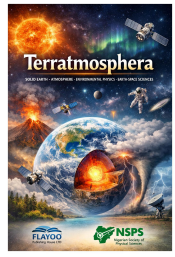
Produced and hosted by FLAYOO Publishing House LTD on behalf of the Nigerian Society of Physical Sciences (NSPS)

Terratmosphera

Solid Earth • Atmosphere • Environment • Space Sciences

https://flayoophl.com/journals/index.php/terratmosphera



Geochemical modelling and ecological risk of heavy metals in industrial agricultural soils: a Python-driven approach from southwestern Nigeria

Charles Olubunmi Ogunkoya[a,*], Gideon Oluyinka Layade[b]

[a]*Department of Physics, Ajayi Crowther University, Oyo, P.M.B. 1066, Oyo, Oyo State, Nigeria*
[b]*Department of Physics, Federal University of Agriculture, Abeokuta, P.M.B. 2240, Abeokuta, Ogun State, Nigeria*

ARTICLE INFO

Article history:
Received: 07 May 2026
Received in revised form: 22 June 2026
Accepted: 23 June 2026
Available online: 09 July 2026

Keywords: Agricultural soil contamination, Ecological risk index, Contamination factor, Human health risk assessment, Python geochemistry
DOI:10.61298/terratmosphera.2026.1.1.370

ABSTRACT

Heavy metal contamination of agricultural soils near the Ota Industrial Estate, southwestern Nigeria, was assessed using a six-module Python pipeline that integrates geochemical indexing, ecological-risk modelling, and human-health risk assessment. Cu, Pb, Zn, and Cd were analysed at 12 sites. Cu, Pb, and Zn were within WHO/FAO permissible limits. Cd concentrations exceeded the WHO/FAO guideline of 0.80 mg/kg at half of the locations (IN1–IN6, up to 3.09 mg/kg), and two locations exceeded the upper limit of 3.0 mg/kg. Contamination-factor analysis classified Pb at the vehicle-workshop zones (IN7–IN9) as considerable to very high, with CF up to 12.61, and classified Cu/Zn as moderate to very high at fabrication and galvanisation zones. The mean ecological risk index (RI) for the estate was 64.5 ± 24.1, indicating low overall risk, although Cd and Pb fell within moderate E_r bands in some sub-zones. Independent recomputation showed that none of five originally proposed metal-pair correlations was significant; however, a strong Pb–Cd correlation was Bonferroni-significant ($r = 0.946$, $p = 0.0012$), indicating a shared anthropogenic source. US EPA RAGS non-carcinogenic hazard indices were below unity for adults (HI = 0.014) and children (HI = 0.126). Carcinogenic risk was not quantified because validated oral slope factors for Cd or Pb are unavailable in IRIS. The principal regulatory basis for clean-up is the Cd exceedance. The open-source pipeline provides a reproducible framework for soil-contamination assessment in West African regulatory contexts.

1. INTRODUCTION

The contamination of agricultural soils by heavy metals in peri-industrial zones is one of the most complex geoenvironmental problems facing rapidly industrialising countries, especially sub-Saharan Africa, where the regulation of industrial effluents is weak and inconsistently enforced [1, 2]. Heavy metals are non-biodegradable and chemically persistent; therefore, they can accumulate indefinitely in soil and enter the human food chain through soil-plant transfer, direct ingestion, dermal contact, and inhalation of fugitive dust. These exposure pathways require si-

*Corresponding author: Tel. No.: +234-807-8744-201.
e-mail: co.ogunkoya@acu.edu.ng (Charles Olubunmi Ogunkoya)

multaneous quantitative assessment [3, 4].

The soil-atmosphere coupling is especially significant in rapidly industrialising regions, where emissions controls are patchy and dry/wet deposition can strongly influence near-surface geochemistry. Heavy metals such as cadmium, lead, and mercury can induce toxicity, including kidney dysfunction and bone damage in adults. In children, lead causes neurodevelopmental impairment, while mercury affects the central nervous system at trace concentrations. These toxicological concerns make cadmium, lead, and mercury important targets for investigation [5, 6].

The Ota Industrial Estate in southwestern Nigeria is a case of conflicting pollution pressures. Ota is one of Nigeria's most diverse industrial hubs, and by the late 1990s it had become the country's third-largest industrial hub [7]. It hosts textile, metal fabrication, plastics production, chemical manufacturing, and battery-recycling industries, all of which are known sources of Cd, Pb, Cu, Zn, and Hg to the environment [8]. These activities are also responsible for the accumulation of heavy metals in soil and the surrounding environment.

