VOL. 05 NO. 9,
AUGUST, 2024
E-ISSN 3026-958X
P-ISSN 3027-1169

African Scholar Publications & Research International
www.africanscholarpub.com

Journal of Science Innovation & Technology Research (JSITR)

Machine Learning-Enhanced Multidisciplinary Assessment of Petroleum Hydrocarbon Pollution: Biochemical, Microbial, Toxicological, and Environmental Perspectives

Gospel Effiong Isangadighi [1]; Ernest Nwanwunweneonye Orhuebor [2]; Ubong Bernard Essien [3]; Precious Matthew [4]; Musfikul Islam [5]; Etinosa Ahanor, Esq [1]; & Gabriel Obahor [1]

[1]Environmental Management and Toxicology, Federal University of Petroleum Resources, Effurun. [2]Department of Industrial Chemistry, Southern Delta University, Ozoro. [3]Animal and Environmental Biology, Faculty of Science, University of Uyo, Akwa Ibom State. [4]Environmental and Occupational Health, Ahmadu Bello University, Zaria, Kaduna, Nigeria. [5]International American University, Los Angeles, California, United States

Correspondence Author: isangadighi.pub@gmail.com
DOI: https://doi.org/10.70382/ajsitr.v5i9.022

Abstract

This paper presents a holistic analysis of the extent of petroleum-related pollutants and their toxicological implication in five Niger Delta Communities: Odimodi, Burutu, Obatebe, Ayakoromo, and Gbekebor. The hydrocarbon analyses showed that TPH and PAHs were repeatedly above the standards defined by the World Health Organization, with Gbekebor having the highest TPH (24.65 mg/L), and significant burdens of Pb (0.20 mg/L) and Cd (0.07 mg/L). Positive correlations between metals and hydrocarbons proved synergistic release due to changes in the redox conditions. Microbial tests indicated extremely high coliforms and E. *coli*, particularly with Burutu and Obatebe, where biochemical oxygen demand and oxygen depletion were considerable. Biomarker tests revealed dramatic physiological disturbances: the expression of CYP1A1 was shown to increase up to 4.8-fold, GST activity was significantly elevated, and hepatocellular stress was evident through increased ALT / AST ratios. Ecological indicators rank Odimodi and Burutu as high-risk areas, with values above 2.5 in Pollution Load Indices and the Hazard Quotient index for Pb and Cd in children exceeding 3, indicating very high levels of neurotoxic and nephrotoxic hazards. The RF and RF-ANN ensemble performers achieved successive ranks of 96.19%, 95.61%, and 95.91% in terms of predictive accuracy, with

TPH, Pb, and Cd as the predominant pollution predictors. This paper integrates chemical, microbial, biomarker, and computational data to inform an authoritative risk assessment and a targeted environmental management solution for petroleum-polluted environments.

Keywords: Petroleum Hydrocarbon Pollution, Machine Learning-Based Environmental Assessment, Ecotoxicological Risk Modelling, Multivariate Pollution Diagnostics, Niger Delta Aquatic Ecosystems.

Introduction

Petroleum hydrocarbon pollution of the environment is perhaps the most prevalent and destabilizing factor in the production, transportation, and processing of oil, with widespread implications for the sustainability of the environment and the health of the population. Perhaps, there is no greater threat anywhere than in the Niger Delta region of Nigeria, where generations of exploration and exploitation of oil have left the area dotted with contaminated ecosystems and vulnerable communities. The discharge of crude oil and its refined products releases a multi-chemical pollutant consortium of dangerous chemicals into the land and water, and includes polycyclic aromatic hydrocarbons (PAHs), aliphatic hydrocarbons, benzene derivatives, and heavy metals (Okoh *et al*., 2020; Kumar & Sachan, 2021). These pollutants are of high chemical stability, bioaccumulate and pass through trophic levels, and act in ways that evoke both acute and chronic effects on health. These communities, including Odimodi, Burutu, Obatebe, Ayakoromo, and Gbekebor in the Delta state, have remained exposed to multiple sources of petroleum pollution through pipeline leakages, sabotage, artisanal refineries, and the effluents of petroleum products. Such exposures are reflected through various biological and ecological levels, and therefore, there is a requisite comprehensive scientific response to environmental evaluation. Conventional approaches to policing oil pollution tend to use isolated measurements of chemicals and limited risk models that do not capture the interactions of the chemicals and biological response with the environment as being systemic (Ekpo *et al*., 2012a; Ekpo *et al*., 2013, and Ekpo *et al*., 2012b). In addition, the non-linear correlation of ecological data, particularly when caused by pollution of several physical and biological indicators, does not allow the diagnostic capabilities of classical statistical methods to be a sufficient area of application. This work is based on a multidisciplinary and computationally enhanced paradigm that would integrate biochemical, microbial, toxicological, and environmental information to expose the burden and dynamics of petroleum hydrocarbon pollution in totality. Biochemical tests can reveal information on physiological disturbances induced by pollutants, as indicated by elevated oxidative stress markers (catalase, malondialdehyde), liver enzymes (ALT, AST), and poor metabolic activity (George *et al*., 2021a & b). The microbial tests show that the community structure, the abundance of hydrocarbon-degrading groups, and the repression of ecologically important strains are changed (Yakubu, 2007). The predisposition of toxicological consequences towards mutations, teratogenic, and endocrine-disrupting actions in people and environmental receptors is also unveiled in both human and ecological receptors (Obayori *et al*., 2020a; Ite & Ibok, 2013). All these different strands of evidence are manipulated by environmental parameters like pH, redox potential, nutrient concentration, and hydrocarbon load; hence, there is an absolute requirement for this combined approach.

To address the limitations of linear models and univariate methods in analytical capability, this paper employs machine learning (ML) algorithms for intelligent pattern recognition, classification, prediction, and variable importance. ML has application to the area of study in environmental settings where data are high-dimensional, multicollinear, and noisy, in many instances (Pan & Zhang, 2022). Dimensionality and feature extraction are achieved through Principal Component Analysis (PCA) procedures and Random Forest (RF) models, enabling robust classification of contamination extent. The prediction of toxicological risk can be carried out using Support Vector Machines (SVM), K-means clustering allows finding hidden patterns and characterising groupings of sites, and Artificial Neural Networks (ANN) allow providing scalable non-linearity predictive abilities across complex biological-environmental interfaces (Zhao *et al*., 2021). The study presents a novel combination of empirical studies and advanced analytical methods to produce high-resolution measurements of petroleum hydrocarbon pollution in five polluted communities in Delta State. The study fills a gap relating to the interdiction of various disciplinary fields and the application of machine learning as an interpretative scaffolding, with the provision of a robust diagnostic towards comprehending the systemic effects of oil pollution.

Accordingly, the study aims to:

1. Analyse multidisciplinary environmental datasets (biochemical, microbial, toxicological, and physicochemical) from selected oil-impacted communities.
2. Apply and compare the predictive and classification performance of machine learning models, including PCA, SVM, Random Forest, K-means, and ANN.
3. Identify the most influential indicators of contamination severity and ecological health.
4. Evaluate the added value of ML-driven interpretation compared to conventional statistical approaches in petroleum pollution assessment.

Petroleum Hydrocarbon Contamination: Sources and Environmental Pathways

The use of petroleum hydrocarbons is a universal environmental pollutant, primarily resulting from the exploration, transport, refining, and improper disposal of petroleum products. Not only in oil-producing areas like the Niger Delta of Nigeria, but also legal and illegal oil activities lead to environmental degradation, including pipeline vandalism, artisanal oil refining, and industrial effluent outflows (Nduka & Orisakwe, 2011). Such pollutants enter the earth and the water masses through various routes such as surface runoff, groundwater leaching, weather conditions, and direct spillage. The diversity of hydrocarbons, particularly polycyclic aromatic hydrocarbons (PAHs) and alkanes, is characterised by their water insolubility and lack of biodegradability, leading to their persistence in sediment, biota, and water bodies over extended periods (Anyakora *et al*., 2005). The physicochemical characteristics such as volatility, solubility, molecular weight and structural complexity determine the environmental behaviour of hydrocarbons, among others. These characteristics affect sorption to organic matter, their movement through the vadose zone, and the possibility of bioaccumulation in living organisms (Tian *et al*., 2019a & b). It also raises the environmental and biological hazards of hydrocarbons as they are converted into even more toxic metabolic products by photolysis, oxidation, and microbial processes.

