IJEEP
INTERNATIONAL JOURNAL OF ENERGY ECONOMICS AND POLICY
EJ EconJournals
ISSN: 2146-4553

International Journal of Energy Economics and Policy

ISSN: 2146-4553

available at http: www.econjournals.com

International Journal of Energy Economics and Policy, 2025, 15(6), 158-171.

EJ EconJournals

# Studying How Macroeconomic Variables Affect Carbon Dioxide Emissions in Saudi Arabia: An ARDL Bounds Testing Approach to Cointegration

**Mohamed R. Abonazel[1]*, Ehab Ebrahim Mohamed Ebrahim[2], Abanoub M. Shafik[3]**

[1]Department of Applied Statistics and Econometrics, Faculty of Graduate Studies for Statistical Research, Cairo University, Giza, Egypt, [2]Department of Economics, College of Business, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia, [3]Department of Statistics, Faculty of Politics and Economics, Beni-Suef University, Beni-Suef, Egypt.
*Email: mabonazel@cu.edu.eg

**Received:** 12 May 2025 **Accepted:** 03 September 2025 **DOI:** https://doi.org/10.32479/ijeep.20770

**ABSTRACT**

Most countries around the world face significant environmental challenges due to their heavy reliance on non-renewable energy sources for economic growth. Therefore, a careful study of air quality in these countries is crucial, given the increasing impact of air pollution on public health and environmental sustainability. So, this study investigates the intricate relationships between carbon dioxide ($CO_2$) emissions from power industry (Energy) and three key economic drivers: manufacturing, trade, and foreign direct investment (FDI) in Saudi Arabia from 1990 to 2023. The Autoregressive Distributed Lag (ARDL) model was employed to explore both short-term and long-term dynamics. The optimal model identified was ARDL(1,4,1,4). The findings suggest that manufacturing and trade exert a negative effect on $CO_2$ emissions, whereas FDI contributes positively to the long run. The error correction model (ECM) indicates that economic adjustments occur over approximately 2 years.

**Keywords:** Trade, Manufacturing, Environmental Sustainability, Foreign Direct Investment; Nonrenewable Energy
**JEL Classifications:** C32, Q43, Q54

## 1. INTRODUCTION

Energy serves as the fundamental cornerstone of human well-being and economic prosperity worldwide, yet limited access to modern, affordable, and reliable energy services continues to represent one of the most significant barriers to social and economic development across many regions of the globe. Currently, approximately 666 million people worldwide lack access to basic electricity, with the majority concentrated in sub-Saharan Africa and rural areas of developing nations.

This energy poverty creates a vicious cycle where households are trapped in deprivation, relying on expensive and unhealthy energy sources that provide poor services while consuming significant portions of their limited income. In stark contrast, resource-rich nations like Saudi Arabia face entirely different challenges, as the Kingdom's abundant fossil fuel reserves have enabled exceptionally high energy consumption patterns, with per capita consumption reaching 7.5 tonnes of oil equivalent in 2023, including approximately 9.2 megawatt-hours of electricity per person. This energy abundance has driven economic prosperity but has also resulted in extraordinarily high carbon dioxide emissions, with Saudi Arabia emitting 17.0 tonnes of $CO_2$ per capita - nearly 5 times the global average. The Kingdom's reliance on fossil fuels for 99% of its electricity generation has contributed to its position among the world's highest per capita emitters, with total emissions increasing by 197% between 1990 and 2017. These contrasting scenarios illustrate the complex global energy landscape, where billions struggle without basic energy access while others face the environmental consequences of excessive fossil fuel consumption, highlighting the urgent need for sustainable energy solutions that can both expand access for the underserved and reduce emissions in energy-intensive economies.

Understanding the relationship between Saudi Arabia's energy usage and economic growth is essential, given the nation's reliance on petrol and oil. The Kingdom is the richest country in the Arab world because of its enormous oil and gas reserves. It has the second-largest proven crude oil reserves in the world and was the biggest oil producer in 2012. Its economy, which is mostly reliant on fossil fuels, has been shaped by the availability of its resources. About 90% of all export earnings in 2011 came from the sale of fossil fuels, according to the Organization of the Petroleum Exporting Countries (OPEC). Furthermore, Saudi Arabia's GDP is in the top twenty global economies (Alshehry and Belloumi, 2015).

However, Saudi Arabia's domestic energy use is heavily subsidized, resulting in inefficiencies, including excessive consumption and improper use of petrol and oil resources. The incentives for the quickly expanding population to embrace energy-saving measures in a variety of economic sectors have been lessened by these subsidies.

The gross domestic product (GDP) of Saudi Arabia was around PPP US$3.688/kg of oil equivalent in 2010. Nearly 6,168 kg of oil equivalent were consumed per person, which was the largest amount in the world and far more than the average of 1,851 kg for the same year. Saudi Arabia is the largest oil exporter in the world, and energy exports play a significant role in the country's economy. According to the World Bank, the country's total natural resource rents in 2010 made up 53.73% of its GDP, of which 50.46% came from oil and 3.24% from natural gas.

These high rents present a calculated chance to support and invest in renewable energy sources. Saudi Arabia's energy usage, however, is still totally reliant on fossil fuels, mainly natural gas and oil, which leads to substantial $CO_2$ emissions that have a major impact on climate change. Compared to the global mean of 4.7 metric tonnes, the nation's $CO_2$ emissions per capita in 2009 were around 16 metric tonnes. The European Union reported 7.22 metric tonnes/capita, compared to 17.27 metric tonnes in the United States (World Bank Indicator, 2012).

According to Hamieh et al. (2022), Saudi Arabia emitted approximately 569 million tonnes (Mt) of carbon dioxide ($CO_2$) from fossil fuel combustion in 2020. Notably, around 75% of these emissions, roughly 437 Mt, originated from stationary industrial sources, primarily concentrated in industrial cities and densely populated areas. Six key sectors are identified as the primary contributors to the country's industrial $CO_2$ emissions: Electricity generation, seawater desalination, oil refining, petrochemicals, cement, and iron and steel production.

Among these, electricity generation is the largest emitter, releasing approximately 184 Mt of $CO_2$ annually due to its significant reliance on fossil fuels (65% natural gas and 35% oil). Seawater desalination ranks second, emitting about 75 Mt/year, followed closely by the petrochemical industry with 73 Mt. Oil refining, cement, and iron and steel sectors contributed approximately 49 Mt, 34 Mt, and 19 Mt, respectively, in 2019 (Ye et al., 2023).

In total, 240 stationary industrial $CO_2$ sources are distributed across the kingdom. These sources form 27 emission clusters, each within a radius of <90 km. Among them, 27 clusters emit more than 1 Mt of $CO_2$ annually. The eight largest clusters, which collectively emit over 20 Mt/year, account for 77% of total stationary $CO_2$ emissions in the country. The most significant cluster is located in Jubail, with emissions of approximately 101 Mt/year, followed by Yanbu (54 Mt/year) and southern Jeddah (43 Mt/year). Other major clusters are found near Rabigh, Dammam, Riyadh, and Jizan, each representing potential sites for cost-effective carbon capture and storage (Hamieh et al., 2022).

Saudi Arabia, one of the biggest emitters of carbon dioxide ($CO_2$), is under increasing pressure to lessen its environmental impact. Testing the conservation hypothesis becomes more important in this situation since it aids in determining whether or not initiatives intended to lower energy usage will negatively impact the nation's economic development. The idea of carbon dioxide ($CO_2$), its main factors, and its negative impacts on economic performance have all been the subject of several studies. Building on this framework, the current study looks at how various macroeconomic factors affect $CO_2$ emissions in Saudi Arabia in an effort to fill a research vacuum. In the Saudi context, the main goal is to investigate the dynamic causal link between these macroeconomic variables and $CO_2$ emissions.

The format of this paper is as follows: The pertinent literature is reviewed in Section 2. The data analysis and the empirical approach are presented in Section 3. The results and policy recommendations are finally presented in Section 4.

## 2. LITERATURE REVIEW

Alkhathlan and Javid (2013) conducted a comprehensive analysis was conducted to investigate the long-run and short-run relationships between energy consumption, carbon dioxide ($CO_2$) emissions, and economic growth in Saudi Arabia using the Autoregressive Distributed Lag (ARDL) model. Their empirical findings confirmed the presence of a long-run equilibrium among the variables. The study revealed that economic growth and increased energy consumption significantly contribute to $CO_2$ emissions.

Sbia et al. (2014) investigated how FDI, clean energy use, trade openness, and economic growth affect $CO_2$ emissions in GCC nations. The dynamic ARDL technique revealed a long-term causation between economic and trade activity and $CO_2$ emissions. The report stated that resource-rich countries, such as Saudi Arabia, should improve their commitment to clean energy regulations and carbon mitigation techniques in order to preserve economic growth while tackling climate problems.