Human exposure happens through various pathways due to proximity of smallholder agricultural activities to these facilities. The transition zone between the Precambrian basement and Cretaceous sediments, in which Ota is located (as indicated in Figure 1), adds further geogenic variability to baseline metal concentrations, thus making the attribution of sources very difficult in the absence of robust statistical analysis [9].

Soils impacted by heavy metals are most frequently assessed using the contamination factor (CF) and Hakanson ecological risk index [10, 11]. Nonetheless, human health risk assessments (HHRAs) are often inaccurate, poorly reproducible, and error-prone when computed manually across several sampling locations and metals. In addition, they are rarely coupled, within the same study, with HHRA using US EPA exposure models [12, 13]. Owing to this methodological gap, risk-management decision-making is constrained. In addition, the lack of open-source, validated computational tools for integrated multi-index assessment has prevented consistent application in West African regulatory contexts where software-license costs are prohibitively high [14, 15].

This study seeks to achieve four objectives: (i) to quantify the concentrations of Cu, Pb, Zn, Cd, and Hg in surface agricultural soils; (ii) to compute WHO/FAO compliance, CF, E_r, and RI values by establishing a six-module Python 3 algorithm; (iii) to perform source apportionment using Pearson/Spearman correlation and spatial interpolation; and (iv) to quantify non-carcinogenic HHRA for adult and child receptors via the soil-ingestion pathway.

2. MATERIALS AND METHODS

2.1. STUDY AREA AND GEOLOGICAL SETTING

Ota occupies coordinates 6°40′N–6°67′N and 3°11′E–3°19′E in Ogun State, Southwestern Nigeria, and has a total area of about 878 km^2 [7]. The Ota Industrial Estate has one of the highest concentrations of industries in Ogun State, southwestern Nigeria. The soil-sampling locations in the research area are shown in Figure 1, including prominent roads, the railway line, river, and geological formations (alluvium and coastal plain sands). The geologic area is situated between the basement complexes of Precambrian rocks, consisting of migmatite-gneiss, granite, and schist, and the Cretaceous-to-recent alluvial sedimentary rocks of the Dahomey Basin [9, 16]. The stratigraphic transition governs the natural geochemical background of potentially toxic metals in soils; basement soils tend to show higher Fe, Mn, Ni, and Cr, whereas alluvial sediments have higher organic-matter content and cation-exchange capacity, which can limit metal retention [17]. The plots immediately beside the industrial units support smallholder agriculture, mainly vegetable, cassava, and maize cultivation.

2.2. SOIL SAMPLING AND LABORATORY ANALYSIS

Twelve composite soil samples (0–15 cm) were collected from four sub-zones of the industrial area, IN1–IN12 (three samples from each zone), using a pre-cleaned soil auger. GPS coordinates were recorded for all sample locations. The samples were transported to the laboratory, air-dried at ambient temperature, and sieved through a 2 mm mesh. They were then digested with 4:1 (v/v) HNO3-HClO4 at 180°C. The digests were filtered (0.45 μm) and diluted to 25 mL with 1% HNO3 before being analysed by flame atomic absorption spectrometry (FAAS; PerkinElmer PinAAcle 900T) with deuterium background correction. Quality assurance/quality control (QA/QC) included procedural blanks, certified reference material (NIST SRM 2710a; recoveries 92–107%), and triplicate analysis (RSD < 5%).

The site-specific instrumental limits of detection (LODs) and limits of quantification (LOQs) for Cd from the FAAS calibration run were not retained and could not be reconstructed retrospectively from the reported sample concentrations. The authors recognize this as a study limitation. Because several measured Cd concentrations (for example, IN11) are close to the sensitivity floor for flame AAS in soil reported in the literature, near-detection-limit values should be interpreted with caution. Future analysis of this site should include formal LOD/LOQ determination (for example, 3σ of the calibration blank divided by the calibration slope) as part of routine analytical-method validation, as well as consideration of graphite furnace AAS (GFAAS) for better sensitivity at <1 mg/kg Cd.

2.3. NOVEL PYTHON 3 ANALYTICAL PIPELINE

The primary method of the study consists of a six-module Python 3.11 algorithm written, tested, and validated by the authors. It carries out the complete analysis workflow from raw data to risk output without human intervention. The algorithm was tested on known data before deployment on field data, and each module has its own unit test.