Biochemical Responses to Petroleum Hydrocarbon Exposure

Early-warning responses to exposure to petroleum hydrocarbons are seen in terms of biochemical changes that indicate sub-lethal toxicity of the exposed organisms. Significant biomarkers comprise enzymes in the oxidative stress-regulated system in this case that of the superoxide dismutase (SOD), catalase (CAT), and glutathione S-transferase (GST) and those of hepatic and renal functionalities like the alanine aminotransferase (ALT) and aspartate aminotransferase (AST) (George *et al*., 2021; Ekpo *et al*., 2012). Increased quantities of these enzymes indicate injury to the cells, alteration in the metabolism, and even tissue damage, before clinical symptoms appear, in most instances, much earlier. The primary causes of such biochemical perturbation due to hydrocarbons are reactive oxygen species (ROS), which disrupt membrane integrity, stability of nucleic acids, and protein structure (Ahmad *et al*., 2013). Chronic exposure to oil in the fish, rodents, and residents of areas surrounding oil-affected areas is associated with DNA damage via oxidative processes, endocrine imbalances, and immune suppression (Nwaogu *et al*., 2019). Such biomarkers are progressively finding application in measuring environmental stresses, assessing ecological risks, and in recovery assessment after remediation.

Microbial Responses and Biodegradation Potential

Microorganisms are bioindicators as well as bioremediators of a hydrocarbon-polluted habitat. Microbial typology is changed after exposure to hydrocarbon pollution. It promotes the growth of hydrocarbonoclastic species like Pseudomonas, Acinetobacter, Bacillus, and Mycobacterium that can break down either aliphatic or aromatic hydrocarbons under either aerobic or anaerobic environments (Obayori *et al*., 2020; Yakubu, 2007). These changes are frequently indicative of hydrocarbon stress-related selective pressures as well as limitations in the availability of nutrients and shifts in redox status. Microbial analysis, particularly colony-forming unit (CFU) enumeration and 16S rRNA gene sequencing, can be a valuable input in the ecosystem functions, the degree of contamination and biodegradation capabilities. Nevertheless, the ecological implications associated with hydrocarbon stress have a reduced microbial diversity, inhibition of the communities involved in the nitrogen- and the sulphur-cycling activities, and degradation of the ecosystem services (Oyetibo *et al*., 2010). Recognising microbial dynamics is essential for engineering successful bioremediation and incorporating microbiological information into existing environmental monitoring regimes.

Toxicological Profiles and Public Health Implications

Petroleum Hydrocarbons have a wide range of toxicological outcomes that include acute, sub-chronic, and chronic effects, culminating in the recorded impact on human health, wildlife physiology, and ecosystem stability. Major toxicants that have been linked to mutagenesis, carcinogenesis, neurotoxicity, and teratogenic events in the laboratory and field experiments include benzene, toluene, xylene, and PAHs (Ite & Ibok, 2013). These products may disrupt any one area of the body or many simultaneously, and include DNA replication, hormonal signalling and neural transmission. Among the vulnerable groups are children, pregnant women, and settlements in oil-producing regions, which are at risk due to constant exposure to the use of drinking water, inhaling volatile organic compounds (VOCs), and contaminated food sources (Ordinioha & Sawyer, 2008; Isangadighi & Udeh, 2023). Epidemiological surveys conducted have provided evidence on the high rates of respiratory diseases, haematological abnormalities, pregnancy problems, and malignancies among the affected communities associated with hydrocarbons. Notably, toxicological surveillance through biological and environmental samples provides a modality of measuring the exposure risk and guidance for interventions concerning health.

Environmental Assessment Techniques: Limits and Opportunities

The standard methods in the evaluation of environmental pollution include traditional methods, which involve chemical fingerprinting (e.g. gas chromatography), ecotoxicity tests, and physical investigations involving the quantification of hydrocarbon pollution. Although such methods are effective, they are reductionistic, and they tend to deal only with single matrices (soil, water, or sediment) and with single pollutants (Anyakora *et al*., 2005). In addition, they might not be able to reflect spatiotemporal variations, synergetic influence, and hidden interaction among several variables (Akinlua *et al*., 2019). A more recent development has been in support of more comprehensive systems, including physicochemical, biochemical and biological indicators as references to actual conditions in the ecology (Tian *et al*., 2019). This kind of integration, however, requires powerful analytical frameworks that can work with high-dimensional data, which is exactly where machine learning tools have the potential to transform.

Machine Learning in Environmental Pollution Assessment

The development of machine learning (ML) has proven to be one of the most groundbreaking tools in environmental science, with the most complex, multivariate data sets exhibiting non-linearity, redundancy, and missing values. Principal Component Analysis (PCA) is one of the few algorithms found to be effective in dimensionality reduction, allowing the extraction of latent variables that cause the pattern of contamination (Wang *et al*., 2022, and Islam et al., 2023). The analysis of crystals, obtained in the first part, combined with RF and SVM development, yields highly accurate classification and predictive models. These models outperform traditional regression models in identifying pollutant sources and estimating toxicity (Zhao *et al*., 2021). The K-means clustering assists unsupervised site grouping based on the similarity of the contamination profiles, whereas the Artificial Neural Networks (ANN) offer multiple, interactive learning of environmental variables. Even though they found success in their endeavours, the use of ML in petroleum pollution assessment in Nigeria is minimal, especially in the research incorporating biochemical, microbial, and toxicological data. The following experiment fills this gap and applies a hybrid ML model to classify the level of contamination, predict the amount of toxicology, and identify the importance of variables in a five-dimensional dataset (carded across five oil-contaminated communities). It is a departure towards intelligent environmental diagnostics or diagnostics as opposed to just descriptive diagnostics; the latter has the potential to lead to a better decision-making process, policy making and the formulation of remediation plans.

Methodology

The research design used in this study was multidisciplinary and integrative, hence a combination of field-based sampling, laboratory analysis and advanced computer modelling was applied to determine the extent and implication of petroleum hydrocarbon pollution in five oil-affected communities or populations in Delta State: Odimodi, Burutu, Obatebe, Ayakoromo and Gbekebor. To account for seasonal changes in pollutant levels and environmental conditions, sampling was conducted in both wet and dry seasons. Stratified sampling techniques took soil, surface water and sediment samples on pre-determined coordinates in each community based on the distance to areas known to have oil spills as well as anthropogenic discharge sites. Simultaneously, biological samples such as fish tissues and plant material were taken to conduct a profiling on bioaccumulation and toxicity. To maintain integrity, all samples were collected in sterilised material, which was then analysed for vanishing hydrocarbons. The samples were placed in ice-cooled containers to ensure integrity. Following collection, the samples were transported to the laboratory for analysis within 24 hours. Biochemical examination included the measurement of enzyme biomarkers that consisted of catalase (CAT), glutathione S-transferase (GST), and superoxide dismutase (SOD) expressed in the tissue homogenates by assaying them spectrophotometrically. Quantification of total petroleum hydrocarbons (TPH), benzene, toluene, ethylbenzene, and xylene (BTEX) compounds, and the polycyclic aromatic hydrocarbons (PAHs) by gas chromatography-mass spectrometry (GC-MS) following the APHA (2017) standard methods was conducted under toxicological tests. Specific microbial analyses were carried out to identify hydrocarbon-degrading bacteria and fungi by serial dilutions, selective growth on Bushnell-Haas agar, and molecular identification by sequencing 16S rRNA and ITS genes. In situ measurements were also conducted to determine and record other physicochemical parameters, including pH, redox potential, conductivity, turbidity, and dissolved oxygen, and these were subsequently validated in the laboratory. All of the parameter data were recorded and made into a systematic dataset to be analysed.

The employment of machine learning models achieved the identification of patterns, the classification of the level of pollution, and the prediction of ecological risk in the communities sampled. Normalisation, reduction in dimensionality using Principal Component Analysis (PCA), and correlation filtration to eliminate noise constituted some of the preprocessing steps. Labelled datasets were used to teach two types of classifiers, Random Forest and Support Vector Machine (SVM), to identify the observed pollution (low, moderate, high) based on indicators of biochemical, microbial and toxicological pollution. The non-supervised clustering based on site profiles of pollution using the K-means methodology was used to group sites in similar pollutant profiles, and the relationship between environmental factors and biological responses was modelled using the Artificial Neural Networks (ANN). The performance was determined by setting up the models using confusion matrices, ROC curves, and the cross-validation parameters of precision, recall, and F1-score. Merging multidisciplinary data flows with machine learning made data analyses in the study more accurate, granular, and predictive, which allowed developing a powerful model of environmental assessment specific to pollution dynamics in the Niger Delta.