Alshehry and Belloumi (2017) investigated the causal link between fossil fuel usage, $CO_2$ emissions, and economic development. In a research specifically focused on Saudi Arabia, the ARDL bounds testing technique was applied. The study found a strong link between fossil fuel usage and increased $CO_2$ emissions over time. The report determined that Saudi Arabia's strong reliance on fossil fuels posed a significant environmental risk.

Raggad (2020) used the NARDL model to study the asymmetric impacts of energy consumption, economic growth, and financial development on $CO_2$ emissions in Saudi Arabia between 1971 and 2014. Using the RALS-LM test, which takes structural breaks into account, the study tested variable stationarity and discovered an uneven long-run cointegration relationship. Emissions were increased by both positive and negative shocks to economic development, but the former had a larger effect. Emissions also rose as a result of negative financial shocks and positive energy shocks. While economic contractions enhanced environmental quality, development in the near term resulted in environmental damage. Positive growth shocks Granger-cause $CO_2$ emissions, which in turn drive energy usage, according to the Hatemi J (2012) causality test. However, no significant causation was discovered for financial development.

Alajlan and Alreshaidi (2022) used the ARDL model and ECM framework to investigate the causal links between urbanisation, $CO_2$ emissions, and economic development in Saudi Arabia between 1985 and 2019. The limits test showed both short- and long-term links among the variables after the ADF and PP tests confirmed a mixed order of integration. The findings demonstrated that while urbanisation has little influence on emissions, but has a beneficial impact on growth, economic growth raises $CO_2$ over the long run. Additionally, the study examined the Environmental Kuznets Curve (EKC) hypothesis and used diagnostic tests to validate the validity of its conclusions.

Abonazel et al. (2023) examined the long-term connection between the trade balance and the exchange rate deficit using the autoregressive distributed lag (ARDL) model, elucidating the influence of probable factors in Egypt from 1990 to 2021. The findings showed that ARDL was the most effective model (2,1,2,3). It is discovered that the exchange rate significantly degrades the trade balance, supporting the economic theory's hypotheses. The trade balance and the money supply have a positive and strong relationship, while FDI has no long-run impact on the trade deficit. The four variables undergo short-term economic modifications (after about 1 year and 2 months or so).

Derouez et al. (2024) studied the short- and long-term effects of renewable and non-renewable energy consumption, technological progress, population growth, foreign direct investment (FDI), energy exports, energy prices, and $CO_2$ emissions on Saudi Arabia's economic growth from 1990 to 2022. Using the Autoregressive Distributed Lag (ARDL) model and the Vector Error Correction Model (VECM) Granger causality framework, the study found that non-renewable and renewable energy sources, population, FDI, energy exports, and energy prices all have a significant positive impact on Saudi Arabia's economic growth. In contrast, technical innovation and $CO_2$ emissions were proven to negatively affect growth. Furthermore, the VECM causality findings revealed four bidirectional causal links between economic growth and nonrenewable energy usage, FDI, energy exports, and energy pricing.

Daly et al. (2024) used the ARDL bounds testing approach to examine the short- and long-term relationships between $CO_2$ emissions, economic growth, and different forms of energy consumption from 1970 to 2021 in Saudi Arabia, including liquefied petroleum gas (LPGC), natural gas (NGC), and other refined products (ORPC). The investigation verified that there was a long-term cointegrating link between the variables. The empirical findings demonstrated that LPGC had a positive and statistically significant effect on GDP. Bidirectional causation was found between GDP and ORPC, whereas unidirectional causality was found between LPGC and GDP and between NGC and ORPC, according to the Toda-Yamamoto Granger causality test. Natural gas and $CO_2$ emissions were found to have a short-term favourable effect on economic growth, but they did not demonstrate a substantial causal relationship with GDP.

Azazy et al. (2024) examined the relationships of dynamics between Egypt's manufacturing (M), commerce, land, and urban population used for grain cultivation, as well as carbon dioxide ($CO_2$) emissions, from 1990 to 2021. In order to determine the most dependable method, the study benchmarked robust estimating approaches (M, MM, and S estimators) against the traditional ordinary least squares (OLS) method. This created a unique methodological framework for the autoregressive distributed lag (ARDL) model. ARDL(1, 0, 2, 1, 3) was determined to be the most suitable specification by applying the ARDL model to evaluate both short- and long-term associations. Empirical results showed that while urban population expansion and agricultural land usage for cereal production contributed positively to emission levels, manufacturing and commerce had a long-term moderating influence on $CO_2$ emissions. Furthermore, it was shown that $CO_2$ emissions were considerably impacted by the lag effects of land and trade factors. The system should reach equilibrium in around 25 months, according to the error correction model (ECM). The S-estimator outperformed the conventional OLS strategy among the estimating techniques used, producing the smallest values for the Bayesian information criterion (BIC) and the Akaike information criterion (AIC).

Sallam et al. (2025) examined the dynamic causal link between Egypt's inflation rate from 1990 to 2022 and four key economic factors: The money supply (M), foreign direct investment (FDI), exchange rate (EX), and balance of trade (BOT) variables. Given the dynamic nature of the interaction between these variables, the ARDL approach was employed to test this hypothesis both in the long and short run. Them findings showed that ARDL (2,2,2,2,2) was the best model to fit the data. Moreover, the inflation rate is significantly impacted negatively by the EX, FDI, and BOT factors. Meanwhile, long-term inflation has a positive and strong relationship with the money supply.

Ben-Ahmed and Amamou (2025) investigated the long-term correlations between Saudi Arabia's 1990-2019 carbon dioxide

($CO_2$) emissions, renewable energy consumption (RENC), and non-conventional energy resources (NCER). The investigation started by utilising the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests to examine the time series' stationarity features. The results showed that the variables had varying degrees of integration. Consequently, the cointegration dynamics were investigated using the Autoregressive Distributed Lag (ARDL) model, where $CO_2$ emissions were considered the dependent variable. Three different lag selection criteria were used in the study, and causality directions and their statistical significance within the ARDL framework were also examined. To make sure the results were dependable and broadly applicable, robustness checks were carried out. The results verified that there is a substantial long-term correlation between emissions linked to climate change and both RENC and NCER. In particular, a 1% increase in the use of renewable energy was linked to a 0.21% drop in $CO_2$ emissions, but a 1% increase in non-conventional energy resources resulted in an astounding 53.4% reduction in emissions.

Nadia et al. (2025) examined the long-term effects of Saudi Arabia's energy use, both renewable and non-renewable, on climate change over the years 1990-2019. The Autoregressive Distributed Lag (ARDL) method, created by Pesaran et al. (2001), was used in the study to investigate if a long-term equilibrium connection between the variables existed. Types of energy usage and indicators of climate change were shown to be significantly correlated over the long run. The country's energy structure has an impact on the environment, as evidenced by the long-term effects of both renewable and non-renewable energy sources on climatic trends.

Ebrahim et al. (2025) investigated the relationship between $CO_2$ emissions and key economic factors in Egypt using data from 1990 to 2022 and the ARDL econometric model. Their study finds that manufacturing and GDP per capita have a negative impact on $CO_2$ emissions, suggesting that economic growth and industrial development may support environmental improvement when clean technologies are adopted. In contrast, trade openness and increased land for cereal production are linked to higher emissions, highlighting the environmental cost of expanded trade and agriculture. The results show both short-term and long-term effects, with trade and land use changes influencing emissions beyond immediate periods. The study emphasizes the need for sustainable policies in trade and agriculture and supports the potential for economic growth that does not harm the environment. These findings provide important guidance for policymakers aiming to balance development and environmental protection in Egypt.

The goal of this study is to find out how trade, manufacturing, and foreign direct investment affect $CO_2$ emissions in Saudi Arabia from 1990 to 2023. This study looks at these connections in the context of the Saudi economy, which is different from previous research that looked at them in cross-country or global settings. Using the ARDL framework, the study looks at both short-term and long-term changes, giving us useful information about how economic growth and environmental outcomes are connected in a place that relies on oil.

# 3. DATA SOURCE AND THE MODEL OF METHODOLOGY

## 3.1. Source of Data

The $CO_2$ data for Saudi Arabia and its economic determinants (Manufacturing, Trade, and foreign direct investment) between 1990 and 2023. The World Bank's official database provided the economic variables utilised in this analysis, while the International Energy Agency (IEA) provided the information on $CO_2$ emissions.

An analysis of the relationship between manufacturing, trade, and foreign direct investment (FDI) and $CO_2$ emissions was conducted using a linear econometric model. The designated model looks like this:

$$CO_2 = \beta_0 + \beta_1 M + \beta_2 Trade + \beta_3 FDI + e \quad (1)$$

$CO_2$ emissions are the dependent variable in this formulation, while manufacturing (M), trade, and foreign direct investment (FDI) are the explanatory variables. The calculated coefficients that quantify the strength and direction of each variable's impact on $CO_2$ emissions are denoted by the parameters $\beta_1$, $\beta_2$, and $\beta_3$. The error component, represented by *e*, captures the variance that the model is unable to explain, whereas $\beta_0$ is the intercept term.