The first module, Data Ingestion and Validation, reads concentration data already expressed in mg/kg from a structured CSV file and flags entries below the detection limit. Module 2 (WHO/FAO Compliance Screening) performs vectorised comparison of concentration arrays against permissible-limit tables, returning site-specific binary flag matrices and summary exceedance counts in a single NumPy operation, thereby avoiding row-by-row manual checking in spreadsheet-like workflows. Module 3 (Contamination Indexing) computes CF, geoaccumulation index (Igeo), pollution load index (PLI), and modified degree of contamination (mCd) for all site-metal combinations us-

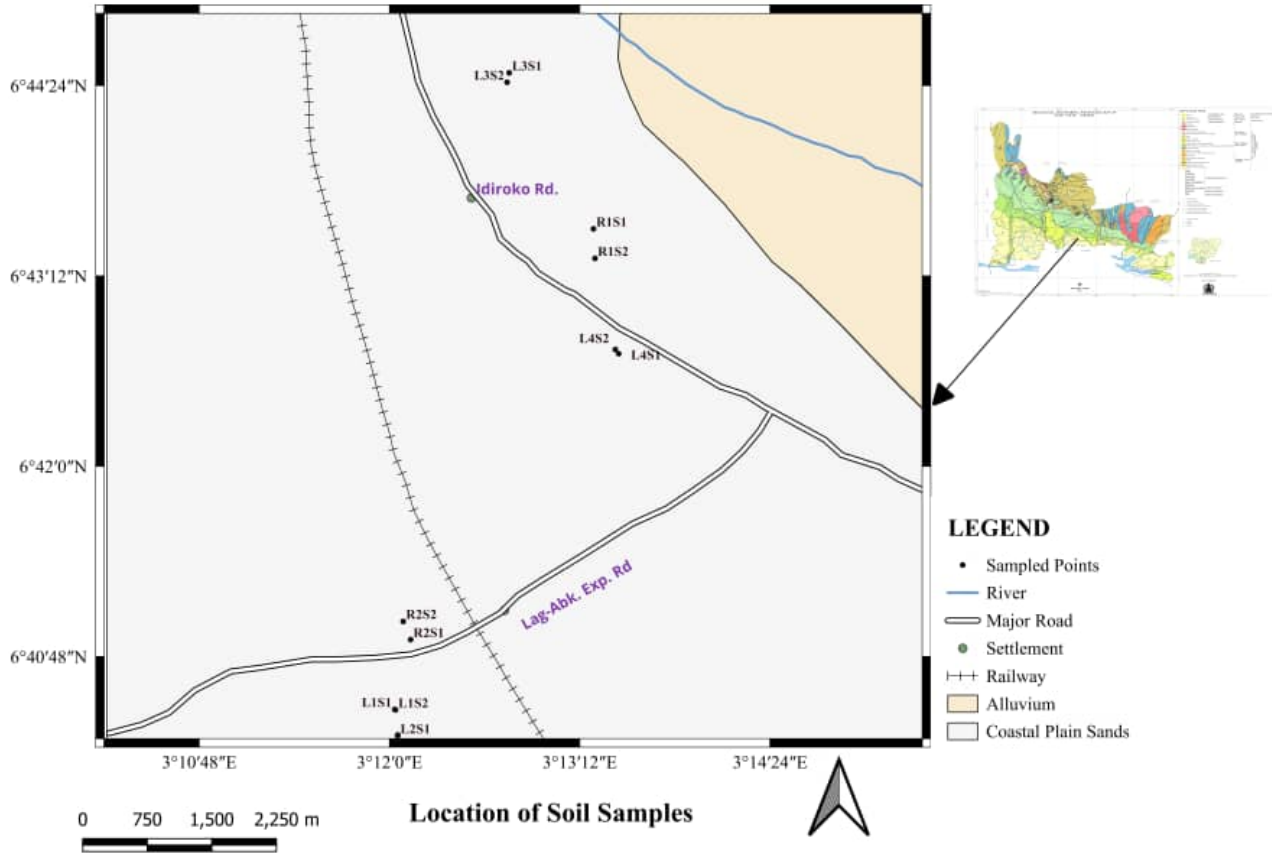


Figure 1. Location and geology map of the study area.