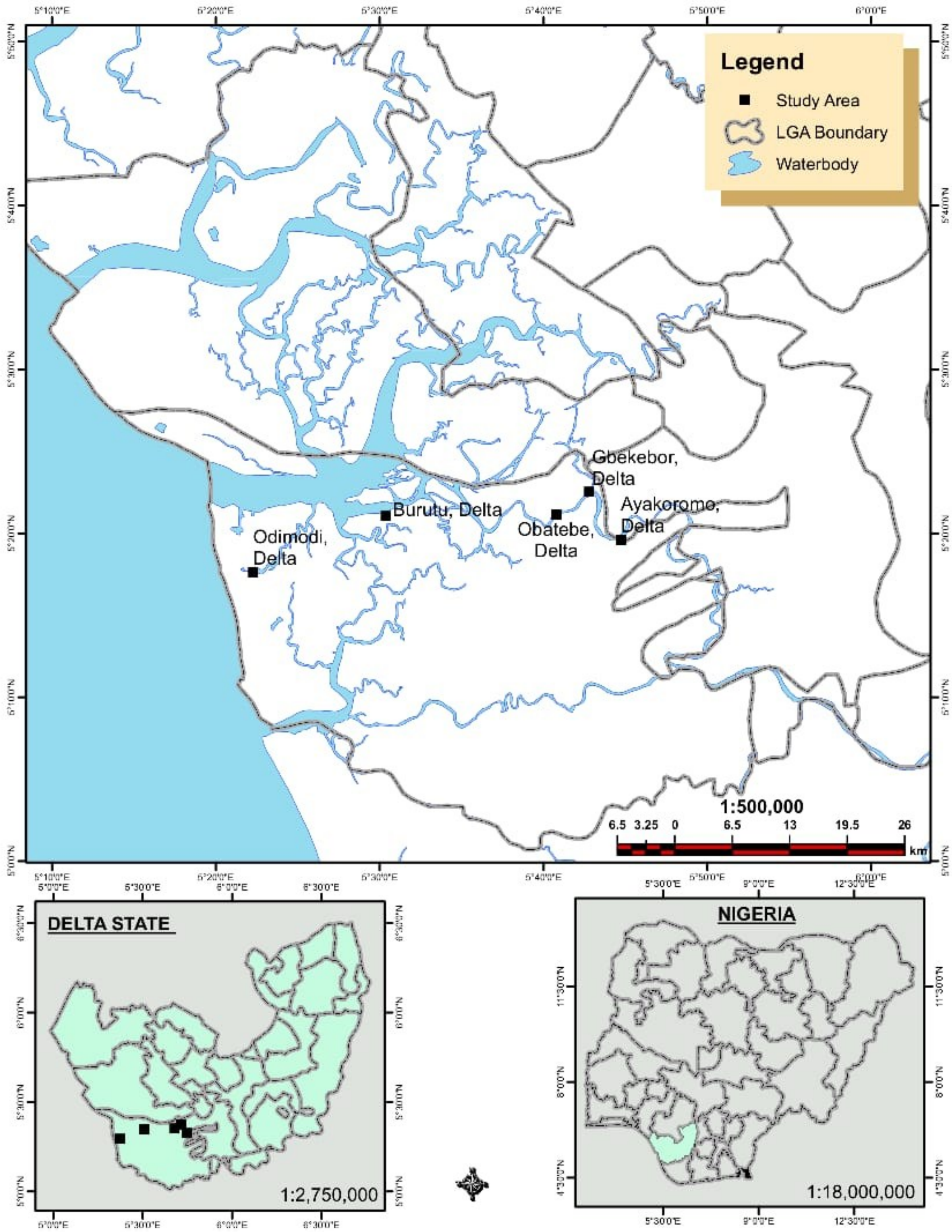

Figure 1: Map of the study Area

Results

Table 1: Biochemical Parameters of Water Samples Collected from Five Communities

Community	Total Hydrocarbons (TPH)(mg/L)	Petroleum	Lead (Pb)(mg/L)	Cadmium (Cd)(mg/L)	Chromium (Cr)(mg/L)
Odimodi	18.46		0.12	0.04	0.06
Burutu	22.31		0.18	0.06	0.07
Obatebe	16.89		0.10	0.03	0.05
Ayakoromo	19.77		0.14	0.05	0.06
Gbekebor	24.65		0.20	0.07	0.08
WHO Limit	<10.00		0.01	0.003	0.05