Important information about the nature of these associations may be gleaned from the computed coefficients' sign and size. A positive coefficient suggests a direct association between rising $CO_2$ emissions and rising levels of the related independent variable. A negative coefficient, on the other hand, suggests an inverse connection in which $CO_2$ emissions decrease as the independent variable rises.

## 3.2. Methodology

For econometric analysis of dynamic single-equation regressions, the Autoregressive Distributed Lag (ARDL) model is a crucial tool. For investigating long-term correlations between variables that are integrated of different orders, namely I(0) and I(1), it is especially well-suited. The adaptability of the ARDL technique is one of its main advantages; it starts with a thorough dynamic specification and then applies a series of linear and nonlinear restrictions to the model to make it more concise and understandable. Unlike earlier cointegration techniques, the ARDL framework does not require all variables to be integrated of the same order, allowing for a mix of stationary and first-order integrated series. Furthermore, the ARDL model can handle any autocorrelation and endogeneity problems in the dataset by including both current and lagged values of the explanatory variables in addition to lagged values of the dependent variable (Ghouse, 2018).

Additionally, ARDL offers a single framework that encompasses both the long-term equilibrium and the short-term dynamics in a single calculation. Its structure, which integrates temporal dynamics through lags and differences, enhances model specification and helps minimise issues such as omitted variable bias or incorrect functional form. As noted by Ebrahim et al. (2025), this makes the ARDL methodology a robust and interpretable approach for empirical time series modelling.

Let $p$ is the number of lags in the dependent variable ($y$) and $q$ is the number of lags of the independent variable , then the general form of the ARDL ($p$, $q_1$, $q_2$,...,$q_k$) model for $k$ independent variables is given by the following equation:

$$y_t = \beta_0 + \sum_{i=1}^{p} \beta_i y_{t-i} + \sum_{j=1}^{k} \sum_{h=0}^{q_j} \alpha_{j,h} x_{j,t-h} + e_t \qquad t = 1,2,...,T \tag{2}$$

Where $e_t$ is the error term and $\beta_0$, $\beta_i$, $a_{j,h}$ are the parameters of the model.

The ARDL ($p$, $q_1$, $q_2$,...,$q_k$) model in Eq.(2) is based on a set of fundamental statistical assumptions. One key condition is that the model must be linear in its parameters, which means the relationship between the dependent variable and the independent variables is expressed using linear coefficients. This ensures that the estimated results reflect the true nature of the relationship without systematic bias in the residuals. Another important assumption is that the error term has a constant mean equal to zero, which helps maintain the accuracy and neutrality of the estimates. Additionally, the variance of the error term should remain constant over time, a condition known as homoscedasticity, which supports the reliability and efficiency of the model's estimators. It is also assumed that the error terms are not correlated with each other at different periods, meaning there is no autocorrelation. This is necessary to avoid distortions in the estimation process. Furthermore, the error term must be independent of the explanatory variables at all time points. This final condition supports the assumption of exogeneity, which is essential for ensuring that the relationships identified by the model are not driven by feedback effects or omitted variable bias. The notion that estimates is an appropriate estimating approach is supported by the premise that it corresponds to the explicit method of least squares (Hill et al., 2011).

### 3.2.1. Estimation of the ARDL model

Pesaran (2015) states that the OLS method can estimate the ARDL (p, q) model in Eq.(2). The use of ordinary least squares (OLS) estimation provides several benefits, including the ability to compute coefficients with a high degree of precision under classical assumptions. This is particularly important because OLS estimates are unbiased estimators of the coefficient, and the Gauss-Markov theorem states that it is the best linear unbiased estimator (BLUE) due to its guarantee of the lowest variance among all linear unbiased estimators. Especially, let,

$$\theta = (\beta_0, \beta_1, \beta_2., \beta_p, \alpha_0, \alpha_1, \alpha_2., \alpha_q)^T$$

and

$$z_t = (1, y_{t-1}, y_{t-2}., y_{t-p}, x_t, x_{t-1}., x_{t-q})^T \qquad \text{for } t = 1,2.,T$$

The OLS estimation for $\theta$ is

$$\hat{\theta} = \left( \sum_{t=1}^{T} z_t z_t^T \right)^{-1} \sum_{t=1}^{T} z_t y_t \tag{3}$$

We could rewrite the model as a matrix from:

$$Y = Z\theta + E \tag{4}$$

where

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{bmatrix}_{T\times 1} = \begin{bmatrix} 1 & y_0 & y_{-1} & \cdots & y_{1-p} & x_1 & x_0 & \cdots & x_{1-q} \\ 1 & y_1 & y_0 & \cdots & y_{2-p} & x_2 & x_1 & \cdots & x_{2-q} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & y_{T-1} & y_{T-2} & \cdots & y_{T-p} & x_T & x_{T-1} & \cdots & x_{T-q} \end{bmatrix}_{T\times(p+q+1)}$$

$$\times \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \\ \alpha_0 \\ \vdots \\ \alpha_q \end{bmatrix}_{(p+q+1)\times 1} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_T \end{bmatrix}_{T\times 1}$$

When the model is rewritten as a matrix, the OLS estimation of $\theta$ is.

$$\hat{\theta} = (Z^T Z)^{-1} Z^T Y$$

The F-test has been used to evaluate the ARDL model, and the t-test has been used to assess the significance of each explanatory variable. Autoregressive is a component in the form of adequate lags of the explanatory variables, and the model is fitted with a distributed lag since the $y_t$ may be described by the delayed values. The OLS estimate of the ARDL (p, q) model is predicated on the existence of lagged values of the dependent variable as regressors. If there are outliers in the data, the OLS will be a skewed estimate (Pesaran, 2015).

### 3.2.2. Cointegration F-bounds test

The F-Bounds test for an ARDL model is a statistical procedure used to determine whether a long-run equilibrium (cointegration) relationship exists among a set of variables, regardless of whether the regressors are integrated of order zero, I(0), or order one, I(1), but not I(2). This test was developed by Pesaran et al. (2001). The ability to ascertain if the variables in question are cointegrated, that is, whether they have a shared stochastic tendency across time, makes this method very useful in time series analysis. Researchers can determine if long-term relationships exist by comparing the calculated F-statistics to the crucial boundaries. Therefore, when investigating the long-term dynamics and interdependencies among macroeconomic variables, the limits test is a crucial diagnostic tool.

The ARDL bounds test is based on estimating an unrestricted error correction model (UECM). For a dependent variable $y$ and a regressor $x$, the model is:

$$\Delta y_t = \beta_0 + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + \sum_{j=0}^{q} \alpha_j \Delta x_{t-j} + \phi_1 y_{t-1} + \phi_2 x_{t-1} + e_t$$

The test examines the joint null hypothesis that the coefficients of the lagged level variables are zero (i.e., no long-run relationship):

$$H_0: \phi_1 = \phi_2 = 0$$

An F-statistic is computed for the joint significance of the lagged level terms. The computed F-statistic is compared to two sets of critical values (bounds):

- Lower bound (I(0)): Assumes all variables are stationary.
- Upper bound (I(1)): Assumes all variables are integrated of order one.

The rule of the decision is

- Reject $H_0$ when F > upper bound, then cointegration exists
- Fail to reject $H_0$ when F < lower bound, then no cointegration
- Inconclusive when F between bounds.

If cointegration is found, we can proceed to estimate the long-run coefficients and the associated error correction model, which quantifies the speed at which the system returns to equilibrium after a shock.

### *3.2.3. The error correlation model*

The error correction model (ECM) is a foundational tool in time series econometrics, designed to analyze and quantify the relationship between variables that are non-stationary but share a long-run equilibrium, a property known as cointegration. ECMs are particularly valuable because they allow researchers to model both the immediate, short-term adjustments and the gradual, long-term convergence of variables toward equilibrium after a shock or change in the system. At its core, ECM describes how deviations from a long-term relationship between variables are corrected over time. The model achieves this by including an error correction term, which represents the previous period's deviation from equilibrium. The coefficient on this term, often called the speed of adjustment, indicates how quickly the dependent variable responds to restore balance after a disturbance. A significant and negative coefficient means that the system corrects itself efficiently, moving back toward equilibrium in subsequent periods.

Overall, ECMs are essential for understanding dynamic systems where short-term fluctuations are inevitable but long-term stability is expected. They provide a rigorous framework for forecasting, policy analysis, and empirical research, ensuring that both the transient and persistent components of economic relationships are accurately captured.

Estimate the ECM linked to the corresponding long-term associations in order to ascertain the explanatory factors' short-term effects on the dependent variable. ECM is a technique that uses a lag polynomial for the differences to represent a linear time series. Considering the autoregressive distributed lag, or ARDL (1,1), which is expressed as follows:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \alpha_0 x_t + \alpha_1 x_{t-1} + e_t \quad (5)$$

where $\beta_0$, $\beta_1$, $\alpha_0$, $\alpha_1$ are the parameters of the model, and $e_t$ is the error term.

