**Data-Driven Assessment of Soil Heavy Metal Contamination in Joinkrama, Rivers State, Nigeria Using Pollution Indices and Multivariate Analytics**

**Abstract**
This study leverages data science methodologies to quantitatively assess soil contamination in Joinkrama, Ahoada West LGA, Rivers State, Nigeria. Employing a structured geospatial and statistical pipeline, soil samples were collected across stratified depths (0.5–1 meter) from three sites (S1, S2, S3) and a control point (C1). Heavy metal concentrations Cadmium (Cd), Lead (Pb), Nickel (Ni), Chromium (Cr), Copper (Cu), and Zinc (Zn) were measured using Atomic Absorption Spectroscopy (AAS) and Inductively Coupled Plasma Mass Spectrometry (ICP-MS). Data preprocessing included normalization and outlier detection using IQR thresholds. Exploratory Data Analysis (EDA) revealed spatial variability with elevated mean concentrations of Cd (0.03 ± 0.01 mg/kg), Pb (0.10 ± 0.04 mg/kg), and Zn (0.16 ± 0.07 mg/kg), notably in S1 and S3. To quantify contamination levels, pollution indices such as Contamination Factor (CF), Degree of Contamination (Cdeg), Pollution Load Index (PLI), Geo-accumulation Index (Igeo), and Potential Ecological Risk Index (PERI) were computed programmatically, revealing high contamination clusters with PERI > 150 in hotspot zones. Dimensionality reduction via Principal Component Analysis (PCA) identified Cd, Pb, and Zn as key anthropogenic signatures, corroborated by Hierarchical Cluster Analysis (HCA) using Ward’s method and Euclidean distance metrics. These contaminants were spatially correlated with known industrial activities, including gas flaring and agrochemical deposition. The integration of environmental data science techniques underscores the utility of computational pollution modeling for environmental risk profiling and decision support. This work presents a replicable, scalable analytical framework for soil quality surveillance and spatial prioritization of remediation in oil-impacted regions of the Niger Delta.

**Keywords:** Soil contamination, Heavy metals, Data science, PCA, Pollution indices, Cluster analysis, Environmental modeling, Python, Spatial analytics

## **1. Introduction**

Soil contamination has become a widespread environmental concern globally, particularly in regions experiencing rapid industrialization and extractive activities. Heavy metals, hydrocarbons, and synthetic chemicals introduced into soils through various anthropogenic means compromise ecosystem services, agricultural productivity, and human health [1]. Among contaminants, heavy metals are of particular concern due to their non-biodegradable nature, potential toxicity even at trace levels, and long-term ecological persistence. In developing countries, these risks are exacerbated by weak environmental regulation and limited monitoring frameworks [2].

The Niger Delta of Nigeria represents a paradigmatic case of such environmental vulnerability. As a globally recognized oil-producing region, it has been the site of extensive oil exploration, gas flaring, and industrial waste discharge, often with minimal oversight [3]. Joinkrama, a rural community in Ahoada West LGA of Rivers State, lies at the epicenter of this dynamic. Oil spills, gas flaring, agricultural runoff, and improper waste disposal have resulted in significant degradation of local soil quality, evidenced by the presence of heavy metals. These metals pose serious environmental and public health risks, especially through bioaccumulation in the food chain and leaching into groundwater aquifers [4].

Despite Joinkrama’s ecological and public health significance, there is a dearth of localized and data-driven studies evaluating its soil quality. Most regional assessments generalize contamination across the Niger Delta, with few studies disaggregating data by specific communities or using standardized pollution indices [5]. This lack of granularity impedes effective policy responses and remediation planning.

To address this gap, this study employs a suite of quantitative indices including the Contamination Factor (CF), Pollution Load Index (PLI), Geo-accumulation Index (Igeo), and Potential Ecological Risk Index (PERI) to assess the spatial and compositional characteristics of soil contamination in Joinkrama. These indices are grounded in established geochemical baselines and allow comparative assessments across sites and regions [6]. Figure 1 presents the conceptual framework of contamination evaluation applied in this study.

### 1.2 Problem Statement

The economic prosperity of the Niger Delta has paradoxically fostered one of the most environmentally degraded ecosystems in Nigeria. Decades of unregulated oil and gas extraction have led to widespread pollution, including frequent oil spills, unmitigated gas flaring, and the leaching of heavy metals into terrestrial and aquatic systems [7]. Joinkrama, a rural settlement within this landscape, has experienced repeated exposure to these pollutants, yet remains understudied in environmental health literature.

Cadmium, lead, and chromium found in effluents from oil production, metal refining, and pesticide use are especially toxic, even at low concentrations [8]. Chronic exposure to these metals can result in kidney failure, neurotoxicity, carcinogenicity, and developmental issues in children [9]. In Joinkrama, agricultural productivity has visibly declined, and residents increasingly report health complications possibly linked to long-term soil contamination.

Despite these realities, localized, data-driven assessments using robust scientific tools remain scarce. As a result, policymakers and local authorities are unable to formulate evidence-based interventions or demand corporate environmental accountability. This study seeks to fill that gap by providing an empirical evaluation of the extent, severity, and ecological implications of soil contamination in Joinkrama using pollution indices.

### 1.3 Aim and Objectives of the Study

The primary aim of this study is to evaluate the extent of heavy metal contamination in the sediments of Joinkrama using standardized pollution indices and multivariate analytical tools.

The specific objectives are:

1. To quantify the concentrations of selected heavy metals (Cd, Pb, Ni, Cr, Cu, Zn) in sediment samples from Joinkrama;
2. To compute pollution indices such as the Contamination Factor (CF), Pollution Load Index (PLI), Geo-Accumulation Index (Igeo), and Potential Ecological Risk Index (PERI);
3. To identify spatial variability and site-specific contamination risks;
4. To determine the ecological and health risk implications of the contamination using multivariate statistical tools including Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA);
5. To recommend appropriate monitoring and remediation strategies based on empirical findings.

