Journal of Research in Chemistry

E-ISSN: 2709-9423
P-ISSN: 2709-9415
Impact Factor (RJIF): 5.29
JRC 2025; 6(2): 282-286

www.chemistryjournal.net
Received: 10-08-2025
Accepted: 15-09-2025

Adebobola Ololade Agbeja
a) Department of Chemical Sciences, Faculty of Natural and Applied Sciences, Lead City University, Ibadan, Nigeria
b) Department of Sustainable Forest Management, Forestry Research Institute of Nigeria, PMB 5054, Ibadan, Nigeria

Olumuyiwa Olufisayo Ogunlaja
a) Department of Chemical Sciences, Faculty of Natural and Applied Sciences, Lead City University, Ibadan, Nigeria
b) African Centre of Excellence for Water and Environmental Research (ACEWATER), Redeemer's University, Ede, Osun State, Nigeria

Olayinka Oderinde
Department of Chemistry, Nile University of Nigeria, Plot 681, Cadastral Zone C-OO, Research & Institution Area, Jabi Airport Bypass, Abuja, FCT 900001, Nigeria

Adetola Abiola Ajayi
a) Department of Chemical Sciences, Faculty of Natural and Applied Sciences, Lead City University, Ibadan, Nigeria
b) Department of Science Laboratory Technology, School of Pure and Applied Sciences, Federal Polytechnic, PMB 50, Ilaro, Nigeria

Olubunmi Ayoni Ogundiran
a) Department of Chemical Sciences, Faculty of Natural and Applied Sciences, Lead City University, Ibadan, Nigeria
b) Department of Chemical Sciences, Faculty of Sciences, Taraba State University, PMB 1167, Jalingo, Nigeria

Khairat Abiola Olaifa
Department of Sustainable Forest Management, Forestry Research Institute of Nigeria, PMB 5054, Ibadan, Nigeria

Oyedoyin Bukola Oyedeji
Advance Level Programme Chemistry Unit, Bowen University, Iwo, Nigeria

Correspondence
Adebobola Ololade Agbeja
a) Department of Chemical Sciences, Faculty of Natural and Applied Sciences, Lead City University, Ibadan, Nigeria
b) Department of Sustainable Forest Management, Forestry Research Institute of Nigeria, PMB 5054, Ibadan, Nigeria

Modelling vehicular exhaust emissions and their environmental impacts using machine learning in Ibadan, Nigeria

Adebobola Ololade Agbeja, Olumuyiwa Olufisayo Ogunlaja, Olayinka Oderinde, Adetola Abiola Ajayi, Olubunmi Ayoni Ogundiran, Khairat Abiola Olaifa and Oyedoyin Bukola Oyedeji

DOI: https://www.doi.org/10.22271/reschem.2025.v6.i2d.237

Abstract
Ibadan faces growing air pollution due to increased vehicle use, old tokunbo cars with poor emission controls. These Diesel truck vehicles emit harmful gases, impacting urban air quality and health risks. This study investigated vehicular exhaust emissions and their environmental implications across nine major transport hubs in Ibadan, Southwest Nigeria. Using a handheld Kane 5-gas analyzer, real-time measurements were taken from 1,034 diesel-powered trucks to assess the concentrations of carbon monoxide (CO), carbon dioxide (CO_2), oxygen (O_2), and nitrogen oxides (NO_x). The study employed descriptive and inferential statistical analyses, Euro IV compliance evaluation, and machine learning classification models to analyze emission patterns and regulatory conformity. Satellite-derived atmospheric parameters from the Atmospheric Infrared Sounder (AIRS) were integrated to explore the relationship between emissions and climatic variables. Results revealed that Iwo Road recorded the highest mean CO and CO_2 concentrations (0.24% and 2.33%, respectively), while Toll-Gate exhibited the highest NO_x mean level (463.3 ppm). The Random Forest, k-Nearest Neighbors, and Neural Network models demonstrated superior predictive accuracy for emission compliance classification (AUC = 1.000). These findings highlight the substantial influence of vehicular emissions on local atmospheric dynamics and climate processes. The study underscores the urgent need for stringent emission regulations, clean transport technologies, and integrated air quality management strategies to mitigate urban air pollution and climate risks.

