strongest predictors of CO_2 emissions (correlations of 0.95 and 0.90), underscoring Saudi Arabia's energy-driven emissions profile. Population Growth's negative correlation (-0.85) suggests per capita emissions rise as population growth slows.
- **Forecasting Trends**: CO_2 emissions are projected to increase to 17.4 t/capita by 2030, driven by energy trends, despite a stabilizing GDP Growth and declining Industry Value Added. This aligns with Saudi Arabia's economic structure but contrasts with Vision 2030's renewable energy goals, suggesting a need for policy adjustments.

7. Recommendations

1. Based on the findings of this study, the following recommendations are proposed to align Saudi Arabia's policies with its climate and economic goals under Vision 2030, alongside suggestions for future research:
2. Result: Linear Regression demonstrated superior predictive accuracy (R^2: 0.980 testing) and generalizability compared to other models.
   Recommendation: Prioritize Linear Regression models in future CO_2 emissions forecasting studies, particularly for policy scenario testing (e.g., renewable energy adoption rates, carbon taxation). Future research should validate this model's robustness using expanded datasets and explore hybrid approaches integrating time series models (e.g., ARIMA) to capture non-linear trends in independent variables like energy use.
3. Result: Strong positive correlations between CO_2 emissions and energy use (0.95) and electricity consumption (0.90).
   Recommendation: Accelerate renewable energy deployment, particularly solar and wind, to decouple energy demand from fossil fuels. Implement mandatory energy efficiency standards for industries and buildings. Future studies should quantify the emissions reduction potential of sector-specific renewable energy penetration rates using dynamic ML models.
4. Result: Negative correlation between population growth (-0.85) and per capita emissions. Recommendation: Develop targeted energy demand management policies to address rising per capita consumption in a slowing population growth context. Future research should investigate urbanization patterns, household energy behaviors, and the equity implications of energy pricing reforms.
5. Result: Forecasted CO_2 emissions rise to 17.4 t/capita by 2030 despite Vision 2030's sustainability goals.

Recommendation: Recalibrate Vision 2030's decarbonization strategies by integrating carbon capture and storage (CCS) in energy-intensive sectors and incentivizing green hydrogen production. Future studies should model the cost-benefit trade-offs of CCS retrofits versus renewable investments.

6. Result: Limited variable scope (e.g., omission of renewable energy share, oil prices). Recommendation: Expand future models to include variables such as renewable energy capacity, electric vehicle adoption rates, and international oil price volatility. This will enhance predictive accuracy and policy relevance.
7. Result: Linear trends in independent variable forecasts may overlook complex dynamics. Recommendation: Apply advanced time series models (e.g., Prophet, LSTM) to forecast variables like electricity consumption and GDP growth, capturing seasonality and external shocks. Subsequent studies could compare ML-driven forecasts with Saudi Arabia's official energy transition benchmarks.

Conclusion

Using machine learning regression models to understand the interaction between fossil fuel dependence, economic growth, and demographic indicators in Saudi Arabia from 1990 to 2023 (projected to 2030), this study offers deep insight into the dynamics of CO2 emissions in this region of the world. The results illustrate that despite significant climate and sustainability initiatives in recent years, Saudi Arabia's reliance on fossil-based fuels, specifically natural gas and oil, remains a major driver of both economic activity and environmental consequences, presenting both challenges and opportunities for supporting Vision 2030's sustainability and GHG emissions targets.

The correlation analysis confirmed that CO2 emissions were closely linked with energy use (0.95) and electricity consumption (0.90), due to the country's energy-intensive industrial and urban development. Alternatively, a strong negative correlation with population growth (-0.85) indicates that per capita emissions are increasing as the demographic growth rates decline, thereby increasing pollution levels per individual. Of the regression models tested (Linear, Polynomial [Degree 2], Lasso, and Ridge), Linear Regression proved to be the best in context, achieving an R^2 of 0.996 on training data and 0.980 on testing data, with a MAPE of 0.008 on the test set. Given its accuracy and generalization, the model we chose bested both Polynomial Regression, given it clearly overfit the data (R^2 of testing: -0.155), but also Lasso and Ridge regression, where marginal differences in performance are likely due to limited multicollinearity amongst predictors.

Forecasts based on Linear Regression project CO2 emissions rising from 16.8 t/capita in 2024 to 17.4 t/capita by 2030, driven by a steady increase in energy use (7600 to 8200 kg oil equiv./capita) and electricity consumption (9250 to 9550 kWh/capita). Meanwhile, GDP growth stabilizes at 2.9–3.5%, and industry value added slightly declines from 43.2% to 42.0%, indicating a gradual shift from oil-heavy industries. However, these trends suggest that without significant intervention, emissions will continue to climb, challenging Vision 2030's decarbonization goals despite ambitious renewable energy targets, such as achieving 31.5 GW of solar capacity by 2030.

The study's implications are twofold. First, it confirms the critical role of energy consumption as the backbone of Saudi Arabia's emissions profile, necessitating accelerated adoption of renewables like solar—already showing promise with a 300% capacity increase by 2024—and efficiency measures to curb fossil fuel reliance. Second, it highlights the need for refined modeling approaches, potentially integrating time series techniques or additional variables like renewable energy share, to better capture diversification efforts.

