Policy Research Journal
ISSN (E): 3006-7030 ISSN (P) : 3006-7022 Volume 4, Issue 2, 2026

AI-ASSISTED PREDICTIVE ANALYTICS FOR ROAD-TO-RAIL FREIGHT MODAL SHIFT AND TRANSPORT DECARBONIZATION

Ali Nawaz Sanjrani[*1], Attaullah Narejo[2], Sadiq Ali Shah[1], Nouman Qadeer Soomro[3], Ameer Amanullah Sanjrani[4], Waqar Hussain[5]

*[*1,4]Department of Mechanical Engineering, MUET, Shaheed Zulfiqar Ali Bhutto Campus, Khairpur Mir's 66020, Sindh, Pakistan*
[2]School of Computer Science and Engineering, University of Electronic Science and Technology of China,
[3]Department of Software Engineering Mehran University of Engineering and Technology, SZAB Campus, Khairpur Mir's, Sindh, Pakistan
[4]Department of Computer Science, Shaikh Ayaz University of Shikarpur, Sindh, Pakistan
[5]Sindh Environmental Protection Agency, Karachi, Sindh, Pakistan

[*1]alinawaz.sanjrani@muetkhp.edu.pk

DOI: https://doi.org/10.5281/zenodo.18667446

Keywords
Road-to-rail shift; CO_2 emissions; ARIMA forecasting; machine learning; sustainable transport.

Article History
Received: 17 December 2025
Accepted: 01 February 2026
Published: 17 February 2026

Corresponding Author: *
Dr. Ali Nawaz Sanjrani

Abstract
Pakistan's transport sector faces mounting sustainability challenges due to rapid motorization and a freight system dominated by aging, diesel-powered road vehicles. Although motorcycles account for more than 80% of the national vehicle fleet, heavy-duty freight vehicles contribute nearly 60% of transport-sector CO_2 emissions, indicating a fundamental structural inefficiency. This study presents a hybrid predictive analytics framework combining an ARIMA-based time-series model with machine learning techniques to forecast transport emissions and evaluate road-to-rail freight modal shift scenarios. Using national fleet composition, fuel consumption, and emissions data for financial year 2023–24, ARIMA is employed to establish a business-as-usual emissions trajectory, while supervised machine learning models capture nonlinear relationships between freight activity, fuel use, and modal share. Multiple modal shift scenarios are simulated to quantify the emissions and energy impacts of increased rail freight penetration. The results demonstrate that even moderate shifts from road to rail can yield substantial reductions in carbon emissions and fuel demand. The proposed AI-assisted framework provides a data-driven decision-support tool for transport policy planning, supporting climate-resilient and energy-efficient freight systems in emerging economies.

1. Introduction

The global transport sector stands at a critical juncture, functioning as both a vital artery for economic development and a primary driver of environmental degradation. Globally, transportation accounts for approximately one-quarter of total CO_2 emissions [1, 2], with road transport alone responsible for the vast majority of these greenhouse gases (GHG) [2, 3]. This sector is widely classified as "hard to abate" due to its heavy reliance on fossil fuels and complex logistical constraints. In the specific context of Pakistan, these challenges are exacerbated by rapid motorization and a fundamental structural inefficiency: while motorcycles and small vehicles constitute over 80% of the national fleet, the freight sector dominated by aging, diesel powered heavy duty vehicles contributes nearly 60% of transport sector CO_2 emissions . This imbalance is compounded by a logistics network that has historically underperformed, as evidenced by Pakistan's low ranking and subsequent exclusion from recent World Bank Logistics Performance Indices [4].

To mitigate this environmental burden, the "Avoid Shift Improve" (ASI) framework identifies modal shift specifically moving freight from road to rail as a pivotal decarbonization strategy [5]. Rail transport offers substantially superior carbon cost efficiency compared to road haulage, particularly for long-distance and bulk freight [6]. Research indicates that rail freight is significantly less carbon intensive, with some studies suggesting rail emits up to 80% less CO_2 than road transport [7] and requires approximately half the energy for comparable freight work [8]. Furthermore, shifting to rail alleviates road congestion and reduces infrastructure maintenance costs, which are substantial given that logistical bottlenecks in Pakistan are estimated to cost the economy between 4% and 6% of GDP annually [4, 9, 10]
.

Methodological Framework: The Rationale for a Hybrid ARIMA-ML Approach

Navigating this transition requires robust decision-support tools. This study employs a hybrid predictive analytics framework combining an Autoregressive Integrated Moving Average (ARIMA) model with supervised machine learning. The selection of ARIMA as the foundational baseline is deliberate and addresses specific modeling requirements that other techniques cannot meet in isolation.

1.1 Why ARIMA? (Establishing the Baseline):
ARIMA is a standard statistical tool for analyzing time-series data, particularly effective when the series follows linear trends [11-13]. It decomposes data into autoregressive, integrated, and moving average components to project future points based on historical patterns [14]. In the context of this study, ARIMA is utilized to establish a "business as usual" (BAU) emissions trajectory . It is favored for its simplicity, statistical efficiency on smaller datasets, and interpretability [15]. Recent studies have successfully used ARIMA to forecast transitions in energy sectors, such as the phase-out of fossil fuels [16]. Given that Pakistan's historical transport data is often limited to annual observations rather than high-frequency big data, ARIMA minimizes the risk of overfitting that can occur with more complex models .

1.2 Why Not Deep Learning Alone? (The Limitation of LSTM):
While Deep Learning models like Long Short Term Memory (LSTM) networks are powerful for capturing long term dependencies and non-linear patterns [17], they possess significant limitations for this specific application. LSTM models are "data-hungry," typically requiring new data to improve accuracy with attention mechanism and also massive datasets to train effectively without overfitting [18, 19] . In

developing economies where freight activity data can be scarce or fragmented [20], relying solely on LSTM can lead to unreliable predictions. Furthermore, deep learning models often function as "black boxes," making it difficult for policymakers to interpret the specific drivers of a forecast , whereas ARIMA offers parametric transparency regarding trends and shocks .

1.3 The Hybrid Innovation:

To address the limitations of ARIMA regarding non-linearity, this study integrates supervised machine learning. While ARIMA captures the linear "momentum" of emissions growth, it often struggles with complex, non-linear interactions such as sudden fuel price shocks or policy interventions [21]. By feeding the residuals of the ARIMA model into machine learning algorithms, the hybrid framework captures both the linear baseline and the non-linear variations caused by external economic factors [22].

1.4 Research Novelty and Contribution

The primary innovation of this study lies in applying this hybrid analytics framework to the specific, data-scarce context of an emerging economy. The majority of existing research on freight decarbonization focuses on developed regions like the European Union, the United States, or China [23, 24] , where data availability allows for different modeling approaches. There is a distinct lack of frameworks specifically designed to simulate road to rail shifts in environments dominated by unregulated, aging diesel fleets .