Eq.(5) can be rewritten by taking the first difference between the explanatory variable and the dependent variable.

$$\Delta y_t = \beta_0 + \alpha_0 \Delta x_t - (1-\beta_1)(y_{t-1} - \theta x_{t-1}) + e_t \quad (6)$$

where $\Delta y_t = y_t - y_{t-1}$ is the first difference for the dependent variable, $\Delta x_t = x_t - x_{t-1}$ is the first difference for the explanatory variable, and $\theta = \frac{\alpha_0 + \alpha_1}{1-\beta_1}$ is the slope coefficient in the long-run relationship between $(y_t, x_t)$. The equilibrium relationship is

$$\Delta y_t = \beta_0 + \alpha_0 \Delta x_t + e_t$$

and the equilibrium error is

$$(1-\beta_1)(y_{t-1} - \theta x_{t-1})$$

Illustrating how one pair of variables deviates from a specific equilibrium. The model states that an amendment $y_t$ from the prior period is made up of changes related to movement with $x_t$ along the long-run equilibrium route, which accounts for a portion $(1-\beta_1)$ of the departure from the equilibrium $(y_{t-1} - \theta x_{t-1})$. With a model in logs, this connection will be relative (Greene, 2003).

## 3.3. Descriptive Statistics

Table 1 shows the descriptive statistics that offer a preliminary understanding of the dispersion and central tendency of the primary variables being examined. Emissions of carbon dioxide ($CO_2$) have a wide range, ranging from 57.93 to 262.37, with a mean of 154.91 and a standard deviation of 69.25. This points to significant fluctuations in emissions over time, suggesting possible structural or policy-driven shifts that might impact environmental results.

With values ranging from 8.28 to 14.79 and a mean of 10.58, the statistics for the manufacturing sector (M) reveal a rather low spread. Throughout the period examined, industrial activity was comparatively constant, as indicated by the standard deviation of 1.79. On the other hand, FDI exhibits significant volatility, as indicated by a mean of 0.76, a standard deviation of 1.15,

**Table 1: Variables' descriptive statistics from 1990 to 2023**

| Variable | Abbreviation | | Mean | Median | Standard deviation | Maximum |
|---|---|---|---|---|---|---|
| $CO_2$ emissions | $CO_2$ | 57.93270 | 154.9136 | 138.2136 | 69.24687 | 262.3692 |
| Manufacturing | M | 8.282510 | 10.57986 | 9.916685 | 1.790667 | 14.78812 |
| Foreign direct investment | FDI | −1.307818 | 0.759482 | 0.520153 | 1.148954 | 3.296522 |
| TRADE | Trade | 49.71348 | 70.78728 | 67.98671 | 11.65428 | 96.10263 |

and a negative minimum value (−1.31). This discrepancy can be explained by changes in investor confidence or by the way macroeconomic policies impact capital inflows.

Lastly, a mean of 70.79 and a standard deviation of 11.65 for the trade variable indicate a considerable degree of dispersion around the mean. With a high value of 96.10, the relatively narrow interquartile range (as indicated by the median of 67.99) indicates steady trading success, but with sporadic peaks.

Overall, our descriptive results support the need for more econometric modelling to examine the connections between manufacturing, trade, FDI, and $CO_2$ emissions since such variability may have important ramifications for economic and environmental policy.

Table 2 presents the findings of the Jarque-Bera test, which was used to assess whether the data were normal. The main inference made from this data is that all variables are normally distributed because the P-value for each variable was larger than 0.05 for the Jarque-Bera test. According to the Jarque-Bera test, the data is normally distributed, and so the null hypothesis cannot be rejected. A nearly normal distribution of the data is assessed using the Jarque-Bera test. While the alternative hypothesis proposes that the distribution deviates from normality, the null hypothesis asserts that the data are roughly normally distributed.

Table 3 demonstrates the value of the matrix of correlation for the independent variables. It appears that there is a correlation between each pair of independent variables: A moderate negative relationship among the Trade and Manufacturing equal to (−0.45), a low negative relationship among the money supply and the trade balance equal to (−0.07), and a moderate positive correlation between Manufacturing and foreign direct investment equal to 0.43. Since all the relationship coefficients are <0.8, the previous relationship data shows that there is no multicollinearity problem between the independent variables.

The variance inflation factor (VIF) is a diagnostic tool used in regression analysis to detect the presence and severity of multicollinearity among independent variables. When explanatory variables exhibit strong linear correlations with one another, this is known as multicollinearity, and it can impair the accuracy of coefficient estimations and raise their standard errors. The degree to which multicollinearity increases a coefficient's variance in the regression model is gauged by the variance inflation factor (VIF). In general, a VIF value exceeding 10 may suggest problematic collinearity, although more conservative thresholds (e.g., 5) are sometimes applied depending on the context. By evaluating VIF values, researchers may ascertain whether multicollinearity is impacting the precision of the regression results and can think

**Table 2: Jarque-Bera test of all variables in this study**

| Variable | Jarque-Bera | Probability |
|---|---|---|
| $CO_2$ | 3.457085 | 0.177543 |
| M | 4.643858 | 0.098084 |
| FDI | 1.828947 | 0.400727 |
| TRADE | 1.901706 | 0.386411 |

**Table 3: Independent variables correlation matrix**

| Correlation | M | FDI | TRADE |
|---|---|---|---|
| M | 1 | | |
| FDI | 0.4333 | 1 | |
| TRADE | −0.452353 | −0.07015 | 1 |

about suitable corrective actions, including changing the model structure by removing or combining certain variables. Ensuring low VIF values enhances the reliability and interpretability of the estimated coefficients.

Table 4 shows that, according to the VIF results, multicollinearity is not a major problem in the present model. All centered VIF values are well below the commonly accepted threshold of 10, suggesting that none of the explanatory variables, manufacturing, FDI, or trade, are highly correlated with one another in a way that would distort the regression estimates. The centered VIF for manufacturing is the highest at approximately 1.58, while FDI and trade show even lower values of 1.26 and 1.29, respectively. These results suggest that extreme multicollinearity has little effect on the model, enabling more trustworthy interpretation of the predicted coefficients.

It is also worth noting that the uncentered VIF values appear large, especially for the constant term, but these are not typically used to assess multicollinearity, as they include the intercept and do not reflect relationships solely among the independent variables.

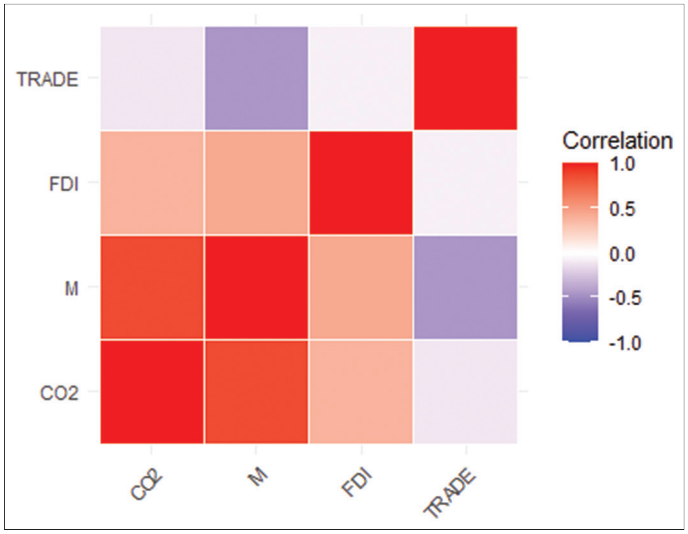
Figure 1 shows the pairwise correlation coefficients between the research variables $CO_2$ missions, manufacturing, foreign direct investment (FDI), and trade are depicted in the heatmap above. The intensity of the colour indicates the strength of the link; the colour scale goes from blue (strong negative correlation) to red (strong positive correlation). The heatmap's deep red colour indicates a strong positive association between manufacturing and $CO_2$ emissions. This suggests that rising levels of carbon emissions are directly linked to increases in industrial activity. Additionally, there is a moderately significant association between FDI and $CO_2$ emissions, indicating that FDI and ecologically intensive economic activity may be related.

The strongly negative, light purple correlation between $CO_2$ emissions and trade suggests that trade liberalisation and environmental

**Table 4: Variance inflation factor (VIF) results for the independent variables**

| Variable | Coefficient variance | Uncentered VIF | Centered VIF |
|---|---|---|---|
| M | 11.74085 | 58.36871 | 1.578974 |
| FDI | 22.79502 | 1.830270 | 1.262090 |
| TRADE | 0.226251 | 50.27938 | 1.288865 |
| C | 3473.078 | 150.0807 | NA |

**Figure 1:** Correlation matrix plot for all variables in this study




deterioration may be decoupled. Furthermore, a modest negative connection between manufacturing and trade suggests that there may be structural changes in the relationship, with greater trade not always translating into higher domestic manufacturing.