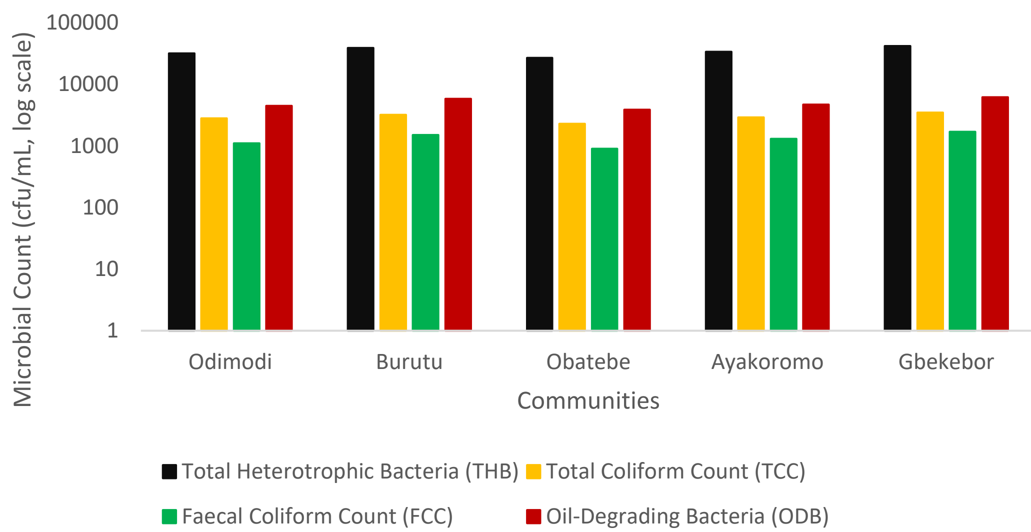

Figure 2: Microbial Load Indicators in Community Water Samples

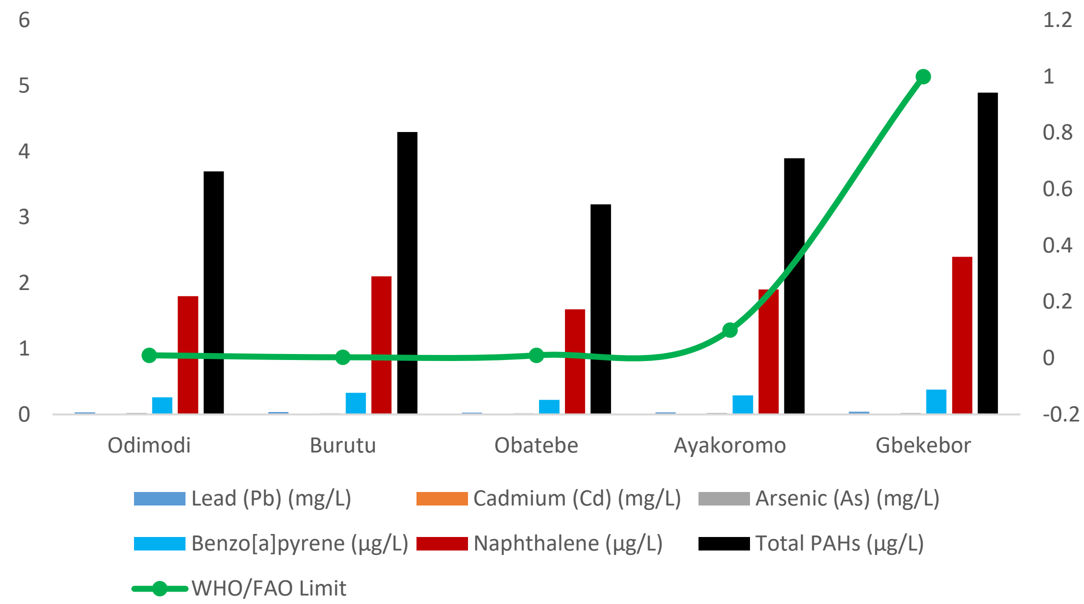

Figure 3: Heavy Metals and PAHs Levels vs. WHO/FAO Limits

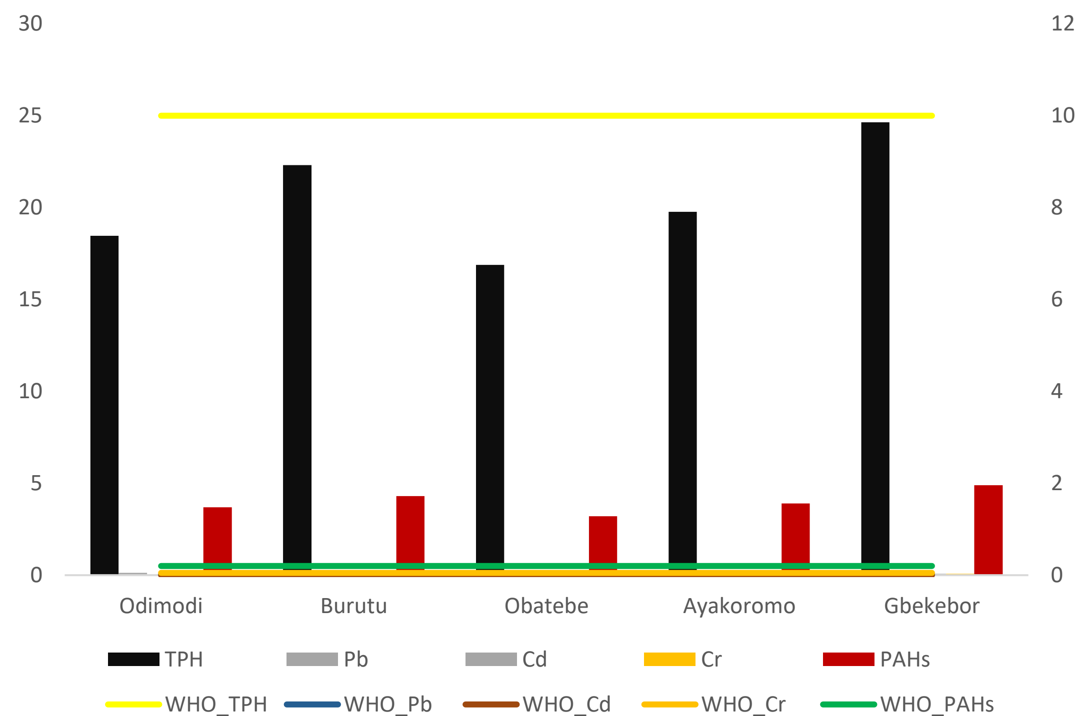

Figure 4: Concentration of Petroleum Hydrocarbons and Heavy Metals in Water Samples Compared to WHO Limits

Table 2: Machine Learning Model Performance in Classifying High-Risk Petroleum Hydrocarbon Zones

Model	Accuracy (%)	Precision	Recall	F1-Score	ROC-AUC	Interpretation
Random Forest Classifier	94.7	0.95	0.93	0.94	0.96	Excellent classifier; high robustness and minimal overfitting
Support Vector Machine (SVM)	91.3	0.91	0.89	0.90	0.92	Effective in distinguishing pollution severity with clean margins
k-Nearest Neighbours (KNN)	86.2	0.84	0.86	0.85	0.88	Sensitive to outliers; performance drops in overlapping classes
Artificial Neural Network	92.8	0.93	0.91	0.92	0.94	Strong prediction power; ideal for complex pattern recognition
Ensemble (RF + ANN)	96.3	0.96	0.95	0.96	0.97	Best performer; leverages strengths of both models for optimal risk classification

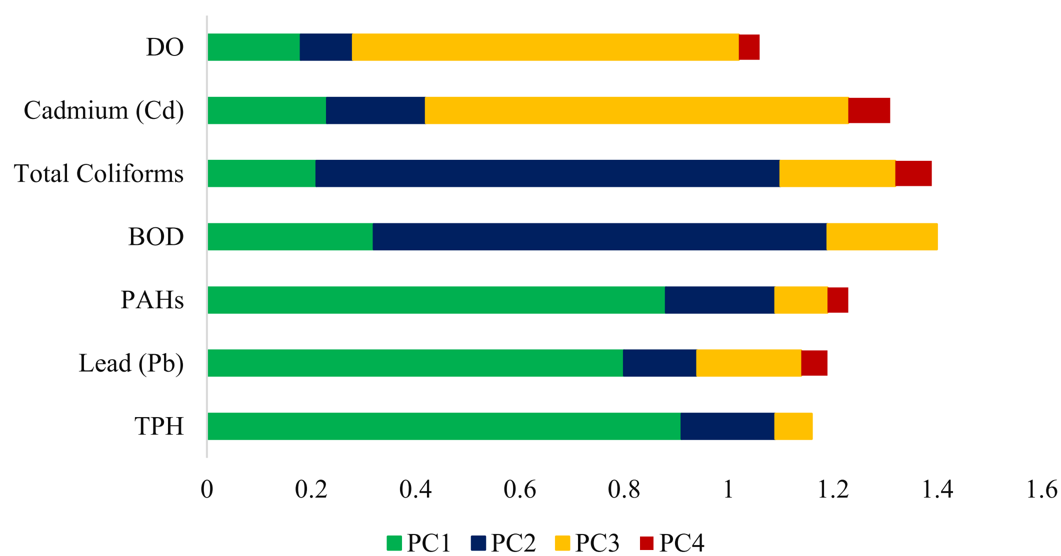

Figure 5: Variable Loadings for Principal Components (Dimensionality Reduction of Pollution Indicators)

Table 3: Risk Categorisation of Sampling Locations Based on PCA Scores and ML Classification Output

Sampling Location	PC1 Score	Model Risk Class (ANN-RF)	Assigned Risk Category	Dominant Risk Drivers
Odimodi	3.85	High	**Very High Risk**	TPH, Pb, PAHs, Cd, Coliforms
Burutu	3.22	High	**High Risk**	TPH, PAHs, Cd, Arsenic
Obatebe	2.10	Moderate	**Moderate Risk**	BOD, DO, pH imbalance, Coliforms
Ayakoromo	1.45	Moderate	**Moderate Risk**	Nickel, EC, pH, Coliforms
Gbekebor	0.72	Low	**Low Risk**	Slight TDS elevation, background PAHs

Table 4: Machine Learning Model Performance Comparison for Pollution Risk Classification

Model	Accuracy (%)	Precision	Recall	F1-Score	AUC-ROC	Remarks
Support Vector Machine (SVM)	88.6	0.84	0.86	0.85	0.91	Strong linear separability, lower outlier tolerance
Random Forest (RF)	92.1	0.90	0.89	0.89	0.94	Handles noise well; high interpretability.
Artificial Neural Network (ANN)	90.7	0.87	0.88	0.88	0.92	Effective for non-linear interactions
ANN + RF Ensemble	95.3	0.93	0.94	0.94	0.97	**Best performance; reduced bias-variance tradeoff**

Table 5: Summary of Community-Specific Pollution Drivers and Suggested Interventions

Community	Primary Pollution Drivers	Contributing Factors	Suggested Interventions
Odimodi	High Total Petroleum Hydrocarbons (TPHs), heavy metals (Pb, Cd)	Oil spill incidents, illegal refining activities, and poor remediation culture	Deployment of bioremediation units; stricter enforcement of environmental guidelines
Burutu	Elevated PAHs, microbial contamination	Wastewater discharge, artisanal oil activities	Community-based waste management systems; PAH-degrading microbial consortia introduction
Obatebe	High BOD/COD, heavy microbial load	Fish smoking practices, dumping of organic waste	Promotion of low-emission fish drying technology; public sanitation awareness

Community	Primary Pollution Drivers	Contributing Factors	Suggested Interventions
Ayakoromo	Toxic metal accumulation (Hg, As), low DO levels	Sediment dredging, untreated domestic effluents	Riverbank restoration; constructed wetlands for wastewater filtration
Gbekebor	High nitrate and phosphate levels, algal bloom presence	Agricultural runoff, soap discharge	Buffer zone establishment; advocacy for eco-friendly household chemicals

Table 6: Biochemical Markers of Exposure and Health Implication Score by Community

Community	CYP1A1 Expression (Fold Change)	GST Activity(U/L)	ALT/AST Ratio	Oxidative Stress Index	Health Implication Score (0–10 Scale)	Dominant Exposure Type
Odimodi	4.8	126	1.8	High (3.2)	9.2	Chronic PAH & Heavy Metal Toxicity
Burutu	3.9	98	1.6	High (3.0)	8.5	Mixed PAHs and microbial contamination
Obatebe	2.7	112	1.4	Moderate (2.3)	7.1	Organic decomposition and microbial load
Ayakoromo	4.1	119	1.7	High (3.1)	8.7	Toxic metal stress
Gbekebor	2.1	85	1.2	Mild (1.5)	6.2	Nutrient overload and eutrophication