By utilizing national fleet composition and fuel consumption data for the financial year 2023–24, this study fills a critical gap. Unlike static discrete choice models that analyze individual shipper behavior [25], this framework generates longitudinal national forecasts. It quantifies the emissions and energy impacts of specific modal shift scenarios, such as the Pakistan Vision 2025 goal of increasing rail freight share from 4% to 20% [24]. The resulting AI-assisted tool provides data-driven evidence for policymakers, demonstrating that even moderate shifts to rail can yield substantial decarbonization, thereby supporting climate-resilient infrastructure planning.

ARIMA-based national CO_2 forecast was combined with a road-transport modal structure to evaluate the potential impact of road-to-rail modal shift scenarios. Freight transport was identified as the primary candidate for rail substitution. Moderate (20%) and aggressive (40%) freight rail-shift scenarios were simulated using emission-factor differentials between road and rail transport, revealing substantial long-term CO_2 reduction potential

2. Methodology

2.1 Research Design and Workflow

This study develops a hybrid forecasting–scenario evaluation framework to quantify Pakistan’s transport-sector CO_2 emissions under a Business-as-Usual (BAU) trajectory and under road-to-rail freight modal shift pathways. The method integrates: (i) statistical time-series forecasting using ARIMA to produce an interpretable baseline; (ii) supervised machine learning to model non-linear variations by learning from ARIMA residuals; and (iii) scenario-based policy analysis to estimate emissions reductions for alternative rail-share targets.

2.2 Data Sources and Preparation

National-level transport and energy statistics for the year 2023–24 and earlier historical years were compiled, including total road transport activity before suitable proxies such as fuel consumption, fleet composition by vehicle class, and fuel-wise consumption (diesel/petrol) where available. In parallel, rail freight activity indicators (e.g., ton-km, traction fuel/electricity use, or published rail emission factors) were collected to quantify road–rail emissions differentials and support modal shift scenario design. As illustrated in Figure 1, the study workflow begins with a Data Input stage, integrating the national CO_2 time series and road fleet structure with road and rail emission factors, followed by Preprocessing steps that include sorting by year, imputing missing values, and selecting the target emissions series. The processed dataset is then used for Statistical Forecasting, where stationarity diagnostics (ADF test) guide differencing and ARIMA order selection (p, d, q) using information criteria, producing a baseline emissions trajectory with uncertainty bounds. Subsequently, the framework performs Transport Emissions Allocation to estimate the road transport share of national CO_2 emissions and isolate the freight-related component as the primary candidate for intervention. Building on this decomposition, Policy Analysis implements road-to-rail freight shift scenarios, and the Emission Reduction Module computes scenario-based reductions using road–rail emission factor differentials, producing a revised CO_2 trajectory under each modal shift assumption. Finally, the framework outputs Evaluation metrics, including scenario trajectories and annual CO_2 reductions ($MtCO_2$), enabling quantitative comparison of business-as-usual and road-to-rail transition pathways. Throughout the process, data were harmonized to a consistent annual time step and converted to CO_2 emissions using standard fuel-to-CO_2 conversion factors (or nationally adopted emission factors), while variables were normalized or scaled where required for machine learning.

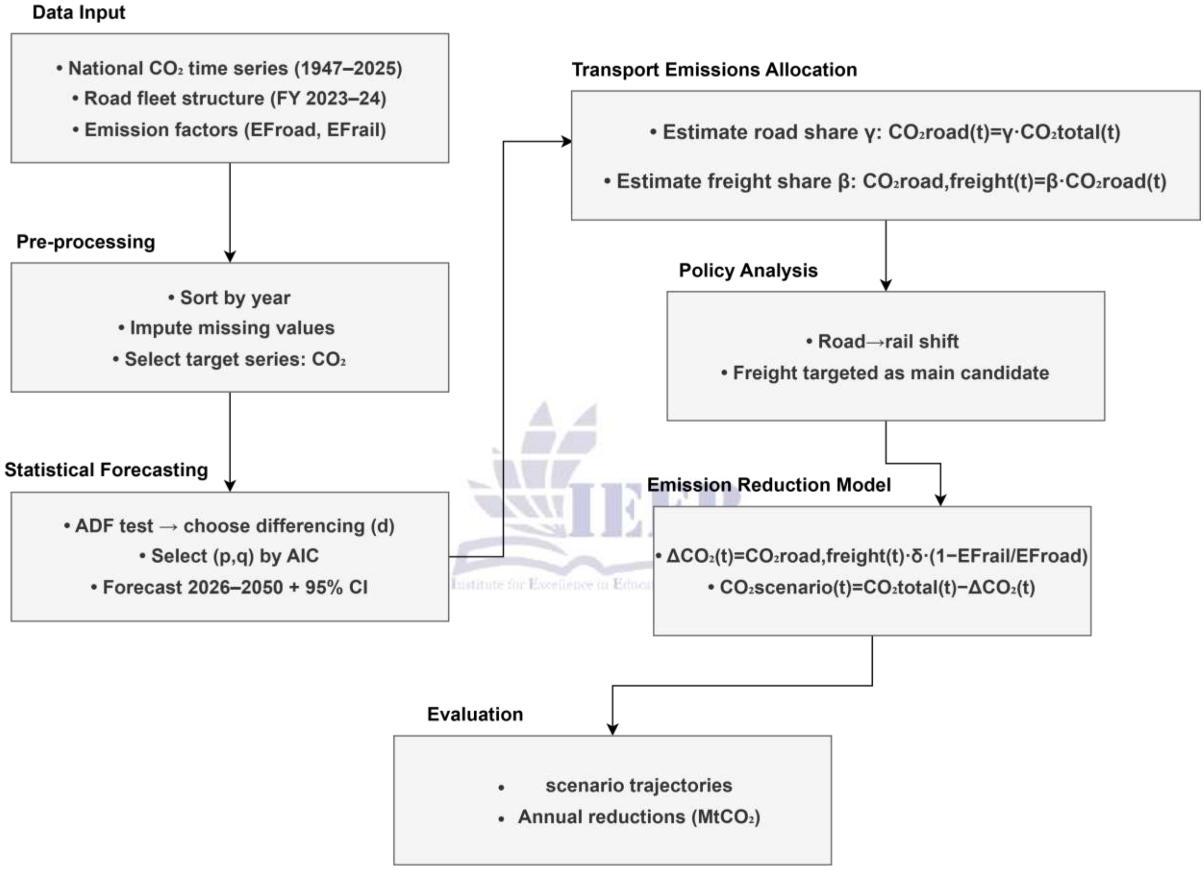

Figure 1. AI-Assisted Road-to-Rail Freight Modal Shift Forecasting Flow diagram

2.3 Baseline Emissions Estimation

Annual CO_2 emissions for the historical period were estimated as:

$$E_t = \sum_f (FC_{t,f} \times EF_f) \quad (1)$$

where $FC_{t,f}$ is fuel consumption for fuel type f in year t, and EF_f is the corresponding emission factor. If fuel consumption was unavailable, emissions were estimated using activity-based factors (e.g., vehicle-km or ton-km) with class-specific emission factors.