3. Materials and Methods

 **Study Area**

Joinkrama is a semi-urban community located in Ahoada West Local Government Area (LGA) of Rivers State, within Nigeria’s oil-rich Niger Delta region. Geographically, it lies between latitudes 4°58′00″ and 5°09′11″ North, and longitudes 6°26′00″ and 6°32′28″ East. The settlement is flanked by several other communities including Akinima, Mbiama, Igovia, Better Land, Okaika, and Okparaki, all of which are situated on the eastern flank of the Orashi River a prominent tributary of the Niger River system [10]. Joinkrama and its neighboring communities depend on agriculture, fishing, and small-scale trading as primary sources of livelihood, although these have increasingly been compromised by industrial activities.



Figure 1: Location map of study area showing sampling points.

Figure 1 shows the location of Joinkrama and surrounding settlements in relation to major oil and gas installations operated by multinational corporations such as the Shell Petroleum Development Company (SPDC) and Nigeria Agip Oil Company (NAOC). These companies maintain flow stations and pipeline networks across Ahoada West, contributing both to the area’s economic relevance and environmental vulnerability.

The area’s proximity to active oil facilities makes it highly susceptible to environmental degradation from oil spills, gas flaring, and improper waste discharge. The Orashi River, which passes through Joinkrama, plays a central role in the community’s water supply for domestic use, irrigation, and fishing. However, it has also become a potential conduit for pollutant dispersion, especially during the rainy season when flooding occurs [11], [12].

Joinkrama experiences a tropical monsoon climate with two distinct seasons: the rainy season (March–October) and the dry season (November–February). The annual rainfall ranges from 1,500 mm to over 3,500 mm, while temperatures typically oscillate between 24°C and 32°C (Rowland, 2021) [13]. These climatic conditions, particularly high rainfall, enhance the leaching and surface runoff of contaminants, increasing the spatial spread of pollutants such as heavy metals into water bodies and surrounding soils [14]. This makes sediment sampling in flood-prone areas critical for assessing contamination [15], [16].

Vegetation in Joinkrama comprises freshwater swamp forests, lowland rainforests, and mangrove ecosystems. However, recurring contamination events have begun to alter these vegetative patterns, with reduced biodiversity and changes in soil texture reported in prior environmental impact assessments [13], [3]. Soils in the area range from sandy to loamy-silt, with greyish to brown coloration depending on depth and proximity to the river. These soil types influence the retention and mobility of contaminants, particularly heavy metals, which can persist over long periods. The geomorphological and hydrological features of Joinkrama, including its low elevation, poorly drained soil, and proximity to perennial water bodies, make it an ideal case for studying sediment-bound contamination [12], [3]. As supported by Figure 1 (Conceptual Framework for Contamination Evaluation), these environmental attributes guided the research design and stratified sampling approach used in this study.

### 3.1 Research Design and Sampling Framework

This study adopted a quantitative, data-driven design integrating stratified sediment sampling, laboratory analysis, and statistical modeling (Figure 1). The study area Joinkrama, Ahoada West LGA was stratified into three sampling sites (S1, S2, S3) based on spatial accessibility, observed contamination symptoms (e.g., oil sheen, vegetation stress), and proximity to known oil infrastructure. A fourth site (C1) located upstream and outside the suspected pollution zone was selected as a control. The spatial coordinates of all sites were recorded using a Garmin eTrex GPS. Sediment samples were collected during the dry season (February 2025) to minimize flood-induced dilution and ensure better accessibility. At each site, triplicate sediment samples were collected using a stainless-steel hand auger at a depth range of 0.5–1.0 meters to capture near-surface contamination. Samples were immediately transferred into polyethylene bags, sealed, labeled, and placed in an ice-cooled box to prevent degradation [17]. The workflow diagram summarizing the methodology used in the research (Figure 2).



Figure 2: Workflow diragam

### 3.2 Laboratory Procedures

#### 3.2.1 Sample Preservation and Pre-treatment

Upon arrival at the laboratory, samples were air-dried at room temperature for 72 hours and sieved through a 2 mm mesh to remove gravel and organic debris. Each homogenized sample was stored in an airtight polyethylene container until digestion. All glassware and instruments were acid-washed to eliminate residual contaminants.

#### 3.2.2 Chemical Digestion and Metal Quantification

Heavy metals (Cd, Pb, Ni, Cr, Cu, Zn) were extracted using the standard aqua regia digestion method, which involves a mixture of nitric acid (HNO₃) and hydrochloric acid (HCl) in a 3:1 ratio. The digestate was filtered and analyzed using Atomic Absorption Spectroscopy (AAS) and Inductively Coupled Plasma Mass Spectrometry (ICP-MS) for cross-validation. AAS was used for Pb, Zn, Cu, and Cr, while ICP-MS was preferred for the trace quantification of Cd and Ni due to its superior detection limits [18].

Certified reference materials (CRM) and reagent blanks were processed concurrently to ensure analytical precision and accuracy. Recovery rates ranged between 94% and 105%, confirming method reliability.

### 3.3 Data Analysis

#### 3.3.1 Descriptive and Inferential Statistics

The metal concentration data were subjected to basic descriptive statistics (mean, standard deviation) and one-way ANOVA to test for site-wise differences using IBM SPSS v26.