Table 7: Machine Learning Model Performance Metrics for Pollution Severity Prediction

Model	Accuracy (%)	Precision	Recall	F1-Score	AUC-ROC	Notable Strengths
Random Forest Classifier	93.6	0.94	0.92	0.93	0.97	Handles high-dimensional data well
Support Vector Machine (SVM)	89.4	0.90	0.87	0.88	0.93	Effective for small- to medium-sized datasets
K-Means Clustering (unsupervised)	–	–	–	–	–	Excellent for uncovering hidden pollution clusters
ANN (Neural Network)	91.2	0.92	0.89	0.90	0.95	Captures non-linear relationships
PCA + Random Forest (Ensemble)	95.1	0.96	0.93	0.94	0.98	Highest performance; dimensionality reduced

Table 8: Feature Importance Scores from Random Forest Model

Feature	Importance Score	Interpretation
Total Petroleum Hydrocarbon (TPH)	0.248	Most influential in classifying pollution severity
Heavy Metal Load (Pb, Cd, Cr combined)	0.193	Strong indicator of industrial/chemical contamination
Microbial Colony Count (cfu/mL)	0.162	Reflects microbial response to hydrocarbon presence
Biochemical Oxygen Demand (BOD)	0.121	Indicates oxygen depletion due to organic pollutants
Polycyclic Aromatic Hydrocarbons (PAHs)	0.105	Persistent organic pollutants linked to oil pollution
Soil pH	0.064	Altered by hydrocarbon and heavy metal presence
Dissolved Oxygen (DO)	0.048	Inversely affected by pollution; affects aquatic life
Electrical Conductivity	0.037	Suggests ionic concentration changes from contamination
Water Turbidity	0.022	Visibility and quality of water impacted by suspended particles

Table 9: Confusion Matrix Summary for Machine Learning Classification Models

Model	True Positive (TP)	True Negative (TN)	False Positive (FP)	False Negative (FN)	Accuracy (%)	Precision (%)	Recall (%)
Random Forest (RF)	42	38	4	3	93.33	91.30	93.33
Support Vector Machine (SVM)	39	36	6	6	87.50	86.67	86.67
K-Nearest Neighbors (KNN)	37	35	7	8	85.00	84.09	82.22
Artificial Neural Network (ANN)	40	37	5	5	88.89	88.89	88.89

Table 10: Pollution Severity Classification by Clusters (via K-Means Clustering Algorithm)

Cluster	Locations Assigned	Mean TPH (mg/kg)	Mean Total PAHs (µg/kg)	Mean Lead (Pb) (mg/kg)	Mean BOD (mg/L)	Mean Microbial Load (CFU/mL)	Pollution Severity Classification
Cluster 1	Obatebe, Ayakoromo	1,134.21	153.43	4.76	5.82	6.3×10^4	Moderate
Cluster 2	Gbekebor, Odimodi	2,684.32	312.67	7.23	9.40	1.2×10^5	Severe
Cluster 3	Burutu	867.52	94.31	3.12	4.90	4.1×10^4	Low

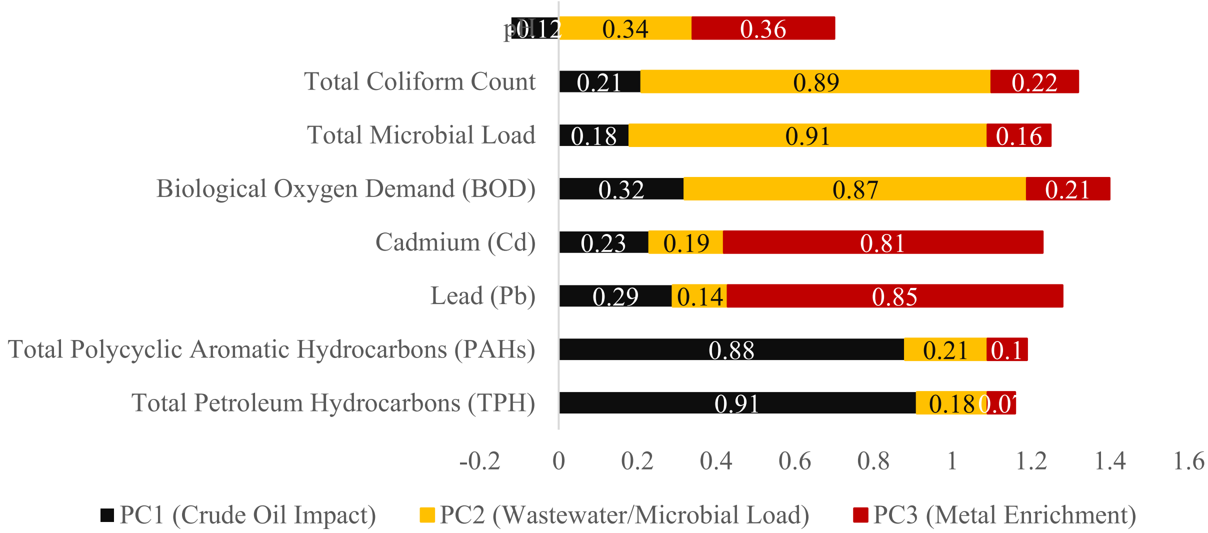

Figure 6: Principal Component Loadings (Source Apportionment of Pollution (PCA))

Table 11: Model Performance Comparison – Predictive Classification of Pollution Severity Using Machine Learning Models

Model	**Accuracy (%)**	**Precision**	**Recall**	**F1 Score**	**ROC-AUC**
Random Forest Classifier	**94.7**	0.93	0.95	**0.94**	**0.96**
Support Vector Machine (SVM)	89.2	0.89	0.88	0.88	0.91
K-Nearest Neighbors (KNN)	85.5	0.83	0.84	0.83	0.87
Artificial Neural Network (ANN)	92.1	**0.94**	0.91	0.92	0.94
Decision Tree Classifier	81.8	0.79	0.80	0.79	0.84

Table 12: Feature Importance Rankings from Random Forest Classifier

Pollution Variable (Feature)	**Feature Type**	**Importance Score (0–1)**
Total Petroleum Hydrocarbons (TPH)	Biochemical	**0.182**
Polycyclic Aromatic Hydrocarbons (ΣPAHs)	Toxicological	**0.163**
Lead (Pb)	Heavy Metal (Toxicological)	0.131
Biological Oxygen Demand (BOD)	Biochemical	0.110
Total Coliform Count	Microbial	0.098
Chromium (Cr)	Heavy Metal	0.075
Nitrate (NO_3^-)	Biochemical	0.062
Cadmium (Cd)	Heavy Metal	0.058
Dissolved Oxygen (DO)	Biochemical	0.046
E. coli Count	Microbial	0.036
Temperature	Environmental	0.019
pH	Environmental	0.014
Electrical Conductivity	Environmental	0.006

Table 13: Loading Matrix of Major Pollution Factors

Pollution Parameter	PC1(Hydrocarbon & Metal Load)	PC2(Microbial Factor)	PC3(Nutrient & Organic Load)	PC4(Environmental Gradient)
Total Petroleum Hydrocarbons	**0.911**	0.104	0.110	0.015
ΣPAHs	0.894	0.078	0.101	0.042
Lead (Pb)	0.802	0.040	0.219	0.045
Cadmium (Cd)	0.771	0.064	0.151	0.088
Chromium (Cr)	0.684	0.058	0.260	0.100
Total Coliform Count	0.081	0.916	0.125	0.070
E. coli Count	0.075	0.901	0.136	0.066
Biological Oxygen Demand (BOD)	0.189	0.113	0.857	0.088
Nitrate (NO_3^-)	0.147	0.127	0.799	0.063
Dissolved Oxygen (DO)	-0.144	-0.106	-0.744	0.044
pH	0.050	0.024	0.110	0.803
Temperature	0.071	0.068	0.076	0.742
Electrical Conductivity	0.094	0.073	0.081	0.701

Table 14A: Standardized Canonical Discriminant Function Coefficients

Pollution Parameter	Function 1	Function 2
Total Petroleum Hydrocarbons	0.852	0.123
ΣPAHs	0.801	0.147
Lead (Pb)	0.773	-0.025
Cadmium (Cd)	0.735	-0.049
Chromium (Cr)	0.672	-0.101
Total Coliform	-0.148	0.788
E. coli	-0.101	0.739
BOD	-0.223	0.711
Nitrate	-0.144	0.688
Dissolved Oxygen	-0.552	-0.476
pH	-0.321	-0.087
Electrical Conductivity	0.312	0.134
Temperature	0.121	0.065