4. ARIMA Modeling for BAU Forecast

To generate a transparent BAU forecast, an **ARIMA** (***p,d,q***) model was fitted to the historical emissions time series E_t. The series was tested for stationarity (e.g., ADF test) and differenced until stationary (*d*). Model orders (*p,q*) were selected using information criteria (AIC/BIC) and residual diagnostics. The finalized ARIMA model produced the baseline forecast $\hat{E}_t^{ARIMA}$ over the policy horizon.

5. Hybrid ARIMA–Machine Learning Residual Learning

Because ARIMA captures primarily linear structure, a supervised ML model was trained to predict ARIMA residuals:

$$r_t = E_t - \hat{E}_t^{ARIMA} \quad (2)$$

Explanatory variables (where available) included macro and transport drivers such as fuel price indices, freight demand proxies, GDP/industrial output proxies, and fleet composition indicators. Multiple algorithms (e.g., Random Forest, Gradient Boosting, SVR) can be evaluated; the final model is selected based on validation performance. The hybrid forecast is then computed as:

$$\hat{E}_t^{Hybrid} = \hat{E}_t^{ARIMA} + \hat{r}_t^{ML} \quad (3)$$

This preserves ARIMA interpretability while incorporating non-linear corrections.

6. Scenario Modeling: Road-to-Rail Freight Modal Shift

Two policy scenarios were simulated relative to BAU:

- **Moderate shift:** 20% of freight activity shifted from road to rail
- **Aggressive shift:** 40% of freight activity shifted from road to rail

Let SSS be the shifted share (0.20 or 0.40). Emissions under a modal-shift scenario are computed using emission-factor differentials:

$$\hat{E}_t^{Scenario} = \hat{E}_t^{Hybrid} - \Delta E_t \quad (4)$$

$$\Delta E_t = S \times A_t^{Freight} \times (EF_{road} - EF_{rail}) \quad (5)$$

where $A_t^{Freight}$ is freight activity (e.g., ton-km) or an equivalent scaling proxy derived from national freight fuel use. If the dataset provides only aggregate emissions, the freight-related portion is estimated from the road modal structure (diesel freight share and heavy-duty contribution), then adjusted using the road–rail emission factor difference. Scenario outputs include annual CO_2 reductions and cumulative savings over the forecast horizon.

7. Model Evaluation and Validation

The models were evaluated using out-of-sample testing (rolling-origin or hold-out split) and error metrics such as MAE, RMSE, and MAPE. Residual diagnostics (autocorrelation checks and normality inspection) were performed to ensure ARIMA adequacy, while ML models were assessed for overfitting using cross-validation and hyperparameter tuning.

8. Outputs and Policy Indicators

The framework reports: (i) BAU emissions forecast; (ii) hybrid forecast improvement relative to ARIMA-only; (iii) emissions trajectories under 20% and 40% rail-shift scenarios; and (iv) cumulative CO_2 savings and percentage reductions. These outputs provide an evidence base to support modal-shift planning, investment prioritization, and climate policy assessment.

The proposed AREMA model estimates transport-sector CO2 emissions by integrating macro and transport drivers, road and rail emission factors, freight activity, and a defined shift scenario. The workflow begins with preprocessing, where the time-series data are prepared using stationarity checks, differencing, and model selection criteria such as AIC and BIC. A baseline "business-as-usual" ARIMA model is then developed to generate the initial emissions forecast and extract residual errors. In parallel, a hybrid residual learning module trains a machine-learning model using the same drivers to learn and predict these residual errors.

The final hybrid forecast is obtained by combining the baseline ARIMA forecast with the predicted residual correction. After forecasting, scenario modeling is applied to evaluate road-to-rail freight shifting under moderate and aggressive shift assumptions, producing scenario-based emissions projections. The model outputs include emissions trajectories over time, estimated emissions reductions and cumulative savings, percentage reduction, and overall model performance indicators, as shown in Figure 2.

3. Result and Discussion

3.1 Provincial Fleet Distribution and Dataset Overview

This study first summarizes the registered road-transport fleet across four provinces of Pakistan

(Sindh, Punjab, KPK, and Balochistan) using the compiled provincial counts reported by NTRC. The dataset covers 15 vehicle categories plus a reported total row. After cleaning the entries (removal of commas, blank spaces, and dash symbols used for missing values), the combined fleet size across all provinces equals 7,514,155 vehicles.

The provincial distribution shows a strong spatial concentration. Sindh accounts for 4,116,292 vehicles (54.78%) of the national fleet, followed by Punjab with 1,684,253 (22.41%), KPK with 1,271,082 (16.92%), and Balochistan with 442,528 (5.89%). This indicates a clear dominance of Sindh in terms of total registered vehicles, which has important implications for transport-driven emissions concentration.

3.2 National Fleet Composition and Dominant Categories

Across Pakistan, fleet composition is heavily dominated by motorcycles and scooters, totaling 4,803,871 vehicles (63.93%) of the national stock. Trucks represent the second largest category with 1,154,634 vehicles (15.37%), followed by motor cars at 526,744 (7.01%).

Secondary contributors include pickups (2.95%), tractors (2.50%), and motor rickshaws (1.90%), while all other categories contribute less than two percent individually. These results confirm that Pakistan's transport system is structurally two-wheeler intensive in terms of count, but heavy vehicles form the second largest operational block. This dual structure is highly relevant for emissions modeling, as heavy vehicles contribute disproportionately to CO_2 despite lower numerical representation.

3.3 Freight-Oriented Vehicle Stock

To isolate the logistics-relevant component of the fleet, freight-oriented categories trucks, pickups, delivery vans, oil tankers, and water tankers were aggregated. The total freight-oriented stock equals 1,550,493 vehicles, representing 20.63% of the national fleet.

Thus, approximately one-fifth of registered vehicles are directly linked to goods movement. This justifies the later modeling focus on freight activity as a major driver of transport-related CO_2 emissions and scenario-based road-to-rail shift analysis.

3.4 Provincial Structural Differences

Although motorcycles dominate nationally, provincial compositions differ significantly.

Sindh shows a freight-heavy structure, with trucks accounting for 26.35%. Punjab is highly two-wheeler centric (86.45% motorcycles). KPK displays a mixed structure, with notable shares of motor cars (12.33%), rickshaws (5.22%), and trucks (4.62%). Balochistan exhibits higher shares of tractors (17.14%) and tanker vehicles, consistent with agricultural and long-haul transport patterns. These structural variations imply regionally differentiated emission intensities and freight dynamics.

Data Consistency Verification, where validation check was conducted by summing all vehicle categories within each province and comparing the result to the reported total. The sums matched exactly for Sindh, Punjab, and Balochistan. However, for KPK, the reported total was 1,280,082, while the internally summed categories equaled 1,271,082 producing a discrepancy of 9,000 vehicles. For transparency and reproducibility, all subsequent shares and national totals use the internally consistent category-wise sum.