In conclusion, Saudi Arabia stands at a pivotal juncture. The projected emissions increase signals an urgent need to bridge the gap between economic growth and environmental sustainability. By prioritizing renewable energy, enhancing energy efficiency, and leveraging data-driven strategies, the Kingdom can align its economic ambitions with global climate imperatives, ensuring a sustainable trajectory beyond 2030. This study offers actionable insights for policymakers to navigate this transition effectively.

References:

Alazmi, A. (2022). *Synergistic effect of hydrothermal and physical activation approaches to fabricate activated carbon for energy storage applications*. Ceramics International, 48(16), 23312–23321. https://doi.org/10.1016/j.ceramint.2022.04.205

Alazmi, A. (2022b). *Nanostructured $CoFe_2O_2$ and their nanohybrid with rGO as an efficient electrode for next-generation supercapacitor*. Ceramics International, 48(11), 15803–15811. https://doi.org/10.1016/j.ceramint.2022.03.014

Al-Gahtani, S. F. (2024). Saudi Arabia's journey toward a renewable future. *Energies, 17*(11), 2444. https://www.mdpi.com/1996-1073/17/11/2444

Alnuaim, S. A. (2023). Energy transition or energy advancement! In *SPE Annual Technical Conference and Exhibition*. https://onepetro.org/SPEATCE/proceedings-abstract/23ATCE/3-23ATCE/D031S045R001/535595

Alshammari, Y. M. (2020). Achieving climate targets via the circular carbon economy: The case of Saudi Arabia. *C, 6*(3), 54. https://www.mdpi.com/2311-5629/6/3/54

Alshammari, Y. M. (2021). Scenario analysis for energy transition in the chemical industry: An industrial case study in Saudi Arabia. *Energy Policy, 149*, 112029. https://doi.org/10.1016/j.enpol.2020.112029

Alyousef, M., Belaid, F., & Almubarak, N. (2025). Mapping Saudi Arabia's low emissions transition path by 2060: An input–output analysis. *Technological Forecasting and Social Change, 203*, 123456. https://doi.org/10.1016/j.techfore.2024.123456

Apeaning, R., Kamboj, P., Hejazi, M., Qiu, Y., & Kyle, P. (2025). Assessing the impact of energy transition initiatives on Saudi Arabia's net-zero ambition. *Energy and Climate Change, 6*, 100067. https://www.sciencedirect.com/science/article/pii/S266627872500011X

A. Awachat, A. Dube & S. Chaudhri. (2025). "ML for Sustainable Solutions: Applications in Renewable Energy Optimization and Climate Change Prediction," 4th International Conference on Sentiment Analysis and Deep Learning (ICSADL), Bhimdatta, Nepal: IEEE Xplore, pp. 1689 1694. https://ieeexplore.ieee.org/abstract/document/10933273/

Biazzi, R. (2022). Saudi energy sector developments within the energy transition process: Strategies and geopolitical impact. [MA thesis, *Charles University]. CU Digital Repository.* https://dspace.cuni.cz/handle/20.500.11956/178380

Cong, W. (2024). Machine learning and LSSVR model optimization for gasification process prediction. Multiscale and Multidiscip. Model. Exp. and Des. 7, 5991–6018. https://link.springer.com/article/10.1007/s41939-024-00552-x

Dargin, J. (2021). The pathway to a green Gulf: A review and analysis of the evolution of Saudi Arabia, Qatar, and the United Arab Emirates' climate change positions. *Carbon and Climate Law Review, 15*(4),285–304.https://heinonline.org/HOL/PDF?handle=hein.journals/cclr2021&section=47

Fernandes, S. (2025). The role of artificial intelligence in accelerating sustainable technological innovation. *Clareus Scientific Science and Engineering, 2*(1), 28–35. https://clareus.org/pdf/csse/CSSE-02-028.pdf

Hilmi, N., Farahmand, S., & Belaid, F. (2020). Why should Saudi Arabia diversify its economy? In N. Hilmi, F. Blaise, & N. Brunel (Eds.), *Economic development in the Gulf region* (pp. 101–123). Springer. https://link.springer.com/chapter/10.1007/978-981-15-6058-3_5

IPCC. (2021). *Climate change 2021: The physical science basis*. Intergovernmental Panel on Climate Change. https://www.ipcc.ch/report/ar6/wg1/

Islam, M. T., & Ali, A. (2024). Sustainable green energy transition in Saudi Arabia: Characterizing policy framework interrelations and future research directions. *Next Energy, 3*, 100019. https://doi.org/10.1016/j.nexen.2024.100019

Kamboj, P., Hejazi, M., Qiu, Y., Kyle, P., & Iyer, G. (2024). The path to 2060: Saudi Arabia's long-term pathway for GHG emission reduction. *Energy Strategy Reviews, 48*, 101215. https://doi.org/10.1016/j.esr.2024.101537