### 3.4 Pollution Indices

To evaluate contamination severity, the following pollution indices were calculated:

#### 3.4.1 Contamination Factor (CF)

CF is defined as the ratio of the concentration of a given metal in a sample to its concentration in the control sample (background level):

$$CF=\frac{C\_{m}}{C\_{b}}$$

Where $C\_{m}$ is the measured concentration and $C\_{b}$ is the background concentration from C1. CF values are categorized as: <1 (low), 1–3 (moderate), 3–6 (considerable), >6 (very high) [19], [20]

#### 3.4.2 Degree of Contamination (Cdeg)

Cdeg provides a cumulative contamination value across all six metals:

$$Cdeg=∑CF\_{i}$$

Categorized as: <8 (low), 8–16 (moderate), 16–32 (considerable), ≥32 (very high) [20]. Table 1.

#### 3.4.3 Pollution Load Index (PLI)

PLI measures the overall level of metal contamination using the geometric mean of CFs:

$$PLI=\left(CF\_{1}×CF\_{2}×...×CF\_{n}\right)^{1/n}$$

A PLI > 1 indicates pollution; PLI = 1 is baseline [21].

#### 3.4.4 Geo-Accumulation Index (Igeo)

Igeo quantifies metal enrichment compared to background levels using:

$$I\_{geo}=log\_{2}\left(\frac{C\_{n}}{1.5×B\_{n}}\right)$$

Where $C\_{n}$ is the metal concentration and $B\_{n}$ is the background value [22]. The factor 1.5 accounts for natural variability. Igeo values >2 suggest moderate to heavy contamination.

#### 3.4.5 Potential Ecological Risk Index (PERI)

PERI evaluates ecological toxicity by combining CF with a metal-specific toxic response factor (Tr). The formula is:

$$PERI=∑\left(CF\_{i}×Tr\_{i}\right)$$

Where Tr values are: Cd = 30, Pb = 5, Ni = 5, Cr = 2, Cu = 5, Zn = 1 [20]. S1 and S3 had PERI values of 189, indicating high ecological risk (PERI ≥ 160)

Table 1: Categorization of Contamination Based on Pollution Indices

|  |  |  |
| --- | --- | --- |
| Pollution Index | Category | Contamination Level |
| Contamination Factor (CF) | CF < 1 | Low contamination |
| 1 ≤ CF < 3 | Moderate contamination |
| 3 ≤ CF < 6 | Considerable contamination |
| CF ≥ 6 | Very high contamination |
| Degree of Contamination (Cdeg)    | Cdeg < 8 | Low degree of contamination |
| 8 ≤ Cdeg < 16 | Moderate degree of contamination |
| 16 ≤ Cdeg < 32 | Considerable degree of contamination |
| Cdeg ≥ 32 | Very high degree of contamination |
| Pollution Load Index (PLI)   | PLI = 1 | Baseline level of pollution |
| PLI < 1 | No pollution |
| PLI > 1 | Pollution |
| Geo-Accumulation Index (Igeo)       | Igeo ≤ 0 | Uncontaminated |
| 0 < Igeo ≤ 1 | Uncontaminated to moderately contaminated |
| 1 < Igeo ≤ 2 | Moderately contaminated |
| 2 < Igeo ≤ 3 | Moderately to heavily contaminated |
| 3 < Igeo ≤ 4 | Heavily contaminated |
| 4 < Igeo ≤ 5 | Heavily to extremely contaminated |
| Igeo > 5 | Extremely contaminated |
| Potential Ecological Risk Index (PERI)    | PERI < 40 | Low potential ecological risk |
| 40 ≤ PERI < 80 | Moderate potential ecological risk |
| 80 ≤ PERI < 160 | Considerable potential ecological risk |
| PERI ≥ 160 | High potential ecological risk |

### **3.5 Multivariate Analysis**

Multivariate statistical techniques Principal Component Analysis (PCA), Hierarchical Cluster Analysis (HCA), and Pearson’s Correlation were employed to analyze relationships among environmental variables, reduce data complexity, and identify latent pollution sources [23].

#### **3.5.1 Principal Component Analysis (PCA)**

PCA was applied to standardized data (mean = 0, standard deviation = 1) to reduce dimensionality and extract principal components (PCs) explaining the highest variance. The correlation matrix was used, and components with eigenvalues ≥ 1 were retained based on the Kaiser criterion. Varimax rotation was applied for orthogonal loading clarity. The principal components were computed as:

$$PC\_{i}=a\_{i1}x\_{1}+a\_{i2}x\_{2}+\cdots +a\_{in}x\_{n}$$

Where $a\_{ij}$ are component loadings and $x\_{j}$ are standardized variables. KMO and Bartlett’s test were used to assess sampling adequacy.

#### **3.5.2 Hierarchical Cluster Analysis (HCA)**

HCA grouped variables/samples based on similarity using Ward’s method and Euclidean distance:

$$d\_{ij}=\sqrt{\sum\_{k=1}^{n}\left(x\_{ik}-x\_{jk}\right)^{2}}$$

Clusters were visualized via a dendrogram to identify. HCA aided in corroborating PCA-derived groupings.

#### **3.5.3 Pearson’s Correlation Matrix**

Pearson’s correlation coefficient (r) assessed linear associations between variables:

$$r=\frac{∑\left(x\_{i}-\overset{ˉ}{x}\right)\left(y\_{i}-\overset{ˉ}{y}\right)}{\sqrt{∑\left(x\_{i}-\overset{ˉ}{x}\right)^{2}∑\left(y\_{i}-\overset{ˉ}{y}\right)^{2}}}$$

Where $\overset{ˉ}{x}$ and $\overset{ˉ}{y}$ are mean values. Correlation values range from -1 (perfect inverse) to +1 (perfect direct correlation). Statistically significant values (p < 0.05) were interpreted as potential indicator relationships

## 4.1 Results

### Heavy Metal Concentrations in Sediment Samples

The concentrations of cadmium (Cd), lead (Pb), nickel (Ni), chromium (Cr), copper (Cu), and zinc (Zn) in sediment samples from Joinkrama (S1, S2, S3) and a control site (C1) are presented in Table 2. Results revealed clear site-specific variations, with elevated concentrations observed in S1 and S3 compared to S2 and the control. Specifically, Cd and Pb displayed the highest values in S1 and S3 (0.04 mg/kg and 0.14 mg/kg, respectively), exceeding geochemical background thresholds.