3.5 Historical CO_2 Emissions

Pakistan's transport-related CO_2 emissions exhibit a clear long-term upward trend from 1947 onward, with accelerated growth after the 1980s and sharp increases during 2005–2021, followed by a recent decline. This historical trajectory is illustrated in Figure 3, which shows the full emissions time series used for ARIMA modeling.

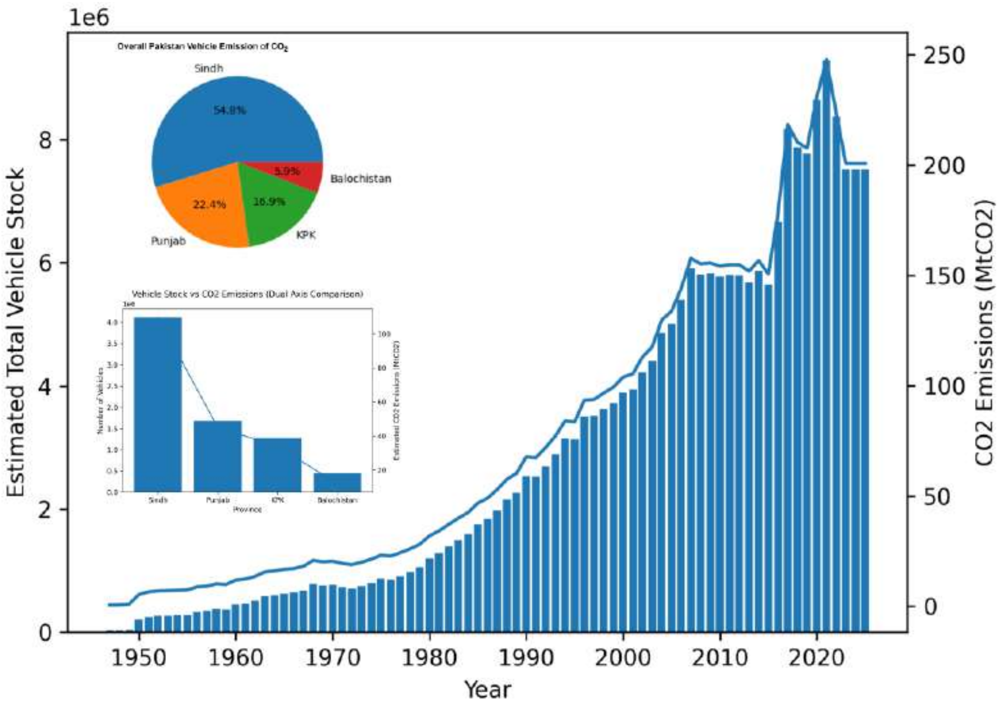

Figure 3. Integrated CO2 Emissions and Contributed Share of Vehicles in Pakistan

3.6 ARIMA Model Identification and Training Performance

3.6.1 Stationarity Assessment and Differencing Strategy

Prior to model estimation, the CO_2 emissions time series was examined for stationarity. Visual inspection of the historical trend (Figure 3) indicates a strong long-term upward trajectory with structural shifts after 2005 and volatility post-2018. Such behavior suggests non-stationarity in mean.

Preliminary differencing tests indicated that:

No differencing ($d = 0$) leaves strong trend persistence.

First differencing ($d = 1$) reduces trend but residual autocorrelation remains.

Second differencing ($d = 2$) produces a more stable mean structure.

3.6.2 Hyperparameter Setting

The hyperparameter setting is used to optimize the performance of the model and to ensure systematic and unbiased model identification, a full grid search was conducted across autoregressive orders ($p = 0–2$), differencing orders ($d = 0–2$), and moving average orders ($q = 0–3$), resulting in 36 candidate ARIMA specifications. This search space was designed to balance model flexibility with parsimony, avoiding excessive parameterization while still capturing potential autocorrelation and trend structures in the emissions series. Given the clearly trending nature of Pakistan's CO_2 emissions, differencing orders up to $d = 2$ were considered to address potential non-stationarity. Models were evaluated using both information-theoretic and predictive accuracy metrics. The Akaike Information Criterion (AIC) was employed to assess model fit while penalizing complexity, ensuring that lower AIC values reflect better trade-offs between goodness-of-fit and overfitting risk.

The hyperparameter setting of the proposed model is shown in Table 1 which present the informative optimized setting of the model with detailed description that confirm the model adequacy. Rigorously to evaluate model performance, Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and the Akaike Information Criterion (AIC) are employed, as they provide complementary insights into accuracy and model adequacy. RMSE penalizes large prediction errors and is particularly effective for assessing overall forecasting precision, while MAPE offers an intuitive, scale-independent measure of relative error that facilitates comparison across datasets and scenarios. AIC further accounts for the trade-off between goodness of fit and model complexity, discouraging overfitting by favoring parsimonious models with strong explanatory power. Together, these metrics ensure a comprehensive and reliable assessment of predictive accuracy, robustness, and generalizability.

Table 1. ARIMA Model Hyperparameter Setting

Component	Setting	Description
Autoregressive order (p)	0 – 2	Captures dependence on up to two lagged observations
Differencing order (d)	0 – 2	Accounts for trend and non-stationarity (none, first, or second differencing)
Moving Average order (q)	0 – 3	Captures up to three lagged error terms
Total candidate models	36	Full grid search across p, d, q ranges (36 Combinations)
Data split strategy	80–10–10	Chronological split (no shuffling)
Training period	1947–2009	Used for parameter estimation
Validation period	2010–2017	Used for model comparison and hyperparameter selection
Test period	2018–2025	Used for out-of-sample performance evaluation
Model selection criteria	AIC	Penalized likelihood criterion for model parsimony
Validation metric	RMSE	Measures average magnitude of forecast error
Test metric	RMSE	Assesses generalization performance
Additional accuracy metric	MAPE	Relative percentage error measure

For predictive evaluation, a chronological 80–10–10 split was implemented to preserve temporal causality. The training set (1947–2009) was used for parameter estimation. The validation set (2010–2017) guided hyperparameter selection, while the final test set (2018–2025) provided an unbiased out-of-sample performance assessment. Root Mean Square Error (RMSE) was used as the primary accuracy metric due to its sensitivity to larger errors, which is important for emissions forecasting. Mean Absolute Percentage Error (MAPE) was additionally computed to provide scale-independent interpretability. This structured evaluation framework ensures reproducibility, prevents data leakage, and allows robust comparison of ARIMA specifications before selecting the final model.

3.6.3 Training Performance Illustration

Figure 4 presents the in-sample training performance of the selected ARIMA (0,2,1) model applied to Pakistan's annual CO_2 emissions during the training period (1947–2009). The blue curve represents the observed emissions, while the orange curve shows the fitted values generated by the model. Overall, the ARIMA (0,2,1) specification demonstrates strong alignment with the historical emissions trajectory. The model successfully captures the gradual growth phase observed between the 1950s and late 1970s, the moderate acceleration during the 1980s and 1990s, and the more pronounced upward expansion in the early 2000s. The fitted series closely follows the curvature of the observed data, indicating that second-order differencing effectively removes the strong deterministic trend present in the raw emissions series.