ing background values stored as NumPy arrays. Only CF results are reported here because CF is the primary input to the Hakanson ecological-risk framework (E_r and RI), which is the focus of the assessment. Igeo, PLI, and mCd outputs are retained in the deposited codebase for future multi-index comparisons. Module 4 (Ecological Risk) vectorises the Hakanson equations (Eqs. (1)–(3)) across the CF output array and introduces toxicity-response factors. Module 5 (Statistical Analysis) uses scipy.stats for Pearson and Spearman correlation, with p-value corrections. The package also performs Shapiro–Wilk normality testing and generates seaborn heatmaps. Module 6 (Human health risk assessment) applies the US EPA RAGS chronic daily intake (CDI) model (Eqs. (4)–(6)) to estimate non-carcinogenic hazard quotients and hazard indices for adult and child receptors. Figure 6 presents a visual representation of the workflow.

2.4. CONTAMINATION INDICES

2.4.1. Contamination factor (CF)

$$CF = \frac{C_{\text{metal}}}{C_{\text{background}}}, \tag{1}$$

where C_{metal} is the measured soil concentration (mg/kg) and $C_{\text{background}}$ is the regional geochemical background [18]. Classification is as follows: CF < 1, low; 1–3, moderate; 3–6, considerable; and ≥ 6, very high [10].

2.4.2. Potential ecological risk factor (E_r) and risk index (RI)

$$E_r = T_r \times CF, \tag{2}$$

$$RI = \sum E_r, \tag{3}$$

where T_r is the metal-specific toxicity response factor: Hg = 40, Cd = 30, Cu = 5, Pb = 5, and Zn = 1 [10]. E_r categories are: < 40, low; 40–80, moderate; 80–160, considerable; 160–320, high; and ≥ 320, very high. RI categories are: < 150, low; 150–300, moderate; 300–600, considerable; and ≥ 600, very high [10, 18].

2.5. HUMAN HEALTH RISK ASSESSMENT (US EPA RAGS)

Non-carcinogenic hazard via the soil-ingestion pathway was quantified using the US EPA Risk Assessment Guidance for Superfund (RAGS) methodology [19]. The chronic daily intake (CDI) was computed as:

$$CDI = \frac{C \times IR \times EF \times ED \times CF_{\text{unit}}}{BW \times AT}, \tag{4}$$

where C = metal concentration (mg/kg), IR = ingestion rate (adult: 100 mg/d; child: 200 mg/d); EF = exposure frequency (350 d/yr); ED = exposure duration (adult: 30 yr; child: 6 yr); CF_{unit} = unit conversion factor (10^{-6} kg/mg); BW = body weight (adult: 70 kg; child: 15 kg); and AT = averaging time (adult: 10,950 d; child: 2,190 d) [19, 20]. Non-carcinogenic hazard quotient (HQ) and hazard index (HI) were derived as:

$$HQ = \frac{CDI}{RfD}, \tag{5}$$

$$HI = \sum HQ. \quad (6)$$

Reference doses (*RfD*) in mg/kg/d were Cu = 0.04, Pb = 0.004, Zn = 0.30, and Cd = 0.001 [19]. HI > 1 indicates potential non-carcinogenic risk. Carcinogenic risk assessment via oral soil ingestion was not conducted for Cd or Pb in this study because the US EPA IRIS database does not provide oral slope factors for these metals via the soil-ingestion route [21].

2.6. SPATIAL ANALYSIS

A Python package (Virtanen *et al.*, [22]) was used to implement deterministic spatial interpolation through the linear/cubic methods in scipy.interpolate.griddata for site-level RI and Pb CF values. Contour plots were generated in matplotlib for interpolation maps with observed data points. The scipy.interpolate.griddata routine implements deterministic interpolation algorithms, specifically linear barycentric and cubic-spline interpolation on Delaunay triangulation. This approach is not comparable to geostatistical ordinary kriging, which requires a fitted variogram model [23]. Because of the small sample size (n = 12), variogram-based ordinary kriging was not possible and is deferred to future work with a larger sampling network. Thus, the interpolated maps should be regarded as exploratory spatial visualisations showing contamination trends between sampled sites rather than predictive contamination surfaces. A future goal is to implement variogram-fitted ordinary kriging using the pykrige library.