Keywords: Vehicular emissions, Euro IV standards, machine learning, satellite data, air quality, Ibadan, Nigeria

Introduction

High-quality air is essential for sustaining human health and overall well-being. When air purity is compromised, even brief exposure can lead to a range of adverse health outcomes, including acute conditions such as respiratory irritation and long-term chronic illnesses affecting the cardiovascular and pulmonary systems (Fussell *et al.*, 2022) [3]. Therefore, continuous monitoring of air quality is crucial to maintaining a healthy and stable environment that supports a safe and peaceful way of life on Earth. Air pollution can originate from both natural phenomena, such as dust storms, wildfires, and volcanic activity, and human activities, including industrial emissions, energy production, and vehicular traffic (WAQ, 2019) [14]. The extent to which these pollutants affect living organisms depends not only on the intensity of nearby emission sources but also on local meteorological conditions, such as wind patterns, temperature inversions, and atmospheric ventilation, which influence the dispersal or accumulation of pollutants (Zhang *et al.*, 2021) [12].

Rapid urbanization, industrial growth, and increasing reliance on motorized transportation have significantly intensified air pollution in many African cities. Diesel-powered vehicles, in particular, dominate both freight and passenger transport in these urban centers, releasing substantial quantities of carbon monoxide (CO), carbon dioxide (CO_2), nitrogen oxides (NO_x), and particulate matter. These emissions collectively degrade air quality and contribute to climate instability (OECD, 2014). In developing countries like Nigeria, the problem is exacerbated by weak enforcement of emission standards, aging and poorly maintained vehicle fleets, and limited adherence to pollution control measures (Kojima & Lovei, 2001) [6]. Ibadan, one of the largest metropolitan areas in Sub-Saharan Africa, illustrates these challenges: unregulated vehicular emissions, chronic traffic congestion, and inadequate maintenance practices have caused rising levels of ambient air pollutants, creating significant health and environmental risks.

Assessing and quantifying the contribution of vehicular emissions to local air pollution is therefore critical, not only for mitigating health impacts but also for addressing broader environmental issues such as greenhouse gas accumulation and changes in local atmospheric conditions (Lelieveld *et al.*, 2020) [7]. Advances in environmental monitoring now allow researchers to integrate multiple data sources for comprehensive assessments. Satellite-derived atmospheric data, including temperature, pressure, wind, and precipitation, provide a powerful tool for modeling pollutant dispersal and predicting environmental impacts across large spatial scales (IPCC, 2021) [5]. By combining ground-based emission measurements with machine learning classification techniques and satellite-derived climate data, this study aims to model vehicular exhaust emissions in Ibadan, evaluate their environmental consequences, and offer insights into the interactions between urban traffic, air quality, and local atmospheric dynamics.