The slight instability observed at the beginning of the series (negative spike near the first observation) is a known boundary effect associated with second differencing, where initial lagged values are limited. This artifact does not affect overall model adequacy and diminishes quickly as the time series progresses. Importantly, no systematic overestimation or underestimation is observed across the training window. Deviations between fitted and actual values appear small and randomly distributed rather than trend-driven, suggesting that the residuals are approximately white-noise during the in-sample period. This indicates that the ARIMA (0,2,1) model adequately captures both long-term structural behavior and short-run fluctuations in emissions. The strong in-sample fit, combined with superior validation and test performance (as shown in Table 2), supports the selection of ARIMA (0,2,1) as the final forecasting model for subsequent scenario and out-of-sample analysis.

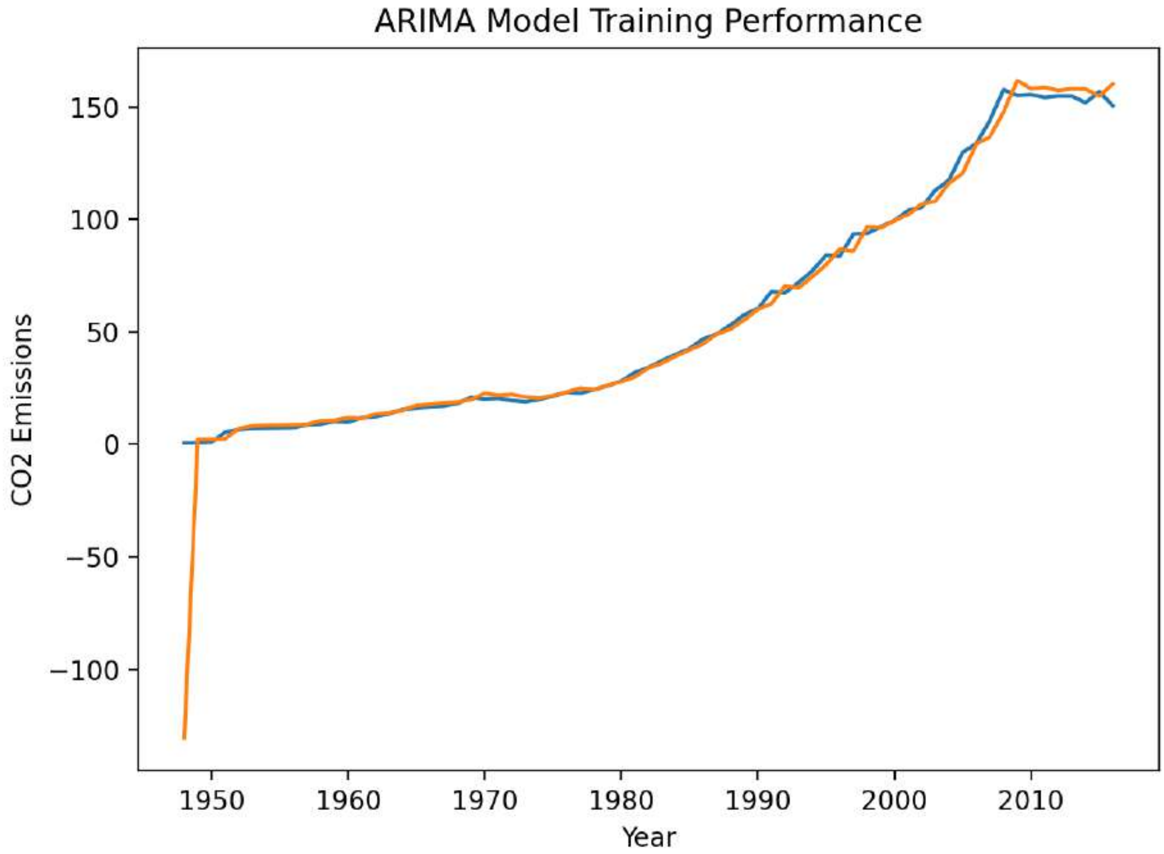

Figure 4. Training performance of the selected ARIMA (0,2,1) model for Pakistan's CO_2 emissions

3.6.4 Selected Model Performance

Multiple ARIMA specifications were evaluated using an 80–10–10 chronological split into training, validation, and test sets, and model performance was compared using AIC along with forecast accuracy metrics including RMSE and MAPE. Based on the combined evidence of strong validation accuracy and parsimonious structure, ARIMA (0,2,1) was selected as the primary baseline model. Its out-of-sample performance, presented in Figure 2 and summarized in Figure 5, shows that the model captures the curvature of Pakistan's long-term CO_2 emissions trajectory while avoiding the excessive divergence observed in lower-differencing alternatives. The second-order differencing effectively addresses the pronounced non-stationarity in the series, leading to stable forecasts that follow the overall direction of observed emissions across the validation and test horizons. Importantly, the 95 percent confidence interval expands gradually rather than explosively, indicating controlled uncertainty propagation, and most observed validation and test values remain within these bounds despite short-term volatility. Overall, the ARIMA (0,2,1) model provides a statistically sound and interpretable baseline that balances trend representation, uncertainty control, and predictive accuracy, making it suitable for subsequent scenario-based emissions analysis.