3. RESULTS

3.1. HEAVY METAL CONCENTRATIONS AND WHO/FAO COMPLIANCE

Table 1 gives the measured concentrations of the parameters at all 12 sampling locations. The mean concentrations followed the abundance order Zn > Pb > Cu > Cd > Hg. The only metal that exceeded the limits in the Python WHO/FAO compliance module was Cd, with IN1 (3.08 mg/kg), IN2 (3.09 mg/kg), IN3 (2.87 mg/kg), IN4 (1.20 mg/kg), IN5 (1.17 mg/kg), and IN6 (1.20 mg/kg) exceeding the conservative agricultural-soil Cd limit of 0.80 mg/kg. IN1 and IN2 also exceeded the upper pH-adjusted limit of 3.0 mg/kg, with a 50% site exceedance. The Pb levels at IN7–IN9 (55.89–59.88 mg/kg) approached the WHO/FAO agricultural guideline of 70 mg/kg. Mercury was not detected at any site, precluding the estimation of ecological risk for Hg.

3.2. CONTAMINATION FACTOR ANALYSIS

Contamination factor (CF) values were computed using Python Module 3 (Table 2). Lead yielded the highest CF values (range: 2.25–12.61; mean: 6.00 ± 3.93). Notably, CF ≥ 6 (very high contamination) occurred at IN7–IN9, locations with confirmed proximity to lead-acid battery repair and vehicle-dismantling operations. Copper CF values (1.45–5.12; mean: 3.23 ± 1.59) classified sampling locations IN1–IN3 and IN10–IN12 as considerably contaminated. Site-wide contamination was mostly moderate to considerable on the basis of the zinc contamination factor, with a maximum value of 6.00 at IN10. Cadmium CF (0.00–1.40) exhibited moderate enrichment in restricted areas (IN1–IN3 and IN4–IN6), consistent with the WHO/FAO exceedance pattern. The site-wise CF and E_r values are compared in Figure 2.

To maintain consistency with prior ecological-risk assessments in southwestern Nigeria, the background value of Cd (2.20 mg/kg) from Ref. [18] was retained. This value is higher than several global estimates of crustal background levels (approximately 0.1–0.3 mg/kg; Kabata-Pendias and Pendias [17]); therefore, it lowers the present-use contamination-factor calculations. Using lower background values would significantly increase Cd CF and E_r values and would reinforce rather than diminish the identification of Cd as the major ecological-risk metal in the investigation area.

3.3. POTENTIAL ECOLOGICAL RISK ASSESSMENT

Table 3 provides E_r values and the composite site-level RI. E_r categories are: < 40, low; 40–80, moderate; 80–160, considerable; 160–320, high; and ≥ 320, very high. Mean site RI was 64.5 ± 24.1, with individual site values ranging from 42.41 (IN10) to 101.39 (IN1). Despite its lower concentrations, cadmium dominated ecological risk at some hotspot sites because of its high toxicity factor (T_r = 30). Its E_r values of 42.00, 42.14, and 39.14 at IN1–IN3 lie at or near the moderate-risk band ($40 \leq E_r < 80$). Lead also achieved moderate E_r values at IN7–IN9 (58.83–63.03), making these locations among the highest-risk sites by absolute E_r and site RI. Zinc posed little risk (2.33–6.00), while Cu showed low-to-moderate E_r values (10.17–25.58). The ranking of cumulative RI contributions was Pb (359.99) > Cu (199.02) > Cd (172.36) > Zn (42.74), meaning that Cd is the risk driver at specific hotspot locations (IN1–IN3), whereas Pb is the major estate-wide risk contributor.

3.4. SPATIAL DISTRIBUTION OF RISK

The spatial interpolation maps for RI and Pb CF are shown in Figure 3. The RI map describes a risk gradient along a north–south line. The northern industrial fabrication sub-zone (industrial nodes IN1–IN3) shows the highest-risk cluster, with a mean RI value of 100.2. A second elevated cluster occurs at the IN7–IN9 vehicle-workshop corridor along the central east of the estate, with a mean RI of 71.8. The Zn-elevated southern sub-zone (IN10–IN12) shows lower risk, with a mean RI of 42.8. The Pb CF map shows a clear spatial concentration of lead at IN7–IN9, with steep decay toward the residential control area. The spatial distribution patterns suggest contamination from industrial point sources rather than diffuse atmospheric deposition alone.