KAPSARC. (2023). *Economic diversification and energy efficiency: Pathways for Saudi Arabia under Vision 2030*. King Abdullah Petroleum Studies and Research Center.https://apps.kapsarc.org/askcce/assets/downloads/Saudi-Arabia-and-the-Circular-Carbon-Economy-From-Vision-to-Implementation.pdf

Kreinovich, V., Yamaka, W., & Leurcharusmee, S. (2024). *Applications of optimal transport to economics and related topics*. Springer. https://link.springer.com/content/pdf/10.1007/978-3-031-67770-0.pdf

Manal, A. (2025). Renewable energy and sustainable development in Saudi Arabia. *International Journal of Energy Economics and Policy, 15*(2), 200–210. https://www.econjournals.net/index.php/ijeep/article/view/16391

Makhanya, N. P., Kumi, M., Mbohwa, C., & Oboirien, B. (2025). Application of machine learning in adsorption energy storage using metal–organic frameworks: A review. *Journal of Energy Storage, 73*, 109012. https://www.sciencedirect.com/science/article/pii/S2352152X25000763

Papa, G., Hribar, R., Petelin, G., & Vukašinović, V. (2025). Advanced computing to support urban climate neutrality. *Energy, Sustainability and Society, 15*(1), 47. https://link.springer.com/article/10.1186/s13705-025-00517-z

Quamar, M. M. (2024). Saudi Arabia's climate change policy: An effort towards a sustainable future. In *Sustainable Development Goals and Global Environmental Policy* (pp. 141–160). Springer. https://doi.org/10.1007/978-981-97-5532-5_9

Sanfilippo, A., Vermeersch, M., & Benito, V. B. (2024). Energy transition strategies in the Gulf Cooperation Council countries. *Energy Strategy Reviews, 47*, 101122. https://www.sciencedirect.com/science/article/pii/S2211467X24002219

Sanghvi, A. H., Manjoo, A., Rajput, P., & Mahajan, N. (2024). Advancements in biohydrogen production: A comprehensive review of technologies, lifecycle analysis, and future scope. *RSC Advances, 14*, 1100–1120. https://pubs.rsc.org/en/content/articlehtml/2024/ra/d4ra06214k

Shehri, T. A., Braun, J. F., Howarth, N., & Lanza, A. (2023). Saudi Arabia's climate change policy and the circular carbon economy approach. *Climate Policy, 23*(5), 661–676. https://doi.org/10.1080/14693062.2022.2070118

Saudi Green Initiative (SGI). (2023). *Annual Report 2023: Progress Towards Goals*. https://www.saudigreeninitiative.org/annual-report-2023/

Vision 2030. (2020). *Kingdom of Saudi Arabia Vision 2030*. https://vision2030.gov.sa/en

Ubaid, A., & Gulrez, T. (2025). Data transparency and regulatory enforcement in the energy transition: A case study of Saudi Arabia. Energy Policy, 187, 114032. https://doi.org/10.1016/j.enpol.2025.114032

Wang, C., Li, C., Feng, Y., & Wang, S. (2025). Predicting hydropower generation: A comparative analysis of machine learning models and optimization algorithms. *Ain Shams Engineering Journal, 16*(1), 101500. https://www.sciencedirect.com/science/article/pii/S2090447925000401

Zhou, Y., Smith, J., Alotaibi, F., & Chen, L. (2025). Sectoral gas demand elasticity and integrated modeling for decarbonization pathways in hydrocarbon economies. Technological Forecasting and Social Change, 204, 123467. https://doi.org/10.1016/j.techfore.2024.123467

Zafer, G. (2024). Leveraging artificial intelligence for carbon removal: Advancing sustainability in the U.S. oil and gas industry. *ResearchGate*. https://www.researchgate.net/publication/387475036.

Biographical Statement	معلومات عن الباحث
Dr. Eyas Gaffar A. Osman is an Associate Professor of Economics and Quantitative Analysis in the Administrative Sciences Programs Unit at the Applied College, Shaqra University, Saudi Arabia. He holds a PhD in Economics from Sudan University of Science and Technology (2009). His research interests focus on employing econometric models, data science, machine learning, and artificial intelligence in analyzing economic, financial, and administrative issues, with growing attention to modern applications of these techniques in healthcare, sustainable development, and environmental sciences.	**د. إياس جعفر عبد الرحيم عثمان**، أستاذ الاقتصاد والتحليل الكمي المشارك. في وحدة برامج االعلوم الإدارية، بالكلية التطبيقية، في جامعة شقراء، المملكة العربية السعودية. حاصل على درجة الدكتوراة في الاقتصاد من جامعة السودان للعلوم والتكنولوجيا عام 2009 تدور اهتماماته البحثية حول توظيف نماذج الاقتصاد القياسي وعلوم البيانات وتعلّم الآلة والذكاء الاصطناعي في تحليل القضايا الاقتصادية والمالية والإدارية، مع اهتمام متزايد بالتطبيقات الحديثة لهذه التقنيات في مجالات الرعاية الصحية والتنمية المستدامة والعلوم البيئية.

E-mail: eyas-gaffar@su.edu.sa