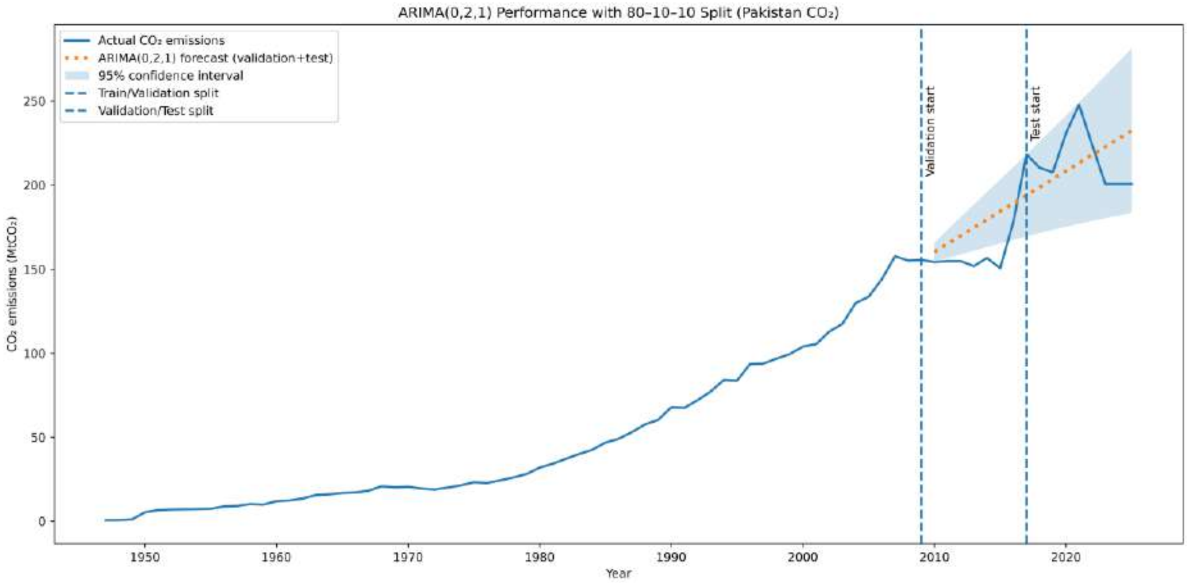

Figure 5. ARIMA Model Performance with 95% Confidence Interval

3.7 ARIMA Model Comparison

A comprehensive ARIMA model comparison was conducted using a full grid search over autoregressive orders from zero to two, differencing orders from zero to two, and moving-average orders from zero to three, resulting in thirty-six candidate specifications evaluated under an eighty–ten–ten chronological split into training, validation, and test sets. Each model was assessed using an information criterion for parsimony together with forecast accuracy metrics on the validation and test horizons. Across the full search space, models incorporating second-order differencing consistently achieved superior performance, confirming strong non-stationarity in Pakistan's CO_2 emissions series. Among all candidates, ARIMA (0,2,1) provided the best overall trade-off between accuracy and simplicity, delivering the lowest validation error with stable out-of-sample behavior, while more complex alternatives such as ARIMA (2,2,3) offered only marginal gains in fit at the cost of additional parameters. Therefore, ARIMA (0,2,1) was selected as the primary baseline model for subsequent forecasting and scenario-based emissions analysis. Figure 6 compares out-of-sample forecasts from multiple ARIMA specifications against the observed CO_2 series. The competing models produce noticeably different post-2010 trajectories: several over-project rapid growth (clear overshoot after 2020), while others under-react and remain too flat. The highlighted ARIMA (0,2,1) provides the best trade-off—its forecast follows the overall direction and curvature of recent emissions without excessive divergence. The 95% confidence band widens gradually with the forecast horizon, indicating controlled uncertainty growth and supporting the model's stability for medium-term projections.

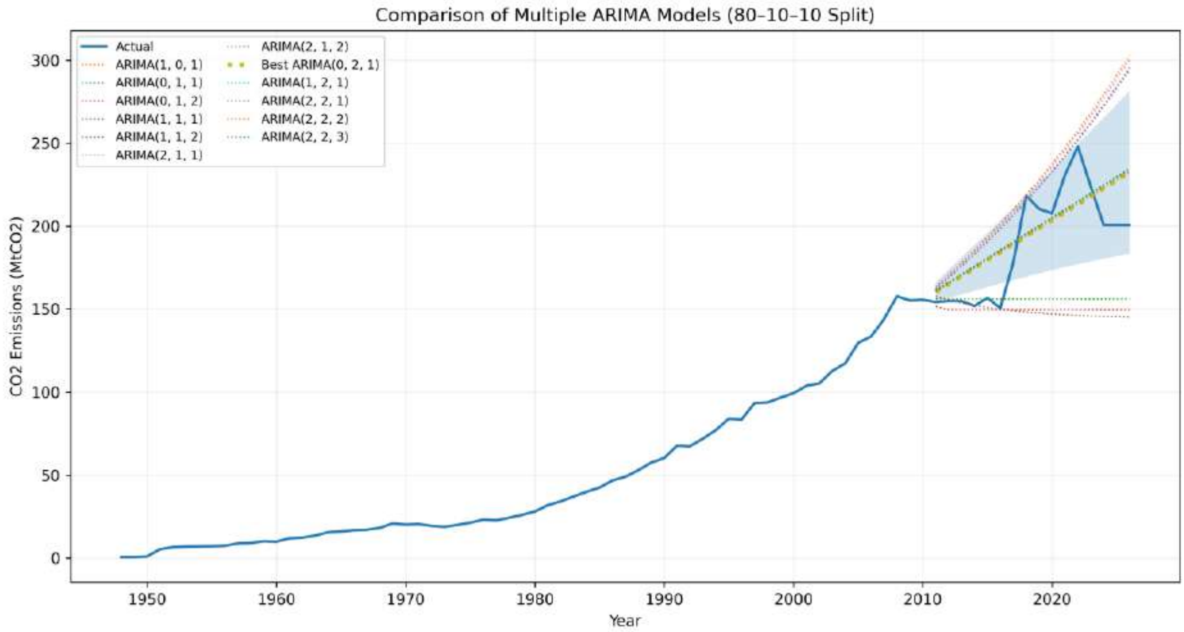

Figure 6. ARIMA best Model Comparison with various Combination

Furthermore; the Table. 2 provides a comparative evaluation of eighteen ARIMA specifications using the Akaike Information Criterion (AIC) together with validation and test accuracy indicators (RMSE and MAPE). The undifferenced baseline family ARIMA (0,0,q) performs poorest, with extremely large errors (validation RMSE above 100 and test RMSE around 163–167), confirming that the emissions series is strongly non-stationary and cannot be modeled without differencing. Introducing autoregressive and moving-average structure without differencing improves results but remains inadequate: ARIMA (1,0,1) still shows high forecast errors (validation RMSE 29.52; test RMSE 59.50), indicating that trend persistence and structural evolution are not handled properly when d is zero. First-difference models provide partial improvement in validation accuracy (ARIMA (0,1,0–2) yields validation RMSE in the 23.49–26.63 range with relatively low validation MAPE near 6–8 percent), yet they generalize poorly, as reflected by large test errors (test RMSE above 61 and MAPE above 27 percent), demonstrating that single differencing is insufficient to stabilize the series for reliable out-of-sample forecasting.

Table 2. ARIMA Model Comparison

Model	Akaike Information Criterion(AIC)	Validation		Test	
		RMSE	MAPE (%)	RMSE	MAPE (%)
ARIMA(0,0,0)	652.11	117.41	69.53	166.56	76.87
ARIMA(0,0,1)	570.38	112.46	65.33	165.51	76.37
ARIMA(0,0,2)	502.48	105.06	57.70	163.61	75.48
ARIMA(0,0,3)	455.89	102.09	55.62	162.86	75.13
ARIMA(1,0,1)	302.41	29.52	16.10	59.50	23.61
ARIMA(2,0,1)	309.17	23.20	13.09	31.78	11.84
ARIMA(1,0,3)	296.64	21.68	12.19	39.18	14.23
ARIMA(2,0,2)	306.02	25.29	14.24	38.08	14.10
ARIMA(0,1,0)	345.94	23.71	6.21	61.94	27.38
ARIMA(0,1,1)	334.20	23.49	6.33	61.38	27.11
ARIMA(0,1,2)	317.82	26.63	7.86	67.76	30.18
ARIMA(2,1,2)	298.27	27.45	15.01	54.47	21.03