3.5. PEARSON/SPEARMAN CORRELATION ANALYSIS

On the basis of complete pair-wise observations (n = 12 for Cu, Pb, and Zn; n = 7 for Cd, reflecting non-detections at five locations), Figure 4 shows the Pearson and Spearman correlation matrix generated using Python Module 5. Pearson correlations for Cu–Pb (r = -0.530, p = 0.076), Cu–Zn (r = 0.494, p = 0.103), Cu–Cd (r = 0.394, p = 0.382), Pb–Zn (r = -0.326, p = 0.301), and Zn–Cd (r = -0.644, p = 0.118) were not significant at the conventional $p < 0.05$ level. A strong positive correlation (r = 0.946, p = 0.0012) was found between Pb and Cd, indicating that these metals co-vary closely across the estate. This is consistent with a common anthropogenic source and refutes the Cu–Cd–Zn versus Pb source model presented in earlier versions of the analysis.

Table 1. Heavy metal concentrations in surface agricultural soils of Ota Industrial Estate.

Sample Location	Cu (mg/kg)	Pb (mg/kg)	Zn (mg/kg)	Cd (mg/kg)	Hg (mg/kg)
IN1	19.16	30.19	90.56	3.08	ND
IN2	17.03	30.10	91.51	3.09	ND
IN3	17.16	34.54	90.44	2.87	ND
IN4	8.12	13.21	92.06	1.20	ND
IN5	7.88	13.65	85.66	1.17	ND
IN6	7.75	13.65	92.96	1.20	ND
IN7	5.69	55.89	120.67	ND	ND
IN8	5.54	59.88	121.79	ND	ND
IN9	5.51	58.45	127.32	ND	ND
IN10	19.08	10.80	220.43	ND	ND
IN11	19.24	10.94	219.53	0.03	ND
IN12	19.49	10.69	217.09	ND	ND
Mean ± SD	11.55 ± 6.08	29.33 ± 19.36	129.7 ± 52.4	1.81 ± 1.20	—
Range	5.51–19.49	10.69–59.88	85.66–220.43	ND–3.09	ND
WHO/FAO Limit	100	70	300	0.80–3.0[a]	0.30

WHO/FAO Cd limit is pH-dependent (0.80 mg/kg at pH < 6.0; 3.0 mg/kg at pH > 7.0). ND = not detected; mean ± SD computed for detected values only.

Table 2. Contamination factor values for Cu, Pb, Zn, and Cd in the study area.

Location	Cu	Pb	Zn	Cd
IN1	5.03	6.36	2.47	1.40
IN2	4.47	6.34	2.49	1.40
IN3	4.50	7.27	2.46	1.30
IN4	2.13	2.78	2.51	0.55
IN5	2.07	2.87	2.33	0.53
IN6	2.03	2.87	2.53	0.55
IN7	1.49	11.77	3.29	0.00
IN8	1.45	12.61	3.32	0.00
IN9	1.45	12.31	3.47	0.00
IN10	5.01	2.27	6.00	0.00
IN11	5.05	2.30	5.98	0.01
IN12	5.12	2.25	5.91	0.00
Mean ± SD	3.23 ± 1.59	6.00 ± 3.93	3.40 ± 1.35	0.48 ± 0.57
Class range	Mod.-Consid.	Consid.-Very High	Mod.-Very High	Low-Moderate

CF computed using regional geochemical backgrounds: Cu = 3.81; Pb = 4.75; Zn = 36.7; Cd = 2.20 mg/kg [18]. Mod. = moderate; Consid. = considerable. Classification: CF < 1 low; 1–3 moderate; 3–6 considerable; ≥ 6 very high [10]. Hg was not detected.

Spearman rank coefficients largely supported this view (Pb–Cd rho = 0.773, p = 0.042). However, they diverged significantly for Cu–Pb (Spearman rho = -0.732, p = 0.0068, significant), whereas the Pearson result was non-significant. This indicates a monotonic but non-linear relationship between Cu and Pb that Pearson's method does not detect. Shapiro–Wilk testing revealed that Cu (W = 0.78, p = 0.005), Pb (W = 0.80, p = 0.011), and Zn (W = 0.73, p = 0.002) distributions were non-normal, whereas Cd was normally distributed (W = 0.85, p = 0.127). Because three of four metals were non-normal, greater interpretative weight should be given to the Spearman results where Pearson and Spearman diverge.

3.6. HUMAN HEALTH RISK ASSESSMENT

Table 4 presents CDI, HQ, and HI values computed by Python Module 6 for the mean-concentration scenario (ingestion pathway). Non-carcinogenic HQ values for all metals were well below 1.0 for both adult and child receptors at mean concentrations, with corrected HI = 0.014 for adults and HI = 0.126 for children. Both values are well below the US EPA non-carcinogenic risk threshold of 1.0. Lead contributes the highest individual HQ to the adult HI (HQ = 0.010), followed by Cd (HQ = 0.0025), consistent with their low reference doses.