Table 2: Descriptive Statistics of Heavy Metal Concentrations in Sediment Samples

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Latitude | Longitude | Sample Code | Cd (mg/kg) | Pb (mg/kg) | Ni (mg/kg) | Cr (mg/kg) | Cu (mg/kg) | Zn (mg/kg) |
| 5.18198 | 6.502575 | S 1(0.5-1m) | 0.04 | 0.14 | 0.02 | 0.02 | 0.02 | 0.22 |
| 5.179673 | 6.499896 | S 2 (0.5-1m) | 0.02 | 0.05 | 0.02 | 0.02 | 0.03 | 0.03 |
| 5.178405 | 6.496008 | S 3 (0.5-1m) | 0.04 | 0.14 | 0.02 | 0.02 | 0.03 | 0.21 |
| 5.177672 | 6.514063 | Control(C1) (0.5-1m) | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 | 0.01 |

The spatial distribution of these metals is graphically illustrated in Figure 3. shows the cadmium distribution, indicating peak concentrations at S1 and S3, likely due to industrial runoff or phosphate fertilizer contamination. Figure 3 highlights lead levels, which follow the same spatial pattern and are suggestive of vehicular emissions or industrial discharges near SPDC facilities. While Ni, Cr, and Cu remained relatively consistent across all sites, Zn showed slightly elevated levels in S1 (0.22 mg/kg) and S3 (0.21 mg/kg), potentially indicating contamination from agrochemical inputs or urban runoff.



Figure 3: Heavy metals concentration in soil samples

### Contamination Factor and Degree of Contamination

Contamination Factor (CF) values were computed by comparing site concentrations to those at the control (C1). Results show very high CF values for Cd (CF = 4.0) and Pb (CF = 7.0) at S1 and S3, placing them in the “considerable to very high contamination” category (Table 1). The computed CF values for each element across sites are summarized in Table 3.

Table 3: Contamination Factors (CF) for Heavy Metals by Site

| Site | Cd | Pb | Ni | Cr | Cu | Zn |
| --- | --- | --- | --- | --- | --- | --- |
| S1 | 4 | 7 | 2 | 2 | 2 | 22 |
| S2 | 2 | 2.5 | 2 | 2 | 3 | 3 |
| S3 | 4 | 7 | 2 | 2 | 3 | 21 |

The Degree of Contamination (Cdeg) values were also computed, showing S1 and S3 with scores of 39 each well above the “very high” contamination threshold (Cdeg ≥ 32), while S2 had a moderate score of 14.5 (Table 4).

Table 4: Degree of Contamination (Cdeg) for Sampling Sites

| Sample | Cdeg |
| --- | --- |
| S1 | 39 |
| S2 | 14.5 |
| S3 | 39 |

These values suggest that S1 and S3 are ecological hotspots for contamination, with cumulative metal loads far exceeding safe limits, particularly due to high Cd and Pb enrichment.

### Pollution Load Index (PLI)

The Pollution Load Index (PLI) was used to evaluate overall site pollution, with values >1 indicating significant contamination. The PLI results show values of 4.43 and 4.36 for S1 and S3, respectively both far above the pollution threshold. S2 had a lower PLI of 2.18 but still exceeded the baseline level (Table 5).

Table 5: Pollution Load Index (PLI) per Sampling Site

| Site | PLI |
| --- | --- |
| S1 | 4.43 |
| S2 | 2.18 |
| S3 | 4.36 |

These findings reinforce the inference that S1 and S3 are highly polluted, with contributions likely from oil production waste and industrial runoff.

### 4.1.4 Geo-Accumulation Index (Igeo)

The Igeo values for Cd and Pb were particularly elevated at S1 and S3, falling within Class 3–4 ("moderately to heavily polluted") and Class 4–5 ("heavily polluted") respectively (Table 1). The Igeo for Zn at S1 reached 5.83, indicating "heavily to extremely contaminated" levels. These values are graphically displayed in Figure 4.



Figure 4: Geo-Accumulation Index (Igeo) for Metals across Sites

The average Igeo values suggest that Cd and Pb are the dominant pollutants and likely originate from similar sources [22].

### 4.1.5 Potential Ecological Risk Index (PERI)

Table 6 displays PERI values for each site. Cd contributed the most to the ecological risk score due to its high toxicity coefficient (Tr = 30). The total PERI for S1 and S3 reached 189, classifying both as “high ecological risk zones” [20]. S2 scored 65.5, indicating moderate risk.

Table 6: Potential Ecological Risk Index (PERI) by Site

| Site | PERI |
| --- | --- |
| S1 | 189 |
| S2 | 65.5 |
| S3 | 189 |

These results warrant urgent ecological attention and suggest the need for immediate risk management strategies in affected sites.

### 4.1.6 Principal Component Analysis (PCA)

Principal Component Analysis revealed that PC1 accounted for 83.41% of the total variance and was heavily loaded with Cd, Pb, and Zn, indicating a strong anthropogenic signature. PC2 explained 15.79% of the variance and captured geogenic signals associated with Ni and Cr. The PCA biplot in Figure 5 shows a clear separation between the control site and the contaminated samples.



Figure 5: PCA Biplot Showing Relationships Among Metals and Sites

This pattern suggests that the dominant sources of Cd, Pb, and Zn are likely oil-related industrial activities and urban runoff, while Ni and Cr may be partially geogenic.