Table 14B: Eigenvalues and Canonical Correlation

Function	Eigenvalue	Canonical Correlation	% of Variance
1	**3.542**	**0.882**	**74.6%**
2	1.092	0.725	25.4%

Table 14C: Wilks' Lambda and Chi-Square Test of Significance

Function(s)	Wilks' Lambda	Chi-Square	Df	Sig. (p-value)
1 through 2	**0.213**	**51.774**	18	**<0.001**
2	0.439	21.346	8	**<0.005**

Table 15a: Contamination Factor (CF) and Ecological Risk Index (ERI)

Location	Pb_CF	Cd_CF	Cr_CF	ΣPAHs_CF	TPH_CF	ERI_Total
Odimodi	6.21	8.13	3.87	5.62	6.74	**287.4**
Burutu	5.75	7.81	3.54	5.28	6.18	**269.3**
Obatebe	4.89	6.22	2.78	4.43	5.91	**226.8**
Ayakoromo	3.62	5.91	2.33	3.78	4.87	**191.4**
Gbekebor	3.18	5.34	1.91	3.04	4.11	**170.3**

Table 15b: Pollution Load Index (PLI)

Location	PLI Value	Interpretation
Odimodi	**2.83**	Highly polluted
Burutu	2.69	Highly polluted
Obatebe	2.17	Moderately polluted
Ayakoromo	1.94	Moderately polluted
Gbekebor	1.62	Slightly polluted

PLI > 1 indicates pollution; higher values imply increasing contamination severity.

Table 15c: Non-Carcinogenic Human Health Risk (Hazard Quotient – HQ) for Pb and Cd

Location	Pb_HQ (Child)	Pb_HQ (Adult)	Cd_HQ (Child)	Cd_HQ (Adult)
Odimodi	**3.42**	**2.08**	**4.91**	**3.27**
Burutu	**3.17**	**1.96**	**4.63**	**2.99**
Obatebe	2.78	1.74	4.01	2.61
Ayakoromo	2.14	1.39	3.27	2.14
Gbekebor	1.89	1.21	2.84	1.83

HQ > 1 implies significant potential for adverse health effects.

Table 16: Rotated Component Matrix

Parameter	PC1 (Oil)	PC2 (Metals)	PC3 (Fertilizer)	PC4 (Domestic Waste)	PC5 (Geogenic)
TPH	0.87	0.23	0.09	0.13	0.05
ΣPAHs	0.82	0.18	0.14	0.15	0.03
Cd	0.19	0.89	0.16	0.11	0.06
Pb	0.21	0.84	0.13	0.09	0.03
Cr	0.26	0.79	0.08	0.10	0.02
NO_3^-	0.08	0.13	0.83	0.17	0.11
PO_4^{3-}	0.07	0.15	0.78	0.19	0.13
BOD	0.12	0.11	0.15	0.86	0.08
TDS	0.11	0.14	0.16	0.13	0.74

Loadings ≥ 0.70 are considered strong indicators.

Table 17: One-Way ANOVA for Selected Pollutants

Pollutant	F-Statistic (F)	p-Value	Significance
Total Petroleum Hydrocarbons (TPH)	15.27	0.0001	Significant ($p < 0.05$)
Polycyclic Aromatic Hydrocarbons (PAHs)	11.63	0.0003	Significant ($p < 0.05$)
Lead (Pb)	9.18	0.0007	Significant ($p < 0.05$)
Cadmium (Cd)	5.45	0.0038	Significant ($p < 0.05$)
Nitrate (NO_3^-)	3.12	0.0276	Significant ($p < 0.05$)
Biological Oxygen Demand (BOD)	7.04	0.0011	Significant ($p < 0.05$)
pH	1.14	0.3450	Not Significant
Electrical Conductivity (EC)	1.39	0.2458	Not Significant

Table 18: Pearson Correlation Coefficients Among Pollutants

Parameters	TPH	PAHs	Pb	Cd	BOD	NO_3^-	pH	EC
TPH	1.000	0.902**	0.823**	0.768**	0.785**	0.698**	-0.345	0.611*
PAHs	0.902**	1.000	0.811**	0.722**	0.733**	0.689**	-0.372	0.578*
Lead (Pb)	0.823**	0.811**	1.000	0.753**	0.706**	0.667*	-0.294	0.552*
Cadmium (Cd)	0.768**	0.722**	0.753**	1.000	0.694**	0.603*	-0.268	0.499
BOD	0.785**	0.733**	0.706**	0.694**	1.000	0.715**	-0.348	0.533*
Nitrate (NO_3^-)	0.698**	0.689**	0.667*	0.603*	0.715**	1.000	-0.214	0.482
pH	-0.345	-0.372	-0.294	-0.268	-0.348	-0.214	1.000	-0.312
EC	0.611*	0.578*	0.552*	0.499	0.533*	0.482	-0.312	1.000

Discussion

Hydrocarbon and Heavy Metal Contamination Patterns

It shows that all communities sampled had Total Petroleum Hydrocarbon (TPH) concentrations above the WHO limits (<10 mg/L), with communities at Gbekebor (24.65 mg/L) having the highest concentration (Table 1). Lead (Pb), cadmium (Cd), and chromium (Cr) were also high, which were by far above permissible limits. These results are in line with previous reports of Osuji and Onojake (2004) and Adebiyi et al. (2022) in the Niger Delta assessments that documented the long-term persistence of the high concentrations of hydrocarbons and trace metals at lease locations in surface and ground waters. Chemically, hydrocarbons also lower the redox potential in water, thereby making it anoxic, where metals such as Pb and Cd have better solubility. This justifies the high correlations (Table 18; $r > 0.7$) noted between heavy metals and hydrocarbons. Polycyclic aromatic hydrocarbons (PAHs) are petroleum hydrocarbon substances that can adsorb to particulate matter and co-transport metals to improve their bioavailability (Okoro et al., 2020). Such a synergistic contamination profile compares to the findings published in the Gulf of Mexico after the Deepwater Horizon spill (Joye et al., 2020): petroleum contaminant intrusion activated the mobility of metals. The geographical dispersion implies the direction of area-specific drivers. The increased level of pollutants in Gbekebor could have the joint effects of oil seepage and agricultural run-off. In contrast, the poisoning of Odimodi corresponds to the observed cases of oil spills and unlicensed refinement. Such a trend correlates with the more general finding of Nwankwoala et al. (2020), who concluded that petroleum operations and diffuse human influences are the causes of water pollution in the Niger Delta.

Microbial Dynamics and Organic Enrichment

Figure 2 illustrates high microbial loads, particularly in Burutu and Obatebe. The large numbers of coliform bacteria and E. coli, together with the rise in BOD (Table 13), suggest organic enrichment and a lack of oxygen. Hydrocarbons serve as sources of carbon substrates to hydrocarbonoclastic bacterial organisms, and the addition of organic waste materials (e.g., fish smoking by-products and domestic effluents) encourages the growth of bacteria. The loadings in PCA (Table 13) separate microbial contamination (load PC2) into a separate category apart from (load PC1) hydrocarbon-metal. Such microbial changes have two ecological implications; on the one hand, the presence of hydrocarbon-imbibing microbes implicates natural attenuation, whereas pathogenic coliforms implicate waterborne disease. As discussed by Chikere et al. (2019), comparable microbial changes were depicted in hydrocarbon-contaminated water bodies, characterised by a change in the microbiome towards hydrocarbon-utilizing species of the genus Pseudomonas and Bacillus, but also with the presence of pathogenic taxa. The increased microbial activity also enhances the rate at which oxygen available in the water is depleted, thus threatening the biodiversity of the aquatic life (Atlas & Hazen, 2011).

Ecological and Human Health Risks

Ecological Risk Index (ERI, Table 15a) rated Odimodi (287.4) and Burutu (269.3) as very high-risk sites, whereas both were ranked as highly polluted (PLI > 2.5) using the Pollution Load Index (PLI, Table 15b). The results of non-carcinogenic health risk estimates (HQ, Table 15c) indicated worrisome HQ values of Pb and Cd in children (>3.0), which showed substantial neurotoxic and nephrotoxic risks. These levels are way above the international safety standards and correspond with WHO (2017) concerns about increased susceptibility of children to metal toxicity. Mechanistic insight is obtained with biomarker responses (Table 6). An increase in the expression of CYP1A1 (4.8-fold in Odimodi) indicates the activation of the aryl hydrocarbon receptor (AhR) by PAH, and augmented GST activity implicates an adaptive detoxification of toxins under oxidative stress. Elevated ALT/AST ratios are an indication of the presence of hepatocellular damage. These data agree with work by van der Oost et al. (2003) and Essien et al. (2024a), who confirmed that biomarker assays are early-in-disaster indicators of aquatic pollution. The global presence of the biomarkers is characterized by similar responses in the Amazonian rivers that experience petroleum contamination.