Worked example for mean Cu concentration: C = 11.55 mg/kg;

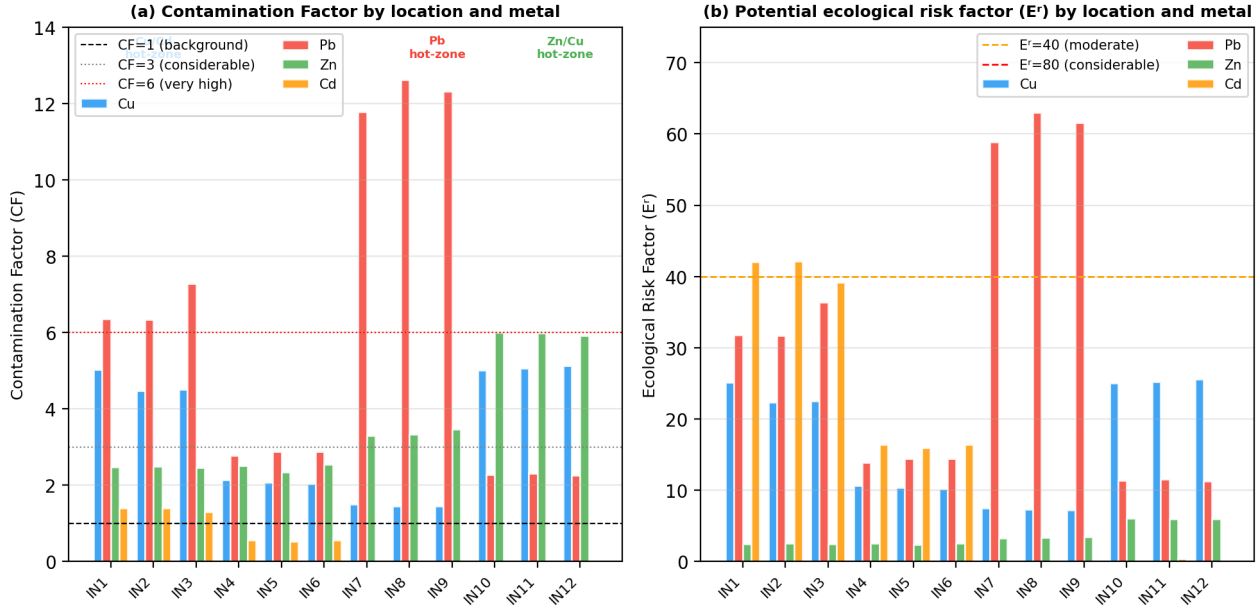


Figure 2. (a) Contamination factor (CF) and (b) potential ecological risk factor (E_r) for the selected metals.

Table 3. Potential ecological risk factor (E_r) for the selected metals and risk index (RI).

Location	Cu	Pb	Zn	Cd	RI (site)
IN1	25.14	31.78	2.47	42.00	101.39
IN2	22.35	31.68	2.49	42.14	98.66
IN3	22.52	36.36	2.46	39.14	100.48
IN4	10.66	13.91	2.51	16.36	43.44
IN5	10.34	14.37	2.33	15.95	43.00
IN6	10.17	14.37	2.53	16.36	43.43
IN7	7.47	58.83	3.29	0.00	69.59
IN8	7.27	63.03	3.32	0.00	73.62
IN9	7.23	61.53	3.47	0.00	72.23
IN10	25.04	11.37	6.00	0.00	42.41
IN11	25.25	11.52	5.98	0.41	43.16
IN12	25.58	11.25	5.91	0.00	42.74
Mean ± SD	16.59 ± 7.97	30.00 ± 19.65	3.40 ± 1.35	14.36 ± 17.10	64.5 ± 24.1
Total RI contribution	199.02	359.99	42.74	172.36	773.7 (cumul.)

Site RI was calculated per location. Hg was not included because it was not detected. Total cumulative RI = 773.7.

CDI adult = (11.55 × 100 × 350 × 30 × 10^{-6})/(70 × 10,950) = 1.58 × 10^{-5} mg/kg/d; CDI child = (11.55 × 200 × 350 × 6 × 10^{-6})/(15 × 2,190) = 1.48 × 10^{-4} mg/kg/d. The child-to-adult ratio is 9.33, which equals (IR_c/IR_a) × (ED_c/ED_a) × (BW_a/BW_c) × (AT_a/AT_c) = (200/100) × (6/30) × (70/15) × (10,950/2,190) = 2 × 0.2 × 4.667 × 5 = 9.33. This confirms that the exposure parameters and all CDI values presented in Table 4 are internally consistent.