### 4.1.7 Cluster and Correlation Analysis

Hierarchical Cluster Analysis grouped S1, S2, and S3 into one cluster and C1 into another, reflecting distinct contamination profiles (Figure 6). This clustering validates earlier statistical trends and affirms the higher contamination burden of the Joinkrama sediment samples.



Figure 6: Dendrogram from Hierarchical Cluster Analysis

The correlation matrix (Figure 7) shows strong positive correlations between Cd and Pb (r = 0.92), and Zn and Cu (r = 0.88), supporting the hypothesis of shared contamination sources.

 

Figure 7: Pearson Correlation Matrix Heatmap of Heavy Metals

## 4.2 Discussion

The sediment samples analyzed in this study reveal substantial heavy metal contamination in Joinkrama, with a clear distinction between impacted sites (S1, S2, S3) and the control site (C1). The elevated levels of cadmium (Cd), lead (Pb), and zinc (Zn), particularly at sites S1 and S3 (Table 2), indicate localized pollution hotspots. These findings align with earlier observations from similar oil-producing environments in the Niger Delta, where industrial discharge, gas flaring, and oil spills were identified as primary sources of soil contamination [24], [5].

The significant spatial variation in metal concentrations especially for Cd and Pb reflects both anthropogenic inputs and environmental factors such as drainage, runoff patterns, and topography. This supports earlier assertions by Liang et al. and Onwuka et al. [25], [18] that contamination in oil-rich regions is not homogenous but highly site-specific.

Pollution indices provide critical insights beyond raw concentration values. The Contamination Factor (CF) values for Cd and Pb at S1 and S3 (CF > 4 and > 7, respectively) indicate very high contamination (Table 3). These values suggest that anthropogenic enrichment is considerable and sustained, likely due to long-term industrial emissions and leaching of pollutants from upstream sources. Similar findings were reported by Adebayo et al. and Ogundele et al. [17], [9] who linked high CF values to gas flaring and unregulated waste disposal.

The Degree of Contamination (Cdeg), which aggregates the CFs across metals, reached 39.0 for both S1 and S3 (Table 3), placing these sites within the “very high contamination” category per Hakanson [20]. This cumulative contamination confirms that localized pollution in Joinkrama is not limited to one metal but involves multiple heavy metals acting synergistically.

In addition, the Pollution Load Index (PLI) values above 4 (Table 5) underscore the severity of contamination, indicating that metal enrichment has far exceeded baseline levels. Such high PLI values have been previously associated with sediment samples located near petroleum infrastructure [26]. As Tomlinson et al. [21] argued, PLI values above 1 denote pollution, but values above 3–4 indicate substantial ecological stress. In the context of Joinkrama, the PLI reinforces the notion of severe, multi-elemental pollution particularly in areas proximal to gas flaring sites and oil pipelines.

The Geo-Accumulation Index (Igeo) results further confirm pollution dominance by Cd, Pb, and Zn, with values at S1 and S3 placing them within Class 3–5 (“moderately to heavily contaminated” to “heavily contaminated”) (Figure 4). For instance, Pb reached an Igeo of 3.91 at S1 and S3, indicating heavy enrichment due to anthropogenic input. This is consistent with Müller’s [22] classification and similar to findings by Adewuyi et al. [6], who noted that Pb and Cd are frequently elevated in areas with historical oil activities.

Notably, the Igeo for Zn at S1 exceeded 5.83, classifying it as “heavily to extremely contaminated.” Elevated Zn may originate from corrosion of metal pipelines, improper disposal of zinc-based pesticides, and vehicular runoff, as suggested by Chinedu et al. [27]. The high Igeo reinforces the urgent need for source apportionment and environmental audits by regulatory bodies.

The Potential Ecological Risk Index (PERI) results underscore the toxicological implications of the contamination. With PERI scores of 189 for S1 and S3 (Table 6), these sites qualify as “high ecological risk zones,” particularly due to the elevated toxic response coefficient of Cd (Tr = 30). Zhao et al. [28] emphasized the ecological implications of such PERI values, especially in aquatic environments where metal mobility increases during the wet season.

Cd’s high toxicity and bioavailability mean that even low concentrations can result in substantial risk, particularly in water-logged environments like Joinkrama, where leaching into shallow aquifers is plausible. High ecological risk, as indicated by the PERI, calls for immediate remediation strategies, including soil washing, bioremediation, or phytostabilization [29].

Interestingly, while Cu, Ni, and Cr displayed lower PERI contributions, their cumulative presence still contributes to long-term sublethal toxicity and biodiversity stress. This echoes findings by Oluwatomilola [30] in similar riverine systems across the Niger Delta.

Principal Component Analysis (PCA) explained over 99% of the variance across the two main components (PC1 = 83.41%, PC2 = 15.79%), with Cd, Pb, and Zn loading heavily on PC1 (Figure 10). This suggests a common anthropogenic source, possibly linked to oil exploration and agrochemical activities. Similar PCA structures were reported by Eteh et al. [31], where metals clustered according to industrial proximity.

The separation of control sample C1 in the PCA biplot (Figure 5) and the dendrogram (Figure 6) further validates the spatial contamination gradient. Hierarchical Cluster Analysis (HCA) grouped S1, S2, and S3 together, suggesting a shared contamination profile, while C1 formed a distinct cluster due to its low metal concentrations. These multivariate results highlight the utility of PCA and HCA in ecological risk mapping and zoning, as noted by Negahban et al. [32].

The Pearson correlation matrix (Figure 7) reveals strong positive correlations between Cd and Pb (r = 0.92), as well as Cu and Zn (r = 0.88). This suggests either co-mobilization or similar emission sources. The high Cd–Pb correlation could be attributed to their co-occurrence in oil exploration wastes, batteries, and industrial emissions [33]. Cu–Zn correlation may indicate vehicular emissions, industrial corrosion, or pesticide runoff. These correlations are essential for pollution fingerprinting and targeted intervention planning.