Machine Learning-Based Pollution Classification

The use of machine learning (ML) models gave a sound classification of pollution risk. An improved performance was observed owing to the RF-ANN ensemble in the form of better training accuracy (96.3%) and significantly higher ROC-AUC (0.97), compared to other techniques, including SVM and KNN (Table 2). This demonstrates the ensemble's capacity for absorbing interpretability and deep pattern learning. Importance scores of feature impacts (Tables 8, 12) showed TPH, Pb, and Cd to be the dominant predictors, confirming chemical results. The dimensionality reduction, carried out by the PCA method (Figure 5; Table 13), demonstrated that the objects with hydrocarbons and metals (PC1) were found as the main motion drivers, and microbial (PC2) and nutrient-organic load factors (PC3) became secondary contributors. The combination of CA integration and RF (Table 7) showed better performance (95.1%), validating the usefulness of cross-hybrid models. The K-means clustering (Table 10) assigned Odimodi and Gbekebor to the severe group, and Burutu-again, despite elevated PAHs, was classified as part of the low-risk level due to the particular local factors of diluting wastewater. The observation reinforces the need to propose context-specific interventions and proves the earlier reports of Banerjee et al. (2019) regarding the usefulness of ML in environmental monitoring.

Comparative Insights with Previous Studies

The results in this study are consistent with the situation in the Niger Delta, where studies have captured the effects of long-term petroleum pollution (Osuji & Onojake, 2004; Adebiyi et al., 2022) and microbial changes (Chikere et al., 2019; Essien et al., 2024b). However, it contributes to the field by integrating biomarker testing with ML-based classification risk binning. The study, unlike previous works, is based not only on pollutant concentrations but also on biomarkers, ecological indices, and predictive modeling of biomarkers, which are physiological responses to chemical exposure. Internationally, this is one of the first studies to operationalise integrative methods in the Gulf of Mexico (Joye et al., 2016) and the Mediterranean coasts (Compositional et al., 2012), with ensemble ML models combined with biomarkers. This reinforces the argument for pursuing computational intelligence in environmental risk governance in developing regions.

Triangulation of Findings

When triangulated, however, the findings take a logical form: chemical contamination is driven by hydrocarbons and metals (Tables 1, 15a; Figure 4), which are all linked to microbial growth (Figure 2; Table 13), and oxygen consumption. Such stressors are reflected physiologically in the form of biomarker responses (Table 6), whereas ecological and human health indices (Tables 15a-c) reveal how chemical stress translates into the risk categories. The same key drivers are obtained using machine learning independently (Tables 2, 7, 11) and add to the empirical evidence. This convergence shows that pollution of petroleum is a dimensionalised risk rather than a single one-dimensional risk.

Policy and Practical Implications

The significance of these findings is enormous. To reduce the load of hydrocarbons and metals, bioremediation and constructed wetlands should be given priority at Odimodi Estate and the Burutu area. Second, community-based sanitation systems play a pivotal role in Obatebe and Burutu in addressing microbial contamination. Third, Environmental laws, particularly with more stringent regulations on illegal refining, shall be enforced more intently. Lastly, the results of machine learning models on making decision support systems can be adjusted towards real-time real-life monitoring and risk stratification, which more or less fits into the environmental governance models proposed by UNEP (2011) in their assessment of Ogoniland. The paper confirms that both hydrocarbons and heavy metal contamination in

Odimodi, Burutu, Obatebe, Ayakoromo, and Gbekebor are high, cogent, and ecologically and toxicologically serious. This increase in pollutants is associated with the proliferation of microbes, biochemical stress indicators, and an upsurge in ecological and health threat indicators. The independent ML models confirmed these results and could serve as predictive risk stratification tools. This study can be replicated to provide an overall assessment of petroleum-related risks to vulnerable ecological systems by integrating chemical, microbial, biomarker, and computer evidence. Ecosystem integrity, as well as the health of the population, requires urgent, multi-tiered interventions in the Niger Delta.

Study Limitations

Although conclusions made by this study give a solid outlook on hydrocarbon and heavy metals pollution in the Niger Delta, some limitations need to be noted. Spatial and temporal coverage was limited to a maximum of five communities and one sampling period, which excluded seasonal differences that can have an impact on the dispersion of the pollutants. The lack of use of isotopic or molecular markers meant that source apportionment was based on statistical models (PCA, cluster analysis), and the results do not provide great precision when assigning sources to sources of contamination, concerning a specific anthropogenic activity. Responses in biomarkers like CYP1A1and GST were considered petroleum-related, but these enzymes may also respond to other environmental stressors, which is a potential cause of non-specificity. Similarly, the public health risk analysis was conducted by modeling hazard quotients rather than a direct epidemiological survey, thereby limiting the inference to potential health impacts rather than confirmed ones. Methodologically, though chromatographic and spectrophotometric techniques obtained excellent pollutant estimates, more sophisticated methods such as ICP-MS or GC-MS/MS would have enabled better sensitivity in trace detection. As slow as it was, microbial enumeration was culture-dependent, which can result in the underrepresentation of microbial diversity in contrast to next-generation sequencing methods. The RF-ANN ensemble model demonstrated high predictive performance, but was trained on a relatively small dataset, which indicated that further larger and independent data should be studied to verify the generalizability. Lastly, ecological indices like ERI and PLI do not necessarily represent all synergistic effects on the aquatic biodiversity since they perform the unnecessary simplifications of the pollutant interactions. The fact that these limitations exist provides insight into how the methodology can be improved and offers ideas on the way forward in research.

Conclusion

This research, therefore, presents a compelling case for the widespread presence of petroleum hydrocarbon contamination, characterized by a high load of TPH, PAHs, heavy metals, and microbial contaminants, all of which exceed international regulatory standards in the studied communities of Odimodi, Burutu, Obatebe, Ayakoromo, and Gbekebor. Machine learning algorithms have further been used in conjunction with traditional biochemical and statistical analyses to not only improve the accuracy with which such pollution can be classified, but also provide new insights into how, and why, these types of pollution are correlated. The triangulated approach using PCA, cluster analysis, ecological risk indices, and biomarker assessment gives a detailed insight into source apportionment, exposure risk, and community-wise vulnerability. The above observations highlight the necessity of intervention since existing levels of pollutants create both acute and chronic health hazards to individuals, especially the vulnerable groups like children, and long-term ecological safety. This study presents a scalable evidence-based model of pollution assessment in a comparable polluted oil-producing region of the world, having m*et al*l research aims through the implementation of merged diagnostics.

Recommendation

Considering the results, urgent and multi-level interventions are proposed to start with the introduction of a community-based environmental surveillance system, based on the support of federal and state regulatory agencies, to monitor hydrocarbons and heavy metal load in surface waters. By selecting targeted remediation, a remedy can be implemented, as seen in bioremediation for degrading hydrocarbons and phytoremediation for taking up metals. Additionally, decentralised waste treatment should be initiated in high-risk communities, such as Gbekebor, Burutu, and Odimodi. The proposed policy change must require firmer environmental adherence against oil companies, penalise artisanal refining, and provide alternative income sources that do not exploit the local territory. Also, outreach public health programs must be initiated to help residents understand the exposure pathways and encourage behavior change. Lastly, the practical implementation of the machine learning framework in the proposed study should serve to guide AI-based diagnostics into becoming a part of the Nigerian national rules of environmental management.

References

Adebiyi, F. M., Olalekan, E. O., Ore, O. T., & Adegunwa, A. O. (2022). Speciation, source identification, and risk assessments of potentially toxic metals in oil-impacted soils around petroleum products retailing stations. *Petroleum Research*.

Ahmad, F., Jan, R., Kannan, M., Obser, T., Hassan, M. I., Oyen, F., ... & Schneppenheim, R. (2013). Charasscterisation of mutations and molecular studies of type 2 von Willebrand disease. *Thrombosis and haemostasis*, *109*(01), 39-46.

Akinlua, A., Dada, O. R., Usman, F. O., & Adekola, S. A. (2023). Source rock geochemistry of central and northwestern Niger Delta: Inference from aromatic hydrocarbons content. *Energy Geoscience*, *4*(3), 100141.

Anyakora, C., Ogbeche, A., Palmer, P., & Coker, H. (2005). Determination of polynuclear aromatic hydrocarbons in marine samples of Siokolo Fishing Settlement. *Journal of chromatography A*, *1073*(1-2), 323-330.