Trade and FDI have a positive correlation, which is to be anticipated in globalised economies where trade flows frequently accompany foreign investment.

## 3.4. Stationarity: Unit Root Test

The cointegration technique is inapplicable when working with integrated variables of various orders, where some series are I(1) and others are I(0), as Pesaran et al. (2001) showed. The ARDL cointegration process, however, may be applied in these situations. According to Johansen and Juselius (1990) reasoning, previous testing for unit roots is not strictly necessary, but it is still crucial to look at each series' stationarity characteristics. Since the inclusion of variables integrated of order two, I(2), may provide erroneous or false results, this step guarantees the correctness of the ARDL model. The mean, variance, and covariance structure of a time series are examples of statistical features that are said to be stationary when they do not change over time. On the other hand, nonstationary time series are stochastic processes that may be identified by structural breakdowns or unit roots. It is mostly unit roots that cause non-stationarity. Whereas its absence suggests stationarity, the existence of a unit root indicates that the time series in question is non-stationary. Dickey and Fuller (1979) developed the unit root-based stationarity test for time series. The unit root test holds that a nonstationary series (X). If a time series doesn't become stable until it has been differenced d times, it is said to be integrated of order d, or I(d). This suggests that there are d unit roots in the series in their original, undifferenced form. According to the null hypothesis, the series is non-stationary as it contains a unit root. According to the alternative theory, the series is stationary because it lacks a unit root. In time series analysis, non-stationary data can result in inaccurate statistical conclusions; therefore, determining whether a unit root exists is essential.

White disturbance in the disturbance period is the underlying assumption of the DF test. Consequently, the error term will exhibit autocorrelation if the dependent variable exhibits autocorrelation. The DF test is therefore deemed invalid as a result. Dickey and Fuller (1981) developed the DF test, which takes p lag values into account, as an improvement on the Dickey-Fuller test (ADF). The identical critical values table and null hypothesis are used in both the DF and ADF tests.

Table 5 shows that the ADF test validates Trade, foreign direct investment, and trade balance were among the variables that were included in the stationarity of I(0) at (integrated of order 0). and I(1) included variables ($CO_2$, and manufacturing) at (integrated of order 1). Figure 2 displays the variables' time series charts from 1990 to 2023, whereas Figure 3 displays the time series plots of the variables from 1990 to 2023 after accounting for discrepancies.

## 3.5. Model Selection

The ARDL model with automated lag collection was estimated using EViews. ARDL (1,4,1,4) was selected based on the lowest AIC. The model used to estimate the ARDL($p$, $q_1$,$q_2$,$q_3$) model with automated lag selection has the lowest AIC value, 6.248745.

Akaike (1973) proposed a particular metric for evaluating the model's excellent fitness. The best model is thought to have the lowest AIC. How it is administered:

$$AIC = ln(\sigma^2) + \frac{2(k+1)}{T} \quad (7)$$

In contrast to the ordinary least squares (OLS) technique, the likelihood function has the following shape when the maximum likelihood estimation (MLE) approach is used:

$$L = -T/2(1+ln(2\pi)+ln(\varepsilon^T \varepsilon/T)) \quad (8)$$

The log-likelihood function value (L), sample size (T), and model residuals (ε) are all represented in this equation. The following formula may be used to calculate the Akaike Information Criterion (AIC) using this likelihood expression:

The formula for

$$AIC = -2L + 2k \quad (9)$$

where K is the number of parameters that the model estimates. Since this version of the AIC directly considers the value of the log-likelihood function from Eq.(8), it is especially helpful when the model is estimated using MLE.

The best ARDL model chosen by Table 6 and Figure 4 is ARDL (1,4,1,4). The dependent variable ($CO_2$) and FDI will be shortened

**Table 5: Results of augmented dickey-fuller test**

| Model | Statistic | $CO_2$ | M | FDI | TRADE |
|---|---|---|---|---|---|
| | | **I. Original data** | | | |
| With intercept | t-value | 0.1913 | 1.0255 | −4.1354 | −3.478 |
| | P-value | 0.9679 | 0.9957 | 0.0029 | 0.0174 |
| | Significant | NSig. | NSig. | *** | ** |
| With intercept and trend | t-value | −2.9189 | −1.6821 | −4.9329 | −1.7392 |
| | P-value | 0.1715 | 0.7366 | 0.0019 | 0.7100 |
| | Significant | NSig. | NSig. | *** | NSig. |
| Without intercept and trend | t-value | 4.9268 | 2.5900 | −3.2291 | −0.493 |
| | P-value | 1.0000 | 0.9967 | 0.0021 | 0.4949 |
| | Significant | NSig. | NSig. | *** | NSig. |
| | | **II. Data after taking the first difference** | | | |
| **Model** | **Statistic** | **d ($CO_2$)** | **d (M)** | **d (FDI)** | **d (TRADE)** |
| With intercept | t-value | −4.9491 | −4.6596 | −4.7763 | −4.4243 |
| | P-value | 0.0003 | 0.0009 | 0.0006 | 0.0014 |
| | Significant | *** | *** | *** | *** |
| With intercept and trend | t-value | −4.8882 | −4.9686 | −4.1038 | −4.4456 |
| | P-value | 0.0022 | 0.0021 | 0.0174 | 0.0068 |
| | Significant | *** | *** | ** | *** |
| Without intercept and trend | t-value | −0.7634 | −4.6521 | −4.7969 | −4.4796 |
| | P-value | 0.3768 | 0.0000 | 0.0000 | 0.0001 |
| | Significant | NSig. | *** | *** | *** |
| III. Integrated order | | I (1) | I (1) | I (0) | I (0) |

NSig. means Not significant, ***means significant at 0.01, and **means significant at 0.05