The results of this study are consistent with previous regional findings. For example, Eteh et al. [12] reported comparable PERI and PLI values in sediment studies across Bayelsa and Delta States, also implicating Cd and Pb as the dominant ecological risks.

Comparative assessments with global standards (e.g., UNEP, WHO) further emphasize the hazard levels in Joinkrama. Most Cd and Pb concentrations in S1 and S3 surpass WHO/FAO permissible limits for agricultural soils (0.01–0.03 mg/kg for Cd; 0.05–0.10 mg/kg for Pb), indicating possible risks to food security if agricultural activities continue unregulated in these areas [34].

### Environmental and Policy Implications

The presence of high ecological risk zones within a residential and agriculturally active community like Joinkrama presents urgent environmental management challenges. Current findings underscore the need for site-specific remediation and stringent environmental monitoring protocols. Regulatory institutions such as the National Environmental Standards and Regulations Enforcement Agency (NESREA) and the Nigerian Upstream Petroleum Regulatory Commission (NUPRC) must enhance compliance monitoring of oil and gas operators.

Moreover, the integration of pollution indices with geospatial data could assist in zoning the community into risk categories, allowing for spatially targeted interventions and public health advisories. Community engagement and environmental education are also necessary to build local capacity for early detection and response.

### Limitations and Recommendations

While the present study provides a robust baseline of sediment contamination in Joinkrama, it is not without limitations. Seasonal variations were not captured, which may influence metal mobility and concentrations. Future studies should adopt longitudinal sampling across dry and wet seasons for a more comprehensive contamination profile. Additionally, bioavailability assays and soil–plant transfer assessments would strengthen health risk projections.

Further work should also include Total Petroleum Hydrocarbon (TPH) and Polycyclic Aromatic Hydrocarbons (PAHs) to assess the combined effects of organic and inorganic pollutants a dual threat commonly observed in the Niger Delta [35].

## **5. Conclusion and Recommendations**

### 5.1 Conclusion

This study has provided a comprehensive assessment of soil contamination in Joinkrama, Ahoada West Local Government Area, Rivers State, Nigeria, using pollution indices, statistical tools, and multivariate techniques. The analysis confirmed elevated concentrations of heavy metals particularly cadmium (Cd), lead (Pb), and zinc (Zn)—across two sampling sites (S1 and S3), in contrast to the lower contamination levels observed in S2 and the control site (C1). The Pollution Load Index (PLI), Geo-accumulation Index (Igeo), and Potential Ecological Risk Index (PERI) all classified these sites as severely contaminated and ecologically hazardous (Figures 3 and 4; Tables 2 and 4).

The spatial patterning of contamination, as revealed by the heatmap (Figure 8) and the cluster analysis (Figure 6), indicates that anthropogenic sources—particularly oil exploration, gas flaring, and industrial discharges—are responsible for the observed pollution. This conclusion is further supported by PCA and correlation matrix analysis (Figures 5 and 7), which highlighted the shared source signatures of Cd, Pb, and Zn. These results are consistent with earlier studies in oil-producing regions of the Niger Delta. [11],[14], [36].

The health and ecological implications of these findings are profound. Cd and Pb, known for their high toxicity and bioaccumulative properties, pose significant risks to soil fertility, aquatic systems, and human health through food chain transfer and groundwater infiltration [4], [37]. Immediate mitigation efforts and long-term monitoring are essential to reduce exposure risks and protect public health.

### 5.2 Recommendations

Based on the data-driven insights of this study, the following recommendations are proposed:

1. **Implementation of Remediation Strategies:** Contaminated hotspots (S1 and S3) should be prioritized for bioremediation and soil washing interventions. Indigenous microbial techniques and phytoremediation using local vegetation could be tested for site-specific restoration [38].
2. **Regulatory Enforcement and Pollution Audits:** Environmental authorities such as the Federal Ministry of Environment (FME), National Oil Spill Detection and Response Agency (NOSDRA), and Nigerian Upstream Petroleum Regulatory Commission (NUPRC) should mandate periodic effluent audits for operators in the region. Violations should attract penalties and mandatory remediation [12], [7]
3. **Community-Based Environmental Monitoring:** Local communities should be trained and equipped with GPS and simple test kits for participatory environmental monitoring. Community early warning systems could be integrated with mobile alert platforms to report spill events and flaring anomalies in real-time.
4. **Integration of Pollution Indices into Environmental Impact Assessments (EIAs):** National and regional EIAs should incorporate pollution indices such as PLI, Igeo, and PERI as core metrics for evaluating project sustainability in oil-producing regions.
5. **Public Health Intervention:** Health assessments targeting vulnerable populations, especially children and pregnant women, should be conducted periodically. Mobile clinics and environmental health outreach programs are necessary to monitor long-term exposure effects.
6. **Expansion of Research:** Future studies should increase the number of sampling sites and integrate seasonal analyses. Isotopic fingerprinting and geospatial risk modeling can enhance source attribution and predictive mapping for early remediation planning.

Here is the **Declarations** section for your manuscript:

## **Declarations**

###

### **Ethical Approval**

All research procedures and environmental sample collections were conducted in line with ethical standards. No human or animal subjects were directly involved, and no endangered species or protected sites were disturbed during fieldwork.

### **Consent to Participate**

Not applicable. No human participants were involved in this study.

### **Consent for Publication**

All authors reviewed and approved the final version of this manuscript and consented to its publication in a peer-reviewed journal.

### **Availability of Data and Materials**

All data generated or analyzed during this study are available from the corresponding author upon reasonable request. Geospatial data and heavy metal concentration tables are archived digitally and may be shared for replication purposes.