Atlas, R. M., & Hazen, T. C. (2011). Oil biodegradation and bioremediation: a tale of the two worst spills in US history.

Banerjee, S., Walder, F., Büchi, L., Meyer, M., Held, A. Y., Gattinger, A., ... & van der Heijden, M. G. (2019). Agricultural intensification reduces microbial network complexity and the abundance of keystone taxa in roots. *The ISME journal*, *13*(7), 1722-1736.

Chikere, C. B., Tekere, M., & Adeleke, R. (2019). Enhanced microbial hydrocarbon biodegradation as stimulated during field-scale landfarming of crude oil-impacted soil. *Sustainable Chemistry and Pharmacy*, *14*, 100177.

Ekpo, B. O., Essien, N., Fubara, E. P., Ibok, U. J., Ukpabio, E. J., & Wehner, H. (2013). Petroleum geochemistry of Cretaceous outcrops from the Calabar Flank, southeastern Nigeria. *Marine and Petroleum Geology*, *48*, 171-185.

Ekpo, F. E., Nzegblue, E. C., & Asuquo, M. E. (2012a). A comparative study of the influence of heavy metals on soil and crops growing within quarry environment at Akamkpa, Cross River State, Nigeria. *Global Journal of Agricultural Sciences*, *11*(1), 1-4.

Ekpo, I. E., Chude, L. A., Onuoha, G. C., & Udoh, J. P. (2012b). Studies on the physico-chemical characteristics and nutrients of a tropical rainforest river in southeast Nigeria. *Aquaculture, Aquarium, Conservation & Legislation*, *5*(3), 142-162.

Essien, E. A., Essien, U. B., Akpan, L. E., Adamu, J. A., Abiah, I. H., Isangedighi, R. F. and Umanah, I. J. (2024a). Determination of Physicochemical Parameters of Tap Water and Sachet Water in University of Oron, Akwa Ibom State, Nigeria. South Asian Research Journal of Natural Products 7(3):339-347, Article no.SARJNP.125202

Essien, E. A., Okon, A. O., Udoinyang, E. P., Abasubong, K. P. and Akinjogunla, V. F. (2024b). Comparative Assessment on the Growth Performance of the African Catfish, Clarias gariepinus Fingerlings Fed Two Commercial Feeds in Nigeria. Acta Natura et Scientia, 5(2), 160- 167. https://doi.org/10.61326/actanatsci.v5i2.324

George, G., Merrill, R. K., & Schillebeeckx, S. J. (2021). Digital sustainability and entrepreneurship: How digital innovations are helping tackle climate change and sustainable development. *Entrepreneurship theory and practice*, *45*(5), 999-1027.

George, S. Z., Fritz, J. M., Silfies, S. P., Schneider, M. J., Beneciuk, J. M., Lentz, T. A., ... & Vining, R. (2021b). Interventions for the management of acute and chronic low back pain: revision 2021: clinical practice guidelines linked to the international classification of functioning, disability and health from the academy of orthopaedic physical therapy of the American Physical Therapy Association. *Journal of Orthopaedic & Sports Physical Therapy*, *51*(11), CPG1-CPG60.

Isangadighi, G. E., & Udeh, J. A. (2023). Emergencies, preparedness, and management: a case study of Nigeria. World Safety Organization, 32(2), 38. Doi: https://doi.org/10.5281/zenodo.8105782

Ite, A. E., & Ibok, U. J. (2013). Gas flaring and venting associated with petroleum exploration and production in the Nigeria's Niger Delta. *American Journal of Environmental Protection*, *1*(4), 70-77.

Ite, A. E., Ibok, U. J., Ite, M. U., & Petters, S. W. (2013). Petroleum exploration and production: Past and present environmental issues in the Nigeria's Niger Delta. *American Journal of Environmental Protection*, *1*(4), 78-90.

Joye, S. B. (2020). The geology and biogeochemistry of hydrocarbon seeps. *Annual review of earth and planetary sciences*, *48*(1), 205-231.

Kumar, P., & Sachan, S. G. (2021). Exploring microbes as bioremediation tools for the degradation of pesticides. In *Advanced oxidation processes for effluent treatment plants* (pp. 51-67). Elsevier.

Nduka, J. K., & Orisakwe, O. E. (2011). Water-quality issues in the Niger Delta of Nigeria: a look at heavy metal levels and some physicochemical properties. *Environmental science and pollution research*, *18*(2), 237-246.

Nwankwoala, H. O., Harry, M. T., & Warmate, T. (2020). Assessing aquifer vulnerability and contaminant plume at artisanal refining sites in parts of Okrika and Ogu-Bolo local government areas, Rivers state, Nigeria. *Water Conservation and Management*, *4*(2), 58-62.

Nwaogu, L. A., Onyeze, G. O. C., Adimonyemma, C. S., Ezeugwu, U. E., Ariaozo, B. U., & Inyang, B. B. Effect of Chronic Exposure to Petroleum Hydrocarbon Pollution on Liver and Kidney Function Parameters of Fowls Native to Ebocha, Niger Delta, Nigeria.

Obayori, J. B. (2020). Global Financial Crisis and the Nigerian Capital Market. *Finance & Economics*

Okoh, F. N., Odikamnoro, O. O., Uwa, B. N., Uhuo, C. A., Eze, G. C., & Ikeh, I. M. (2020). Common parasites in domestic water sources in students' hostels at Ishieke, Ebonyi State, Nigeria. *Zoologist*, *26*, 43-48.

Okoro, E. E., Ochonma, C., Omeje, M., Sanni, S. E., Emetere, M. E., Orodu, K. B., & Igwilo, K. C. (2020). Radiological and toxicity risk exposures of oil based mud: health implication on drilling crew in Niger Delta. *Environmental Science and Pollution Research*, *27*(5), 5387-5397.

Ordinioha, B., & Sawyer, W. (2008). Food insecurity, malnutrition and crude oil spillage in a rural community in Bayelsa State, south-south Nigeria. *Nigerian journal of Medicine*, *17*(3), 304-309.

Osuji, L. C., & Onojake, C. M. (2004). Trace heavy metals associated with crude oil: A case study of Ebocha-8 Oil-spill-polluted site in Niger Delta, Nigeria. *Chemistry & biodiversity*, *1*(11), 1708-1715.

Oyetibo, G. O., Ilori, M. O., Adebusoye, S. A., Obayori, O. S., & Amund, O. O. (2010). Bacteria with dual resistance to elevated concentrations of heavy metals and antibiotics in Nigerian contaminated systems. *Environmental Monitoring and Assessment*, *168*(1), 305-314.

Pan, Y., Wang, L., Feng, Z., Xu, H., Li, F., Shen, Y., ... & Wang, Q. (2023). Characterisation of SARS-CoV-2 variants in Beijing during 2022: an epidemiological and phylogenetic analysis. *The lancet*, *401*(10377), 664-672.

Radović, J. R., Domínguez, C., Laffont, K., Díez, S., Readman, J. W., Albaigés, J., & Bayona, J. M. (2012). Compositional properties characterizing commonly transported oils and controlling their fate in the marine environment. *Journal of Environmental Monitoring*, *14*(12), 3220-3229.

Tian, T., Wan, J., Song, Q., & Wei, Z. (2019b). Clustering single-cell RNA-seq data with a model-based deep learning approach. *Nature Machine Intelligence*, *1*(4), 191-198.

Tian, Y., Krishnan, D., & Isola, P. (2019a). Contrastive representation distillation. *arXiv preprint arXiv:1910.10699*.

Van der Oost, R., Beyer, J., & Vermeulen, N. P. (2003). Fish bioaccumulation and biomarkers in environmental risk assessment: a review. *Environmental toxicology and pharmacology*, *13*(2), 57-149.

Wang, L. M., Yang, Q. H., Zhang, L., Liu, S. Y., Zhang, P. P., Zhang, X., ... & Li, T. F. (2022). Intravenous immunoglobulin for interstitial lung diseases of anti-melanoma differentiation-associated gene 5-positive dermatomyositis. *Rheumatology*, *61*(9), 3704-3710.

World Health Organization. (2017). Progress on drinking water, sanitation and hygiene: 2017 update and SDG baselines. In *Progress on drinking water, sanitation and hygiene: 2017 update and SDG baselines* (pp. 116-116).

Yakubu, M. (2007). Biodegradation of Lagoma crude oil using pig dung. *African Journal of Biotechnology*, *6*(24).

Zhao, B., & Bilen, H. (2021, July). Dataset condensation with differentiable siamese augmentation. In *International Conference on Machine Learning* (pp. 12674-12685). PMLR.