Model	Akaike Information Criterion(AIC)	Validation		Test	
		RMSE	MAPE (%)	RMSE	MAPE (%)
ARIMA(0,2,0)	339.72	22.55	6.30	57.56	25.20
ARIMA(0,2,2)	293.59	21.31	12.06	23.22	9.11
ARIMA(0,2,3)	**290.92**	**19.37**	**10.55**	**22.55**	9.39
ARIMA(2,2,2)	296.64	20.28	11.27	22.69	9.25
ARIMA(2,2,3)	294.31	20.73	11.64	22.86	9.18
ARIMA(0,2,1)	**290.71**	**19.25**	**10.24**	**22.18**	**9.25**

In contrast, models with second-order differencing consistently outperform all other families, producing the lowest validation and test errors and confirming that higher-order non-stationarity dominates the emissions dynamics. Within this competitive d equals two group, ARIMA (0,2,3) attains the lowest AIC (290.92) and strong accuracy (validation RMSE 19.37; test RMSE 22.55), while ARIMA (0,2,2), ARIMA (2,2,2), and ARIMA (2,2,3) show closely comparable performance with only marginal differences in RMSE and MAPE. Notably, ARIMA (0,2,1) achieves the best overall balance of fit and forecast accuracy (AIC 290.71; validation RMSE 19.25; validation MAPE 10.24 percent; test RMSE 22.18; test MAPE 9.25 percent), indicating a parsimonious yet robust structure. Overall, the results demonstrate that second differencing is essential for capturing the underlying emissions trend behavior, and the leading d equals two models deliver stable and accurate medium-term forecasts suitable for subsequent scenario analysis.

3.8 Road-to-Rail Modal Shift Scenarios and Emissions Impact

The ARIMA (0,2,1) model (selected in the forecasting section) provides the business-as-usual baseline projection of Pakistan's transport-related CO_2 emissions (mean forecast with 95% prediction interval). The road-to-rail scenarios were then implemented as a policy overlay on this ARIMA baseline by proportionally reducing the baseline emissions to reflect shifting freight activity from road to rail. Using the study's working assumptions road-freight share of transport CO_2 = 60% and rail freight emitting about nine times less CO_2 per tonne-km than road freight as shown in Table 3, the implied system-level reductions equal 10.7% for a 20% shift and 21.3% for a 40% shift, applied consistently to the ARIMA mean forecast and its 95% bounds.

Table 3. ARIMA baseline forecast and emissions under road-to-rail shift scenarios

Year	Baseline ARIMA mean [95% PI]	20% shift mean	20% reduction	40% shift mean	40% reduction
2026	199.74 [185.03–214.45]	178.44	21.31 (10.7%)	157.13	42.61 (21.3%)
2030	199.89 [155.53–244.24]	178.56	21.32 (10.7%)	157.24	42.64 (21.3%)
2035	199.89 [135.79–263.98]	178.56	21.32 (10.7%)	157.24	42.64 (21.3%)
2040	199.89 [120.84–278.93]	178.56	21.32 (10.7%)	157.24	42.64 (21.3%)
2045	199.89 [108.30–291.47]	178.56	21.32 (10.7%)	157.24	42.64 (21.3%)
2050	199.89 [97.28–302.49]	178.56	21.32 (10.7%)	157.24	42.64 (21.3%)

Building on the selected ARIMA (0,2,1) baseline, the road-to-rail scenarios indicate substantial emissions abatement relative to the business-as-usual trajectory. Under the 20% modal shift, the projected CO_2 level decreases by 10.7%, corresponding to approximately 21.3 $MtCO_2$ per year across the forecast horizon (e.g., 2026 baseline 199.74 $MtCO_2$ reduced to 178.44 $MtCO_2$). Under the more aggressive 40% shift, emissions decrease by 21.3%, equivalent to roughly 42.6 $MtCO_2$ per year (e.g., 2026 reduced to 157.13 $MtCO_2$). While the ARIMA mean forecast remains near 200 $MtCO_2$, the 95% prediction interval widens with time, implying that the absolute abatement also varies with uncertainty; for example, by 2050 the implied reduction ranges from about 10.4 to 32.3 $MtCO_2$ for the 20% case and 20.8 to 64.5 $MtCO_2$ for the 40% case when propagating the baseline interval bounds. Overall, the scenario overlay shows that shifting freight from road to rail can deliver material and scalable CO_2 reductions, with the 40% scenario providing approximately double the abatement of the 20% case, conditional on achieving the assumed modal shift in freight activity.

3.9 Discussion

This study presents an integrated, data-driven assessment of Pakistan's road-transport structure and long-run transport-related CO_2 trajectory by combining provincial fleet statistics, time-series forecasting, and modal-shift scenario analysis. The fleet results reveal a highly uneven provincial distribution, with the largest concentration in Sindh and comparatively smaller shares in Punjab, Khyber Pakhtunkhwa, and Balochistan. Nationally, motorcycles dominate the vehicle stock, while trucks form the most significant heavy-vehicle block, which is important because heavy-duty freight vehicles typically contribute disproportionately to fuel consumption and emissions relative to their counts. These structural findings motivate the modeling strategy adopted in this work: establishing a robust baseline CO_2 projection from historical emissions dynamics and then evaluating policy overlays that target freight activity. From a forecasting perspective, the historical emissions series exhibits strong non-stationarity and changing growth regimes; the ARIMA grid search confirms that undifferenced and first-difference models generalize poorly, whereas second-difference specifications consistently achieve lower validation and test errors. Accordingly, ARIMA (0,2,1) was selected as a parsimonious baseline model with stable out-of-sample behavior, controlled uncertainty growth, and prediction bounds that capture most validation and test observations, making it suitable as a statistical benchmark for scenario comparison rather than a causal representation of underlying drivers.

Building on this ARIMA baseline, the road-to-rail modal shift scenarios translate projected emissions into policy-relevant abatement estimates using transparent emissions accounting assumptions. Under the adopted design, shifting 20 percent of road-freight activity to rail yields an estimated reduction of about 10.7 percent of transport CO_2, while a 40 percent shift yields about 21.3 percent reduction, approximately doubling the abatement as expected under proportional scaling. In addition to emissions reduction, the scenarios imply co-benefits such as reduced highway congestion from heavy vehicles, lower pavement damage, improved safety, and potential reductions in local pollutants along freight corridors; however, feasibility differs substantially between scenarios. A 20 percent shift is more realistic in the near to medium term with targeted terminal upgrades, improved service reliability, and intermodal integration, whereas a 40 percent shift typically requires major capacity expansion (rolling stock, track throughput, terminal handling), operational reforms (dispatching priority, scheduling, reliability standards), and strong last-mile logistics to avoid bottlenecks. Key limitations include the univariate nature of ARIMA (lack of explicit causal drivers), the use of aggregate transport emissions without direct passenger–freight separation, and potential variability in rail-to-road emission intensity by corridor and operating conditions; future work should therefore incorporate exogenous drivers via ARIMAX-type models, adopt bottom-up activity-based emissions accounting in tonne-km, and expand the scenario set to include rail electrification, trucking efficiency improvements, logistics optimization, and vehicle electrification pathways.