**Figure 2:** Variables' time series plots from 1990 to 2023

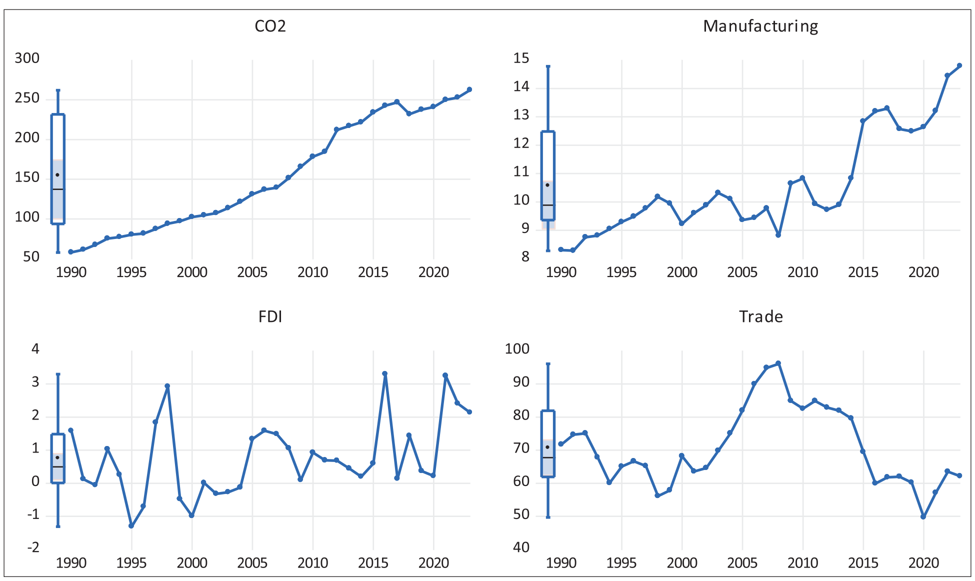



**Figure 3:** Time series plots of the first-differenced variables (1990-2023)

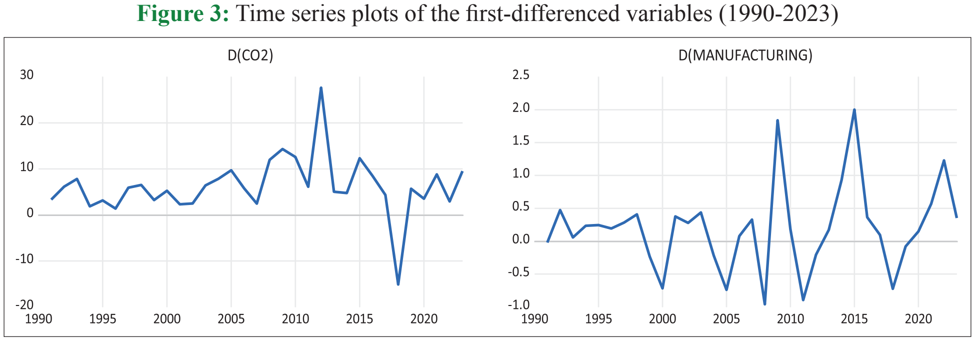
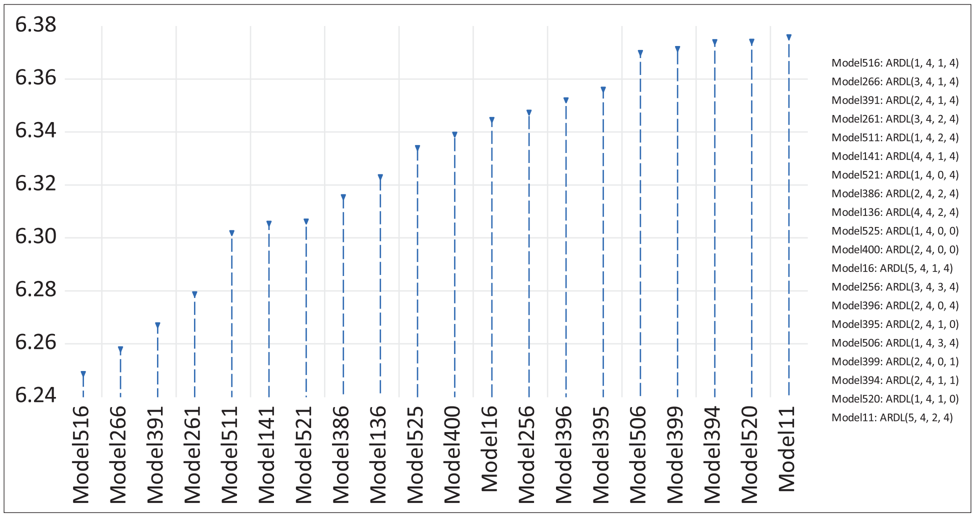


**Figure 4:** Akaike information criteria for the best 20 models




to Lag 1, while manufacturing and trade variables will be shortened to Lag 4. Then the equation of the ARDL model for this data will be:

$$CO_{2t} = \beta_0 + \beta_1 CO_{2t-1} + \sum_{h=0}^{4} \alpha_{1,h} M_{t-h}$$
$$+\sum_{h=0}^{1} \alpha_{2,h} FDI_{t-h} + \sum_{h=0}^{4} \alpha_{3,h} Trade_{t-h} + e_t;\ t = 1, 2, ..., T$$

Table 7 shows the short-run results of the ARDL model, as presented in the table reveals several key insights. The lagged dependent variable $CO_2$ (-1) carries a positive and highly significant coefficient (0.499, P < 0.01), indicating a strong degree of short-run persistence in $CO_2$ emissions. Regarding the manufacturing variable (M), while the current value and most of its lags are statistically insignificant, the first lag of manufacturing (M(-1)) exhibits a significant negative effect (P = 0.0416), suggesting that past industrial activity may contribute to a short-term reduction in emissions. As for foreign direct investment (FDI), both the current and lagged values appear to have positive but statistically insignificant effects, implying a limited short-run impact on emissions. In contrast, the trade variable displays a mixed pattern. Although several lags are insignificant, trade lag four (TRADE(-4)) is statistically significant (P = 0.0381) with a positive coefficient, indicating that changes in trade flows may have delayed short-term effects on environmental outcomes. The overall model demonstrates a strong fit, with an R-squared of 0.997 and a highly significant F-statistic, suggesting that the explanatory variables collectively account for the vast majority of variation in $CO_2$ emissions over the short run. Additionally, the akaike information criterion (AIC) value of 6.18 provides a useful benchmark for model comparison.

## 3.6. Diagnostic Tests

### *3.6.1. Normality test*

The residuals were examined for normality using the Jarque-Bera (JB) test, which was created by Jarque and Bera (1980). The residuals have a normal distribution, according to. To verify whether the residuals from the estimated model follow an approximately normal distribution, a normality test is applied. The null hypothesis assumes that the residuals are approximately

**Table 6: Model selection results based on the akaike information criterion (AIC)**

| Model | $CO_2$ | M | FDI | Trade | AIC |
|---|---|---|---|---|---|
| 1 | 1 | 4 | 1 | 4 | 6.248745 |
| 2 | 3 | 4 | 1 | 4 | 6.258085 |
| 3 | 2 | 4 | 1 | 4 | 6.267126 |
| 4 | 3 | 4 | 2 | 4 | 6.278835 |
| 5 | 1 | 4 | 2 | 4 | 6.301934 |
| 6 | 4 | 4 | 1 | 4 | 6.305537 |
| 7 | 1 | 4 | 0 | 4 | 6.306356 |
| 8 | 2 | 4 | 2 | 4 | 6.315538 |
| 9 | 4 | 4 | 2 | 4 | 6.323123 |
| 10 | 1 | 4 | 0 | 0 | 6.334096 |
| 11 | 2 | 4 | 0 | 0 | 6.339171 |
| 12 | 5 | 4 | 1 | 4 | 6.344703 |
| 13 | 3 | 4 | 3 | 4 | 6.347446 |
| 14 | 2 | 4 | 0 | 4 | 6.352127 |
| 15 | 2 | 4 | 1 | 0 | 6.356170 |
| 16 | 1 | 4 | 3 | 4 | 6.370031 |
| 17 | 2 | 4 | 0 | 1 | 6.371525 |
| 18 | 2 | 4 | 1 | 1 | 6.374091 |
| 19 | 1 | 4 | 1 | 0 | 6.374238 |
| 20 | 5 | 4 | 2 | 4 | 6.375929 |

**Table 7: The results of the ARDL (1,4,1,4) model short-run**

| Variable | Estimate | S.E | t-value | P-value |
|---|---|---|---|---|
| $CO_2$ (-1) | 0.499475 | 0.10635 | 4.696632 | 0.0003*** |
| M | −1.622109 | 1.78017 | −0.911212 | 0.3766 |
| M(-1) | −5.218305 | 2.34228 | −2.227876 | 0.0416** |
| M(-2) | −1.120219 | 2.3607 | −0.474529 | 0.642 |
| M(-3) | −3.700803 | 3.23193 | −1.145075 | 0.2701 |
| M(-4) | −4.208106 | 3.26161 | −1.290191 | 0.2165 |
| FDI | 0.654444 | 0.71315 | 0.917675 | 0.3733 |
| FDI(-1) | 1.462261 | 1.16297 | 1.257355 | 0.2278 |
| TRADE | −0.355441 | 0.19138 | −1.857283 | 0.083* |
| TRADE(-1) | −0.436978 | 0.2328 | −1.877034 | 0.0801* |
| TRADE(-2) | 0.376428 | 0.26839 | 1.402525 | 0.1811 |
| TRADE(-3) | −0.688486 | 0.45761 | −1.50454 | 0.1532 |
| TRADE(-4) | 0.612878 | 0.26956 | 2.273642 | 0.0381** |
| Intercept | 180.7006 | 74.2181 | 2.434723 | 0.0279** |
| Akaike info criterion (AIC) | | | 6.183916 | |
| $R^2$ | 0.997427 | | F-value | 415.352 |
| $R^2$ *adj* | 0.995026 | | P-value of F | 0.00001 *** |

S.E. means standard error, ***means significant at 0.01, **means significant at 0.05, and *significant at 0.1

normally distributed, which supports the validity of certain inferential procedures, such as t-tests and F-tests. On the other hand, the alternative hypothesis indicates a departure from normality. Checking the distribution of residuals is a key diagnostic step to ensure the robustness of model assumptions.

Because the JB test's P-value (0.223) is more than 0.05, Table 8 shows that the residuals are normally distributed. At level 0.05, the null hypothesis is thus not rejected. Figures 5 and 6 also show that the residuals follow a normal distribution.

Figure 5 shows A normal distribution curve is superimposed for comparison over the histogram, which shows the distribution of the regression model's residuals. The residuals appear to be broadly normally distributed based on the histogram's visual alignment with the bell-shaped normal curve. Although there are a few little asymmetries and abnormalities, they don't seem to be very noticeable. The Jarque-Bera test, which produces test statistics of 2.998 and a P = 0.223, provides statistical evidence for normality. We are unable to rule out the null hypothesis of normalcy since the P-value is higher than the standard cutoff of 0.05. This suggests that one of the fundamental presumptions of classical linear regression is met, namely that the residuals do not substantially depart from a normal distribution.