2. Materials and Methods

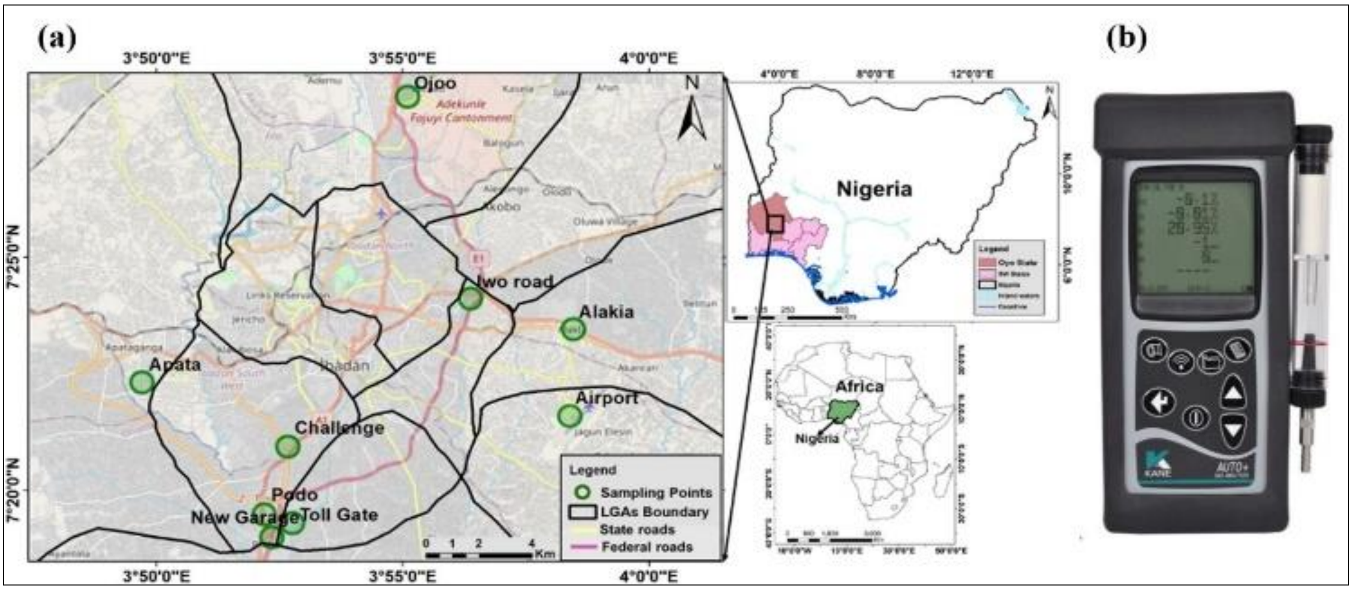

Fig 1: (a) Map of the study area; (b) the KANE gas analyser for automotive emissions used

2.1 The sampling location

The study was conducted across nine major transport hubs in Ibadan: Podo, New Garage, Toll-Gate, Iwo Road, Apata, Ojoo, Challenge, Alakia, and the Airport Road.

2.2 Raw exhaust sampling and data collection measurement: Measurements of the exhaust pollutant emissions from the trucks were performed using a Kane 5-gas analyzer, assessing CO, CO_2, O_2, and NO_x levels in 1,034 diesel trucks. a portable KANE 5-Gas Automotive Analyzer (Model 5-2) (Figure 1b), which has been programmed to detect and determine the concentrations of CO2, O2, CO, and NOx. Before each measurement round, the diesel trucks were allowed to run while" on load" during their working hours, after which the tests were conducted. The instrument probe was inserted into the exhaust pipe of the trucks end and clamped. The obtained data were recorded in percentage (%) volume for CO, CO2, and O2, and parts per million (ppm) for NOx and. Euro IV compliance checks were performed alongside.

2.3 Machine learning analysis

Exhaust emissions are governed by non-linear thermochemical reactions; thus, machine learning models are suitable for the predictions of these non-linear relationships. Machine learning (ML) is an artificial intelligence (AI) subset that focuses on developing algorithms and statistical models enabling computers to perform tasks without being explicitly programmed (Mitchell, 1997; Stanford University, n.d.) [8]. Instead of following predefined rules, ML systems learn data to make predictions through recognizing patterns. ML techniques have multivariable learning that can enable multi-input-multi-output predictions. The acquired exhaust datasets were classified into pass (1) and fail (0) according to EURO IV standard. Machine learning algorithms, including (a) Random Forest, (b) Logistic Regression, (c) Naïve Bayes, (d) Neural Network, (e) k-Nearest Neighbors, and (f) Support Vector Machine, were trained to classify emissions.