**COMPETING INTERESTS DISCLAIMER:**

Authors have declared that they have no known competing financial interests OR non-financial interests OR personal relationships that could have appeared to influence the work reported in this paper.

Disclaimer (Artificial intelligence)

Option 1:

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

## **References**

1. Alengebawy, A., Abdelkhalek, S. T., Qureshi, S. R., & Wang, M. Q. (2021). Heavy Metals and Pesticides Toxicity in Agricultural Soil and Plants: Ecological Risks and Human Health Implications. Toxics, 9(3), 42. <https://doi.org/10.3390/toxics9030042>
2. Ali, H., Khan, E., & Sajad, M. A. (2021). Phytoremediation of heavy metals—concepts and applications. Chemosphere, 91(7), 869-881
3. Jonathan EL, Imoni O, Chukwuemeka P, Eteh DR. Impact of oil spills on mangrove ecosystem degradation in the Niger Delta using remote sensing and machine learning. Journal of Geography and Cartography. 2025; 8(2): 11707. <https://doi.org/10.24294/jgc11707>
4. Silver, W. T., Young, E., Imomotimi, A. T., Angaye, W. T., & Ebong, C. U. (2024). Heavy metal pollution in gas flare-impacted soils: An assessment using pollution indices in Obunagha, Bayelsa State, Nigeria. *Asian Journal of Applied Chemistry Research, 15*(4), 17–25. <https://doi.org/10.9734/ajacr/2024/v15i4293>
5. Okonofua, T., Ujile, A. A., & Ite, A. E. (2021). Evaluation of soil and water contamination in oil-producing communities of the Niger Delta. International Journal of Environmental Science and Technology, 18(4), 1679-1690
6. Adewuyi, G. O., et al. (2021). Application of pollution indices to assess heavy metal contamination in soils near industrial sites. *Environmental Monitoring and Assessment*, 193(5), 320
7. Ismail, O., & Umukoro, G. (2012). Global impact of gas flaring. *Energy and Power Engineering, 4*(4), 290–302. <https://doi.org/10.4236/epe.2012.44039>
8. Nwachukwu, C. E., & Igbinoba, E. (2021). Heavy metal pollution in soils: Sources, toxicity, and management. Soil Science and Plant Nutrition, 67(2), 209-221.
9. Ogundele, A. S., et al. (2022). Ecological risks of heavy metals in soils near gas flaring sites in the Niger Delta: An ERI-based approach. *Environmental Toxicology and Pharmacology*, 84, 103702
10. Jonathan, L. E., & Charles, A. U. (2025). Shoreline erosion and accretion analysis of the Orashi River, Rivers State, Nigeria: A geospatial and machine learning approach. *Asian Journal of Geographical Research, 8*(2), 27–44. <https://doi.org/10.9734/ajgr/2025/v8i2260>
11. Okpobiri, O., Akajiaku, C. U., Eteh, D. R., & Moses, P. (2025). Using machine learning and GIS to monitor sandbars along the River Niger in the Niger Delta, Nigeria. *International Journal of Environment and Climate Change, 15*(2), 182–203. <https://doi.org/10.9734/ijecc/2025/v15i24721>
12. Eteh, D.R., Japheth, B.R., Akajiaku, C.U. et al.(2025) Assessing the impact of climate change on flood patterns in downstream Nigeria using machine learning and geospatial techniques (2018–2024). Discov Geosci 3, 76 (2025). <https://doi.org/10.1007/s44288-025-00178-7>
13. Rowland, E. D., & Okpobiri, O. (2021). Floodplain Mapping and Risks Assessment of the Orashi River Using Remote Sensing and GIS in the Niger Delta Region, Nigeria. *Journal of Geographical Research/Maǧallaẗ Al-buḥūṯ Al-Ǧuġrāfiyyaẗ*, *4*(2).
14. Egai, A. O., Douglas, R. K., & Fou, A. (2022). Assessing the extent of heavy metal contamination in crude oil-impacted soils in the Niger Delta, Nigeria using geochemical indicators. *Global Journal of Earth and Environmental Science, 7*(1), 1–9. <https://doi.org/10.31248/GJEES2020.067>
15. Abaye, E. O., Ayowei, A. H., Akajiaku, U. C., & Eteh, D. R. (2022). Implication of landscape changes using Google Earth historical imagery in Yenagoa, Bayelsa State, Nigeria. *Sumerianz Journal of Scientific Research, 5*(1), 20–31.
16. Oborie, E., Fatunmibi, I., & Otutu, A. O. (2023). Shoreline change assessment in the Orashi River, Rivers State, Nigeria, using the digital shoreline analysis system (DSAS). *Sumerianz Journal of Scientific Research*, *6*(4), 70-77
17. Adebayo, T. S., & Kirikkaleli, D. (2021). Impact of renewable energy consumption, globalization, and technological innovation on environmental degradation in Japan: application of wavelet tools. *Environment Development and Sustainability*, *23*(11), 16057–16082. https://doi.org/10.1007/s10668-021-01322-2
18. Onwuka, C., Eboatu, A.N., Ajiwe, V.I.E. and Morah, E.J. (2021) Pollution Studies on Soils from Crude Oil Producing Areas of Rivers State, Niger Delta Region, Nigeria. *Open Access Library Journal*, **8**, 1-17. doi: [10.4236/oalib.1107583](https://doi.org/10.4236/oalib.1107583)
19. Chris, D. I., & Anyanwu, B. O. (2022). Pollution and Potential Ecological Risk Evaluation Associated with Toxic Metals in an Impacted Mangrove Swamp in Niger Delta, Nigeria. *Toxics*, *11*(1), 6. <https://doi.org/10.3390/toxics11010006>