**Table 8: Jarque-Bera test for normality of the model residuals**

| Test | $\chi^2$-value | P-value |
|---|---|---|
| Jarque-Bera (JB) test of normality | 2.998 | 0.223 |

**Figure 5:** Histogram plot assessing normality of the model residuals

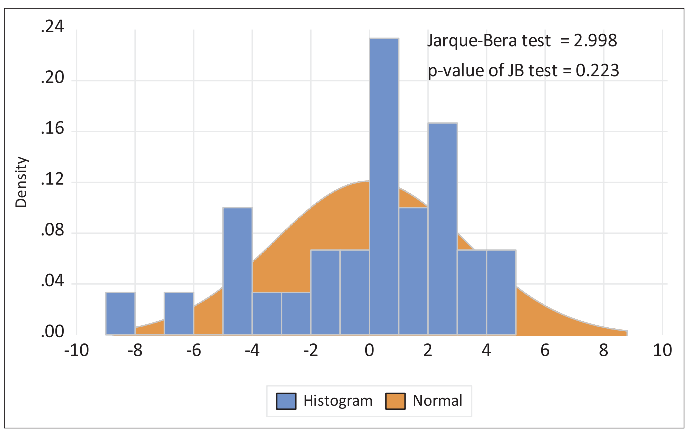



**Figure 6:** Quantile–Quantile plot for residual normality assessment

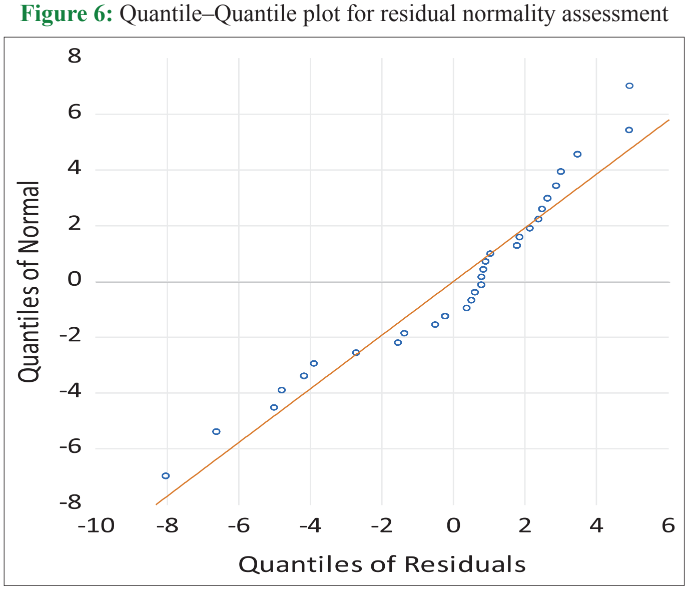



Figure 6 shows the Q-Q plot compares the quantiles of the residuals from the regression model to the theoretical quantiles of a standard normal distribution. In this chart, the points are generally aligned along the 45-degree reference line, especially in the middle range, which suggests that the residuals approximately follow a normal distribution. However, some deviations are noticeable at both tails, particularly at the lower and upper extremes. These slight departures may indicate minor non-normality in the distribution of residuals, but the overall alignment still supports the assumption of normality to a reasonable extent. Such patterns are common in practical applications and do not necessarily invalidate the model, especially when the deviations are not severe.

### *3.6.2. Serial correlation and heteroscedasticity tests*

We have used the Breusch-Godfrey LM and Breusch-Pagan-Godfrey tests to evaluate the serial correlation and heteroskedasticity of the errors.

To detect the presence of serial correlation in the residuals, the Lagrange Multiplier (LM) test is commonly used. The null hypothesis assumes that there is no serial correlation, meaning the residuals are independently distributed over time. In contrast, the alternative hypothesis suggests that serial correlation exists, which may indicate model misspecification or omitted dynamics. Identifying autocorrelation is essential to avoid biased standard errors and misleading statistical inferences.

In parallel, a heteroskedasticity test is conducted to examine whether the variance of the residuals is constant across observations. The null hypothesis states that there is no heteroskedasticity, implying homoscedastic residuals. The alternative hypothesis, however, indicates that heteroskedasticity is present, where the error variance changes with the level of the explanatory variables. Detecting this issue is critical, as heteroskedasticity can affect the efficiency of estimators and the validity of hypothesis tests.

Table 9 shows that the LM test produces a P = 0.3944, which is more than the 5% significance level, demonstrating that the residuals of the estimated model are not affected by serial correlation (Table 9). Therefore, it is not possible to rule out the null hypothesis that there is no serial association. Likewise, the White test findings for heteroscedasticity reveal a P = 0.6432, which is likewise higher than 0.05, meaning that the null hypothesis of homoscedasticity cannot be ruled out. These results corroborate the sufficiency of the model's specification by indicating that the residuals do not show autocorrelation or heteroscedasticity.

### *3.6.3. Stability test*

An essential first step before creating forecasts is evaluating the model's adequacy. As a result, the stability of the model was checked, and the residual behaviors were investigated. The

CUSUM was used to assess structural stability, as shown in Figure 7. According to the plots, both statistics stay within the 5% critical bounds for the duration of the sample, indicating that the estimated model is appropriately described and structurally stable.

### 3.7. F-bounds Test

To examine the existence of a long-term relationship among variables in a time series context, a bounds testing approach is often applied. The null hypothesis suggests that no long-run relationship exists between the variables, implying that they do not move together over time. In contrast, the alternative hypothesis indicates the presence of a long-run association, meaning that despite short-term fluctuations, the variables are connected in the long term. Establishing such a relationship is essential for determining whether a cointegration framework is appropriate for further analysis.

Table 10 shows that the results of the F-Bounds test for the ARDL(1,4,1,4) model indicate strong evidence of a long-run cointegration relationship among the variables, as the calculated F-value (7.089605) exceeds the upper bound critical value (I(1) = 5.23) at the 1% significance level for three independent variables. This means the null hypothesis of no long-run relationship is rejected at the 1% level, confirming that the variables in the model are cointegrated and justifying the estimation and interpretation of long-run coefficients

Table 11 shows the outcomes of the ARDL limits testing strategy for long-run equilibrium relationships. The results imply a clear causal relationship between greater levels of $CO_2$ emissions and two variables (FDI and industrial activity), with FDI having a positive and statistically significant long-term influence on $CO_2$ emissions. On the other hand, there is a negative long-term correlation between industrial activity and $CO_2$ emissions, suggesting that structural changes or improvements in the manufacturing sector might help reduce emissions. While trade openness has no significant long-term impact on $CO_2$ emissions.

**Figure 7:** CUSUM test for parameter stability of the ARDL(1,4,1,4) model

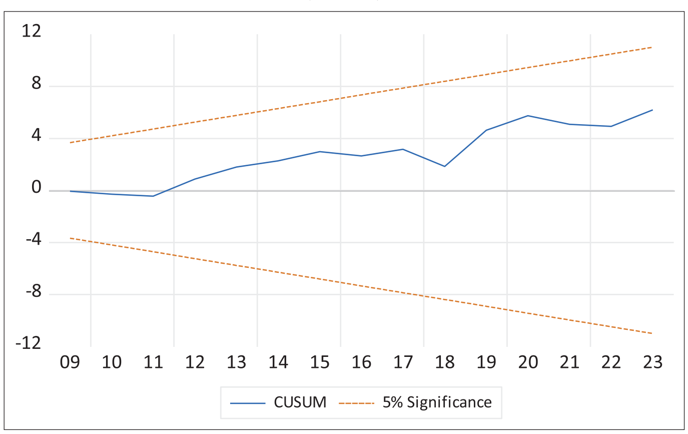



**Table 9: Results of heteroskedasticity and serial correlation tests**

| Test | $\chi^2$-value | P-value |
|---|---|---|
| LM test | 1.08035 | 0.3944 |
| Heteroskedasticity test | 0.81882 | 0.6432 |

### 3.8. Error Correction Model

The error correction term (ECT) serves as a critical indicator of the speed at which deviations from long-run equilibrium are corrected over time. It reflects the proportion of the disequilibrium from the previous period that is adjusted in the current period. A negative and statistically significant ECT coefficient is indicative of convergence toward equilibrium, confirming the existence of a stable long-run relationship among the variables. In contrast, a positive coefficient suggests divergence, signaling that the system is moving away from equilibrium. When the ECT coefficient is equal to 1, it implies that the full adjustment to equilibrium occurs within a single period, indicating an immediate correction. If the value is 0.5, approximately 50% of the disequilibrium is corrected in each period, pointing to a more gradual adjustment process. However, a coefficient of zero would imply that no adjustment is taking place, thereby undermining the evidence of any meaningful long-run relationship.

Table 12 presents the estimation results of the error correction model (ECM) derived from the ARDL (1,4,1,4) specification.