3. Results and Discussion

Table 1: Euro IV Status of Vehicular Exhaust emission (NO_X)

Location	Count	Fail	Pass	% Fail	% Pass
Podo	27	09	18	33	67
New Garage	20	14	06	70	30
Toll-Gate	22	21	01	95	5
Iwo Road	102	36	66	35	65
Apata	36	05	31	14	86

Ojoo	240	19	221	8	92
Challenge	224	22	202	10	90
Alakia	88	08	80	9	91
Airport	273	53	220	19	81
Total	1032	187	845	18.1	81.9

Pass means vehicles that meet Euro IV emission standards. Fail means they do not meet the emission standards

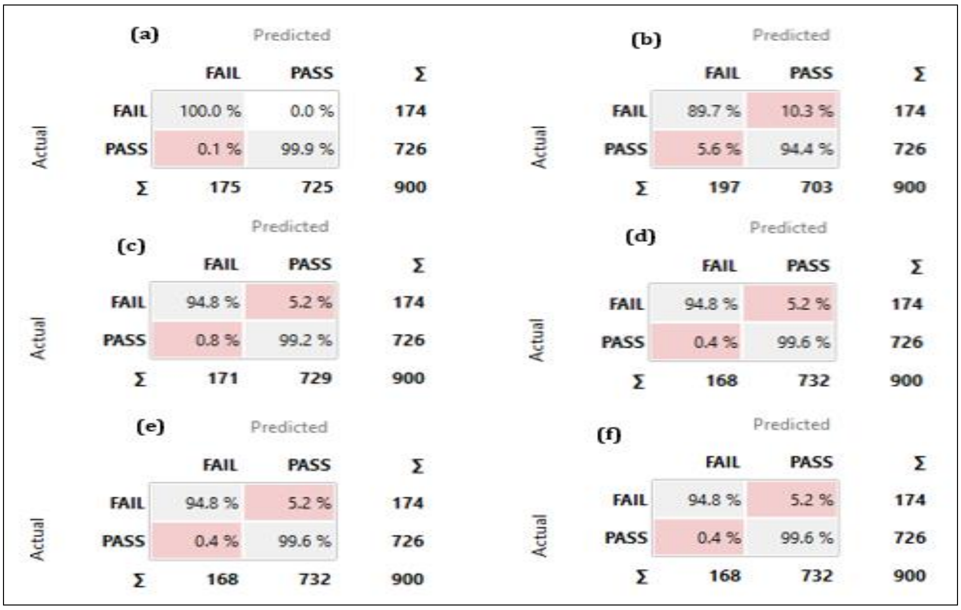

Fig 2: Confusion matrix of predicting models (a) kNN (b) Logistic Regression (c) Naïve Bayes (d) Neural Network (e) Random Forest (f) SVM on the Vehicular Exhaust Emissions

The interpretation model by model in Table 2 kNN shows a Perfect metrics (all ~0.999-1.000) which is Extremely high performance, while Logistic Regression has High AUC (0.997), CA (0.983), F1 (0.983) which indicates Strong performance overall, slightly lower than kNN/RF/NN its MCC (0.946) suggests very reliable predictions with minimal misclassification this makes it a Robust, interpretable model; less likely to overfit than kNN. Naïve Bayes has Lower performance (AUC 0.973, CA 0.934, F1 0.936) with Precision (0.939) and recall (0.934) are decent but not as strong as other models, its MCC (0.802) indicates moderate reliability and May not capture complex feature interactions due to independence assumption. Neural Network is another model Near-perfect performance (AUC 1.0, CA 0.997, F1 0.997) and very strong, its MCC 0.989 shows excellent correlation between predicted and true labels. Random Forest is a model with Perfect metrics (like kNN) it has High robustness to overfitting compared to kNN, but perfect scores still warrant checking dataset size and SVM also has Very strong performance (AUC 0.999, CA 0.987, F1 0.987) though Slightly lower than kNN/RF/NN, but still excellent. Overall, the Top performers: kNN, Random Forest, Neural Network all show near-perfect classification, while we can classify Logistic Regression and SVM as Safe, interpretable options. With High performance, easier to explain and Naïve Bayes is of Lower performer decent but not optimal for this dataset.