4. Conclusion

This study combined provincial vehicle fleet statistics with time-series forecasting to characterize Pakistan's transport structure and project transport-related CO_2 emissions. The descriptive results indicate strong provincial concentration of registered vehicles and a national fleet dominated by motorcycles, while trucks represent the most significant heavy-vehicle block, supporting the emphasis on freight activity in emissions mitigation analysis.

For forecasting, a systematic ARIMA grid search under an 80–10–10 split confirmed pronounced non-stationarity in the historical emissions series, with second-difference specifications consistently outperforming undifferenced and first-difference alternatives. Based on forecast accuracy and model parsimony, ARIMA (0,2,1) was selected as a robust baseline model, providing stable out-of-sample behavior with controlled uncertainty growth. Scenario overlays further suggest that shifting freight from road to rail can deliver meaningful abatement relative to the baseline, with the 40 percent modal-shift case achieving substantially greater reductions than the 20 percent case, conditional on rail capacity expansion, terminal readiness, and service reliability. Future work should strengthen both the predictive realism and policy fidelity of the framework. First, the baseline model can be extended from a univariate ARIMA to multivariate forecasting (e.g., ARIMAX or related models) by incorporating exogenous drivers such as freight activity proxies, fuel consumption, GDP, fuel prices, and policy shocks to better handle structural changes. Second, the scenario module should be upgraded from proportional reductions to a bottom-up activity-based approach that links vehicle stock, utilization (vehicle-km/ton-km), and emission factors, enabling corridor-specific road-to-rail shifts and technology pathways (e.g., rail electrification, cleaner trucking standards). Third, uncertainty analysis can be expanded through sensitivity testing of key parameters (freight share, rail-to-road emission intensity ratio, achievable shift levels) to provide robust ranges for decision-makers. Finally, integrating richer datasets (freight tonnage, logistics flows, rail throughput constraints, and disaggregated passenger vs freight emissions) will improve interpretability and support implementation-oriented decarbonization planning.

REFERENCE

[1] Falanga, A., et al., A roadmap to low-carbon freight transport: A review of EU directives and regulations. complexity, 2025. 15: p. 17.
[2] Flores-Gandur, R., J.M. Vassallo, and N. Sobrino, Assessing the Socioeconomic Impacts of an Inductive Electric Road System (ERS) for Decarbonizing Freight Transport: A Case Study for the TEN-T Corridor AP-7 in Spain. Sustainability, 2025. 17(5): p. 2283.
[3] Aryanpur, V. and F. Rogan, Decarbonising road freight transport: The role of zero-emission trucks and intangible costs. Scientific reports, 2024. 14(1): p. 2113.
[4] Economics, P.I.o.D. Pakistan Institute of Development Economics (PIDE). Available from: http://www.pide.org.pk.
[5] Picone, M., et al., Decarbonizing road transport: A systematic review of methods and models for emission impact assessment. WSEAS Transactions on Environment and Development, 2025. 21: p. 674-688.
[6] Boehm, M., M. Arnz, and J. Winter, The potential of high-speed rail freight in Europe: how is a modal shift from road to rail possible for low-density high value cargo? European Transport Research Review, 2021. 13(1): p. 4.
[7] Fale, M., et al., Evaluating Middle-Distance Passenger Transportation: A Comparison of the Time Efficiency, Sustainability, and Break-Even Distances Between High-Speed-Rail and Air Travel in Sweden. Sustainability, 2025. 17(2): p. 386.
[8] Vaccaro, R., et al., The environmental impact in terms of CO2 of a large-scale train infrastructure considering the electrification of heavy-duty road transport. iScience, 2024. 27(10).
[9] Sumbal, M.S., et al., Logistics performance system and their impact on economic corridors: a developing economy perspective. Industrial Management & Data Systems, 2024. 124(3): p. 1005-1025.

[10] Shoukat, R. Cost, Transit Time, and GHG Emissions Minimisation. Acta Tecnologia 2022 IV]; Available from: https://www.actatecnologia.eu/issues/2022/IV_2022_03_Shoukat.pdf.

[11] Yadav, S. and S. Shukla, A comparative study of ARIMA, Prophet and LSTM for time series prediction. J. Artif. Intell. Mach. Learn. Data Sci, 2022. 1(1): p. 1813-1816.

[12] Kontopoulou, V.I., et al., A review of ARIMA vs. machine learning approaches for time series forecasting in data driven networks. Future Internet, 2023. 15(8): p. 255.

[13] Jiang, C. Comparative analysis of ARIMA and deep learning models for time series prediction. in 2024 International Conference on Data Analysis and Machine Learning (DAML). IEEE. 2025.

[14] Stanislaus, E.O.O., O.U. Adaku, and O.B. Odianonsen, A comparative analysis of five time series models for co2 emissions forecasting in Port-Harcourt and its environs. Euro J Eng Tech, 2024. 12(1): p. 26.

[15] Gerding, E., et al., Proceedings of the international workshop on AI systems for the environment (AISE-2025). 2025.

[16] Liashenko, O., et al., Policy Plateau and Structural Regime Shift: Hybrid Forecasting of the EU Decarbonisation Gap Toward 2030 Targets. Sustainability, 2026. 18(2): p. 1114.

[17] Salem, K.M., et al., A Generalizable Hybrid AI-LSTM Model for Energy Consumption and Decarbonization Forecasting. Sustainability, 2025. 17(23): p. 10882.

[18] Sanjrani, A.N., et al., High-speed train bearing health assessment based on degradation stages through diagnosis and prognosis by using dual-task LSTM with attention mechanism. Quality and Reliability Engineering International, 2025. 41(5): p. 1735-1750.

[19] Kazi, M.-K. and F. Eljack, Practicality of green H2 economy for industry and maritime sector decarbonization through multiobjective optimization and RNN-LSTM model analysis. Industrial & Engineering Chemistry Research, 2022. 61(18): p. 6173-6189.

[20] Kaack, L.H., et al., Decarbonizing intraregional freight systems with a focus on modal shift. Environmental Research Letters, 2018. 13(8): p. 083001.

[21] Quaiyyum, F., K.G. Moazzem, and H.M. Preoty, Market-based Fuel Pricing Using Artificial Neural Network. Available at SSRN 5191714, 2025.

[22] Alharbi, M.H., Prediction of the stock market using LSTM, ARIMA, and hybrid of LSTM-ARIMA models. Journal of Knowledge Management Application and Practice, 2025. 7(1): p. 15-22.

[23] Jonkeren, O., K. Friso, and L. Hek, Changes in external costs and infrastructure costs due to modal shift in freight transport in North-western Europe. Journal of Shipping and Trade, 2023. 8(1): p. 24.

[24] Han, J., H. Lin, and Z. Qin, Prediction and comparison of in-vehicle co2 concentration based on arima and lstm models. Applied Sciences, 2023. 13(19): p. 10858.

[25] Agyemang, K.O. and E. Irannezhad. Uncovering the determinants of shippers' willingness to shift from road to rail freight transport. in Australasian Transport Research Forum (ATRF), 44th, 2023, Perth, Western Australia, Australia. 2023.