**Table 10: Bound test of ARDL (1,4,1,4) model**

| Test statistic | Value | Significant level | I (0) | I (1) |
|---|---|---|---|---|
| F-value | 7.089605 | 0.010 | 4.3 | 5.23 |
| Number of independent variables in the model | 3 | 0.025 | 3.8 | 4.68 |
| | | 0.050 | 3.38 | 4.23 |
| | | 0.100 | 2.97 | 3.74 |

**Table 11: Long-run relationship estimates from the ARDL (1,4,1,4) model**

| Variable | Estimate | S.E | t-value | P-value |
|---|---|---|---|---|
| M | −31.70581 | 9.356882 | −3.388501 | 0.0041*** |
| FDI | 4.228971 | 1.835058 | 2.304543 | 0.0359** |
| TRADE | −0.982169 | 0.577773 | −1.699922 | 0.1098 |
| Intercept | 11.26503 | 1.308954 | 8.606128 | 0.0001*** |

S.E. means standard error, ***means significant at 0.01, and **means significant at 0.05

**Table 12: Estimated short-run dynamics and error correction term**

| Variable | Estimate | S.E | t-value | P-value |
|---|---|---|---|---|
| C | 186.339 | 26.96146 | 6.911311 | 0.0001*** |
| D (M) | −1.622109 | 1.449524 | −1.119063 | 0.2807 |
| D (M(-1)) | 9.029128 | 1.98615 | 4.546046 | 0.0004*** |
| D (M(-2)) | 7.908909 | 1.619562 | 4.883364 | 0.0002*** |
| D (M(-3)) | 4.208106 | 1.625234 | 2.58923 | 0.0205** |
| D (FDI) | 0.654444 | 0.585689 | 1.117392 | 0.2814 |
| D (TRADE) | −0.355441 | 0.16943 | −2.097858 | 0.0533* |
| D (TRADE(-1)) | −0.300819 | 0.156214 | −1.925688 | 0.0733* |
| D (TRADE(-2)) | 0.075608 | 0.173697 | 0.435289 | 0.6696 |
| D (TRADE(-3)) | −0.612878 | 0.195201 | −3.13973 | 0.0067*** |
| ECT | −0.500525 | 0.074696 | −6.700808 | 0.0001*** |

S.E. means standard error, ***means significant at 0.01, **means significant at 0.05, and *significant at 0.1

The error correction term (ECT) carries a statistically significant coefficient and a negative value, confirming the presence of a stable long-run relationship among the variables. This negative sign is a clear indication of convergence, suggesting that deviations from equilibrium are systematically corrected over time. The magnitude of the ECT coefficient implies that approximately 50% of the short-run disequilibrium is adjusted within the current period, indicating that full convergence to long-run equilibrium occurs over 2 years.

## 4. CONCLUSION

This study calculated one of the dynamic causal links concerning Saudi Arabia's manufacturing, commerce, FDI, and $CO_2$ emissions from 1990 to 2023. Because the ARDL model can handle series that integrate from multiple orders, it solves the mixed stationary and nonstationary series problem. Our statistical analysis of the dataset indicated that the best ARDL model to explain the long-term equilibrium connection between $CO_2$ and its determinants (manufacturing, trade, and FDI) is ARDL (1,4,1,4). The ARDL long-run results indicated that manufacturing has a significant negative impact on $CO_2$ emissions, suggesting that increased manufacturing activity is associated with lower $CO_2$ emissions in Saudi Arabia. In contrast, FDI has a significant positive effect on $CO_2$ emissions, meaning that higher FDI leads to increased emissions over the long term. Trade does not have a statistically significant effect on $CO_2$ emissions. Given these findings, it is recommended that the Saudi government focus on policies that encourage cleaner and more efficient manufacturing practices, as this sector appears to contribute to emission reductions. Additionally, stricter environmental regulations should be imposed on foreign direct investment to ensure that it supports sustainable development and does not exacerbate environmental pressures. While trade does not show a significant long-term effect, promoting environmentally friendly trade agreements and facilitating the transfer of green technologies remain important for broader sustainability goals. Future research may enhance air quality forecasting by updated machine learning techniques, or through spatial regression modeling.

## REFERENCES


Abonazel, M.R., Shafik, A.M., Abdel-Rahman, S. (2023), Investigating the dynamic relationship between exchange rate and trade balance in Egypt: ARDL Bounds testing approach. International Journal of Applied Mathematics Computational Science and Systems Engineering, 5, 61-71.

Akaike, H. (1973), Maximum likelihood identification of Gaussian autoregressive moving average models. Biometrika, 60(2), 255-265.

Alajlan, N., Alreshaidi, A. (2022), The nexus of carbon dioxide emissions, economic growth, and urbanization in Saudi Arabia. Environmental Research Communications, 4(12), 125009.

Alkhathlan, K., Javid, M. (2013), Energy consumption, carbon emissions, and economic growth in Saudi Arabia: An aggregate and disaggregate analysis. Energy Policy, 62, 1525-1532.

Alshehry, A.S., Belloumi, M. (2015), Energy consumption, carbon dioxide emissions, and economic growth: The case of Saudi Arabia. Renewable and Sustainable Energy Reviews, 41, 237-247.

Alshehry, A.S., Belloumi, M. (2017), Study of the environmental Kuznets curve for transport carbon dioxide emissions in Saudi Arabia. Renewable and Sustainable Energy Reviews, 75, 1339-1347.

Azazy, A.R., Abonazel, M.R., Shafik, A.M., Omara, T.M., Darwish, N.M. (2024), A proposed robust regression model to study carbon dioxide emissions in Egypt. Communications in Mathematical Biology and Neuroscience, 2024, 86.

Ben-Ahmed, K., Amamou, M. (2025), Climate Change, Energy Consumption and $CO_2$ Emissions in Saudi Arabia: An ARDL Bound Testing Approach [Preprint].

Daly, H., Abdulrahman, B.M.A., Ahmed, S.A.K., Abdallah, A.E.Y., Elkarim, S.H.E.H., Sahal, M.S.G., Elshaabany, M.M. (2024), The dynamic relationships between oil products consumption and economic growth in Saudi Arabia: Using ARDL cointegration and Toda-Yamamoto Granger causality analysis. Energy Strategy Reviews, 54, 101470.

Derouez, F., Ifa, A., Aljughaiman, A.A., Haya, M.B., Lutfi, A., Alrawad, M., Bayomei, S. (2024), Energy, technology, and economic growth in Saudi Arabia: An ARDL and VECM analysis approach. Heliyon, 10(4).e26033.

Dickey, D.A., Fuller, W.A. (1979), Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association, 74(366a), 427-431.

Dickey, D.A., Fuller, W.A. (1981), Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica Journal of the Econometric Society, 49, 1057-1072.

Ebrahim, E., Abonazel, M.R., Ahmed, A.E., Abdel-Rahman, S., Albeltagy, W.A. (2025), Analysis of the economic and environmental factors affecting $CO_2$ emissions in Egypt: A proposed dynamic econometric model. International Journal of Energy Economics and Policy, 15(4), 152-165.

Ghouse, G. (2018), ARDL Model as a Remedy for Spurious Regression: Problems, Performance, and Prospects. Pakistan Institute of Development Economics, Islamabad; [Ph.D. Thesis].

Greene, W.H. (2003), Econometric Analysis. 4th ed. United States: Prentice Hall.

Hamieh, A., Rowaihy, F., Al-Juaied, M., Abo-Khatwa, A.N., Afifi, A.M., Hoteit, H. (2022), Quantification and analysis of the $CO_2$ footprint from industrial facilities in Saudi Arabia. Energy Conversion and Management X, 16, 100299.

Hatemi-j, A. (2012), Asymmetric causality tests with an application. Empirical economics, 43(1), 447-456.

Hill, R.C., Griffiths, W.E., Lim, G.C. (2011), Principles of Econometrics. 4th ed. United States: John Wiley and Sons.

Jarque, C.M., Bera, A.K. (1980), Efficient tests for normality, homoscedasticity, and serial independence of regression residuals. Economics Letters, 6(3), 255-259.

Johansen, S., Juselius, K. (1990), Maximum likelihood estimation and inference on cointegration - with applications to the demand for money supply. Oxford Bulletin of Economics and Statistics, 52(2), 169-210.

Nadia, F., Tabasam, A.H., Sarwar, R.N. (2025), Impact of climate change on energy in Saudi Arabia: An application of the ARDL bound testing approach. Metallurgical and Materials Engineering, 31(3), 326-331.

Pesaran, M.H. (2015), Time Series and Panel Data Econometrics. Oxford: Oxford University Press.

Pesaran, M.H., Shin, Y., Smith, R.J. (2001), Bounds testing approaches to the analysis of level relationships. Journal of Applied Econometrics, 16(3), 289-326.

Raggad, B. (2020), Economic development, energy consumption, financial development, and carbon dioxide emissions in Saudi Arabia: New evidence from a nonlinear and asymmetric analysis. Environmental Science and Pollution Research, 27(17), 21872-21891.

Sallam, M.A., Abonazel, M.R., Shafik, A.M. (2025), Studying the impact of macroeconomic variables on inflation rates in Egypt: An

ARDL approach. Faculty of Mediterranean Business Studies Tivat Montenegro, 21(3), 81-96.

Sbia, R., Shahbaz, M., Hamdi, H. (2014), A contribution of foreign direct investment, clean energy, trade openness, carbon emissions, and economic growth to energy demand in the UAE. Economic Modelling, 36, 191-197.

World Bank Indicator. (2012), World Bank Indicator for Development. Available from: https://data.worldbank.org

Ye, J., Afifi, A., Rowaihy, F., Baby, G., De Santiago, A., Tasianas, A., Hoteit, H. (2023), Evaluation of geological $CO_2$ storage potential in Saudi Arabian sedimentary basins. Earth Science Reviews, 244, 104539.