Table 2: Performance Indicators of Machine Learning Algorithms

Model	AUC	CA	F1	Prec	Recall	MCC	Spec
kNN	1.000	0.999	0.999	0.999	0.999	0.996	1.000
Logistic Regression	0.997	0.983	0.983	0.983	0.983	0.946	0.957
Naïve Bayes	0.973	0.934	0.936	0.939	0.934	0.802	0.906
Neural Network	1.000	0.997	0.997	0.997	0.997	0.989	0.990
Random Forest	1.00	0.999	0.999	0.999	0.999	0.996	1.000
SVM	0.999	0.987	0.987	0.987	0.987	0.957	0.957

Common classification metrics

- **AUC (Area under the ROC Curve):** Measures model's ability to distinguish between classes (1 = perfect, 0.5 = random).
- **CA (Classification Accuracy):** Percentage of correctly classified samples.
- **F1 Score:** Harmonic mean of precision and recall; balances false positives and false negatives.
- **Precision (Prec):** Proportion of predicted positives that are actually positive.
- **Recall:** Proportion of actual positives that are correctly predicted.
- **MCC (Matthews Correlation Coefficient):** Balanced metric for binary classification, even with imbalanced classes (1 = perfect, 0 = random).
- **Specificity (Spec):** Proportion of actual negatives correctly identified.

There's a clear vertical separation between the red and blue clusters in Fig. 2. where PASS (red) points are concentrated at lower NOx levels (mostly below ~250 PPM). FAIL (blue) points dominate at higher NOx levels (above ~250-300 PPM). Whereby the relationship between CO_2 and NOx. The CO_2 (%) values for both groups overlap between roughly 1%-4%, meaning CO_2 alone doesn't distinguish PASS/FAIL well. However, NOx concentration strongly differentiates the two categories; higher NOx emissions are associated with failure. When NOx rises beyond 300 PPM, most results fail, regardless of CO_2 level. This implies that NOx is a dominant variable influencing the pass/fail outcome in emission testing

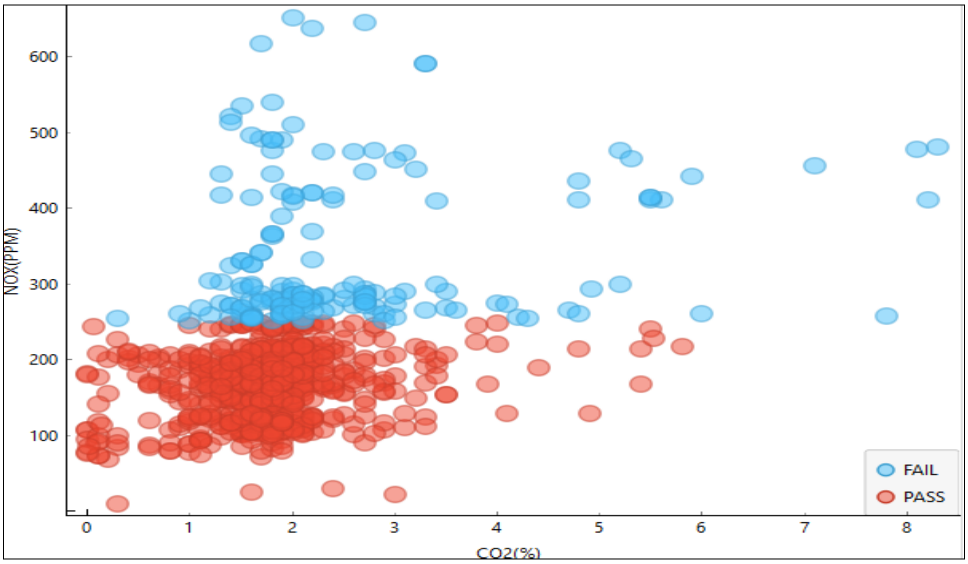

Fig 3: Scatter plots for all the predicting models of gases CO_2 and NOx, where the red dot represents correct prediction, and blue dot represents incorrect prediction.

Table 3: Lower NOx clearly passes, higher NOx fails.