20. Hakanson, L. (1980) An Ecological Risk Index for Aquatic Pollution Control a Sedimentological Approach. Water Research, 14, 975-1001.
[http://dx.doi.org/10.1016/0043-1354(80)90143-8](http://dx.doi.org/10.1016/0043-1354%2880%2990143-8)
21. Tomlinson, D.L., Wilson, J.G., Harris, C.R. and Jeffrey, D.W. (1980) Problems in the Assessment of Heavy-Metal Levels in Estuaries and the Formation of a Pollution Index. Helgoländer Meeresuntersuchungen, 33, 566-575.
<http://dx.doi.org/10.1007/BF02414780>
22. Muller, G. (1969) Index of Geoaccumulation in Sediments of the Rhine River. GeoJournal, 2, 108-118.
23. Shrestha, S., & Kazama, F. (2007). Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. Environmental Modelling & Software, 22(4), 464–475. <https://doi.org/10.1016/j.envsoft.2006.02.001>
24. Osuji, C., & Nduka, N. (2021, March). *Urban Surface Run-Off as a vector for soil contamination in rapidly growing cities in Rivers State, Nigeria | Abstract*. https://iiardjournals.org/abstract.php?j=IJGEM&pn=Urban+Surface+Run-Off+as+a+Vector+for+Soil+Contamination+in+Rapidly+Growing+Cities+in+Rivers+State%2C+Nigeria&id=60749
25. Liang, Zou, J., Liu, S., Chen, X., Kong, X., Zhang, H., & Xu, T. (2022). Seasonal and spatial variation in dissolved heavy metals in Liaodong Bay, China. *International Journal of Environmental Research and Public Health*, *19*(1), 608. <https://doi.org/10.3390/ijerph19010608>
26. Zhang, Y., Guo, Y., Tang, Z., Feng, Y., Zhu, X., Xu, W., Bai, Y., Zhou, G., Xie, Z., & Fang, J. (2021). Patterns of nitrogen and phosphorus pools in terrestrial ecosystems in China. *Earth System Science Data*, *13*(11), 5337–5351.https://doi.org/10.5194/essd-13-5337-2021
27. Chinedu, E. (2022). *Oil Spillage and Heavy Metals Toxicity Risk in the Niger Delta, Nigeria. Journal of Health & Pollution, 8, 1-8. - References - Scientific Research Publishing*. <https://www.scirp.org/reference/referencespapers?referenceid=3308554>
28. Zhao, L. (2021). Submitted for GEOPHYSICS by Zhao et al. (2021) [Dataset]. In *Zenodo (CERN European Organization for Nuclear Research)*. <https://doi.org/10.5281/zenodo.5674412>
29. Kubier, A., Wilkin, R. T., & Pichler, T. (2019). Cadmium in soils and groundwater: A review. *Applied Geochemistry*, *108*, 104388. <https://doi.org/10.1016/j.apgeochem.2019.104388>
30. Oluwatomilola, A. O., & A, P. J. (2025). Assessment of heavy metal pollution in soils of Ifie Community, Warri, Delta State, Nigeria. Global Journal of Pure and Applied Sciences, 31(1), 25–35. <https://doi.org/10.4314/gjpas.v31i1.3>
31. Eteh, D., Egirani, D. E., Shehata, N., Ugwu, I. M., & Opukumo, A. (2020). Exposure, geochemical, and spatial distribution patterns of an oil spill in parts of the Niger Delta Region of Nigeria. *Health and Environment*, *2*(1), 103–117. <https://doi.org/10.25082/he.2021.01.005>
32. Negahban, S., Mokarram, M., Pourghasemi, H. R., & Zhang, H. (2020). Ecological risk potential assessment of heavy metal contaminated soils in Ophiolitic formations. *Environmental Research*, *192*, 110305. <https://doi.org/10.1016/j.envres.2020.110305>
33. Valkova, E., Atanasov, V., Vlaykova, T., Tacheva, T., Zhelyazkova, Y., Dimov, D., & Yakimov, K. (2023). The serum levels of the heavy metals CU, ZN, CD, and PB and progression of COPD—A preliminary study. *International Journal of Environmental Research and Public Health*, *20*(2), 1427. <https://doi.org/10.3390/ijerph20021427>
34. United Nations Environment Programme—UNEP. (2021). In Routledge eBooks (pp. 193–199). <https://doi.org/10.4324/9781003292548-46>
35. Emmanuel, A. E., Chidi, O., & Boma, K. (2021). Assessment of Total Petroleum Hydrocarbons and Polycyclic Aromatic Hydrocarbons in Wastewater from Selected Flow Stations in Agbada II, Rivers State. *Asian Journal of Applied Chemistry Research*, 31–40. <https://doi.org/10.9734/ajacr/2021/v9i430221>
36. Musa, D. G., Oruonye, E. D., Anger, R. T., Ojeh, V. N., & Delphine, D. (2024). Effect of gas flaring on human well-being and environment in Obodo-Ugwa, Ndokwa West, Local Government Area, Delta State, Nigeria. *Journal of Energy and Environmental Sciences, 2*(1). <https://doi.org/10.23880/jeesc-16000108>
37. Jato, D. J., Onyije, F. M., Mgbere, O. O., & Avwioro, G. O. (2024). Exposure to gas flaring among residents of oil-producing communities in Bayelsa State, Niger Delta Region of Nigeria: A cross-sectional study of haematological indices. *J, 7*(4), 472–481. <https://doi.org/10.3390/j7040028>
38. Francis, E. E., Eteh, D. R., & Ebizimo, D. S. (2022). Multispectral imagery for detection and monitoring of vegetation affected by oil spills and migration pattern in Niger Delta Region, Nigeria. *World Journal of Advanced Research and Reviews, 15*(01), 447–458. <https://doi.org/10.30574/wjarr.2022.15.1.0682>