Observation	PASS (Red)	FAIL (Blue)
NOx range (PPM)	Mostly below 250	Mostly above 250
CO_2 range (%)	0-4 (some up to 6)	0-8
Separation clarity	High along NOx axis	Overlap in CO_2
Main inference	Lower NOx → pass	Higher NOx → fail

Descriptive statistics revealed significant spatial variations in emission levels. Iwo Road exhibited the highest mean CO and CO_2 concentrations, while Toll-Gate recorded the highest NO_x levels, indicating poor engine maintenance. Euro IV compliance was lowest at Toll-Gate (5%) and New Garage (30%) but highest at Ojoo (92%), Challenge (90%), and Alakia (91%). Machine learning classification models demonstrated high predictive power, with Random Forest, Neural Network, and kNN achieving near-perfect accuracy (AUC = 1.000). Regression analyses indicated that CO and NO_x significantly influenced atmospheric temperature (p<0.05), while CO_2 and NO_x correlated with mean sea level pressure (p≤0.05), signifying the role of vehicular emissions in atmospheric warming and pressure variation.

4. Conclusion

Vehicular emissions in Ibadan significantly exceed acceptable environmental thresholds, especially in high-traffic zones like Iwo Road and Toll-Gate. The study identified CO and NO_x as key contributors to atmospheric warming. Enforcement of Euro IV standards, retrofit programs for old vehicles, promotion of clean fuel technologies, infrastructure improvements, and integration of satellite-based monitoring systems are recommended to mitigate emission impacts and promote sustainable urban air quality.

References

1. Andreae, M. O., Rosenfeld, D., Artaxo, P., Costa, A. A., Frank, G. P., Longo, K. M., & Silva-Dias, M. A. F. (2004). Smoking rain clouds over the Amazon. Science, 303(5662), 1337-1342. https://doi.org/10.1126/science.1092779
2. European Commission. (2007). Regulation (EC) No 715/2007 on type approval of motor vehicles with respect to emissions from light passenger and commercial vehicles (Euro 5 and Euro 6). Official Journal of the European Union.
3. Fussell, J.C., Franklin, M., Green, D.C., Gustafsson, M., Harrison, R.M., Hicks, W., Kelly, F.J., Kishta, F., Miller, M.R., & Mudway, I.S. A review of road traffic-derived non-exhaust particles: emissions, physicochemical characteristics, health risks, and mitigation measures. Environ. Sci. Technol. 56, 2022, 6813-6835.
4. Garg, A., Shukla, P. R., Bhattacharya, S., & Dadhwal, V. K. (2001). Sub-region (district) and sector level SO_2 and NO_x emissions for India: Assessment of inventories and mitigation flexibility. Atmospheric Environment, 35(4), 703-713. https://doi.org/10.1016/S1352-2310(00)00323-1
5. Intergovernmental Panel on Climate Change (IPCC). (2021). Sixth Assessment Report: The Physical Science Basis. Cambridge University Press.

6. Kojima, M., & Lovei, M. (2001). Urban air quality management: Coordinating transport, environment, and energy policies in developing countries. The World Bank.
7. Lelieveld, J., Pozzer, A., Pöschl, U., Fnais, M., Haines, A., & Münzel, T. Loss of life expectancy from air pollution compared to other risk factors: a worldwide perspective. Cardiovasc. Res. 116, 2020, 1910-1917.
8. Mitchell, T. M. (1997). Machine learning. McGraw-Hill.
9. Stanford University. (n.d.). Machine learning. Retrieved November 11, 2025, from https://cs229.stanford.edu/
10. Organisation for Economic Co-operation and Development (OECD). (2014). The cost of air pollution: Health impacts of road transport. OECD Publishing.
11. Seinfeld, J. H., & Pandis, S. N. (2016). Atmospheric chemistry and physics: From air pollution to climate change. John Wiley & Sons.
12. Zhang, Q., Sun, L., Wei, N., Wu, L. & Mao, H. The characteristics and source analysis of vocs emissions at roadside: Assess the impact of ethanol-gasoline implementation. Atmos. Environ. 263, 2021, 118670.
13. Wallace, J. M., & Hobbs, P. V. (2006). Atmospheric science: An introductory survey (2nd ed.). Academic Press.
14. WAQ. World Air Quality. World Air Quality Report. 2019, 1-22.
15. World Health Organization (WHO). (2018). Ambient (outdoor) air pollution. https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health