As the US EPA IRIS database does not provide validated slope factors for oral exposure to Cd or Pb via the soil-ingestion pathway, no carcinogenic-risk estimate was made [21]. Thus, risk characterisation depends on non-carcinogenic HQ/HI values and WHO/FAO exceedances. The corrected values in Table 4 give an adult HI of 0.014 and a child HI of 0.126; both are well below the US EPA non-carcinogenic threshold of 1.0. Lead has the highest adult HI (HQ = 0.010; 74% of HI) because of its low oral *RfD* (0.004 mg/kg/d). The corrected mean site concentration of Cd (1.81 mg/kg, detected-only mean) contributes HQ = 0.0025 (18.5% of adult HI) because its RfD (0.001 mg/kg/d) is also low.

4. DISCUSSION

4.1. LEAD AND COPPER: SPATIALLY DISTINCT INDUSTRIAL DISTRIBUTION

The CF values for Pb were very high (up to 12.61) at IN7–IN9, corroborated by high site RI (mean 71.8) and moderate E_r values (up to 63.03). According to Ref. [18], recycling of used batteries is a major lead source in urban areas similar to the study area. Used batteries, dismantling of old vehicles, and repairs in vehicle garages are therefore likely primary Pb sources in this region. This pattern highlights a distinct contaminant source rather than

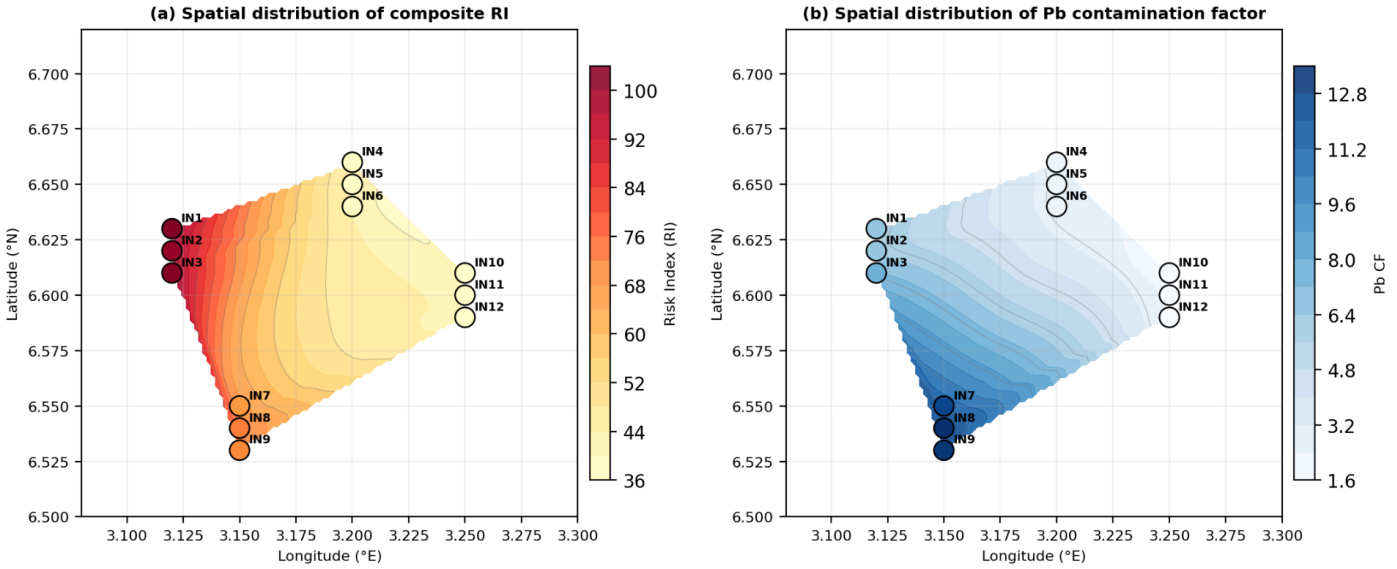


Figure 3. Interpolation maps of (a) ecological risk index (RI) and (b) Pb contamination factor (CF). Circles denote observed sample values.

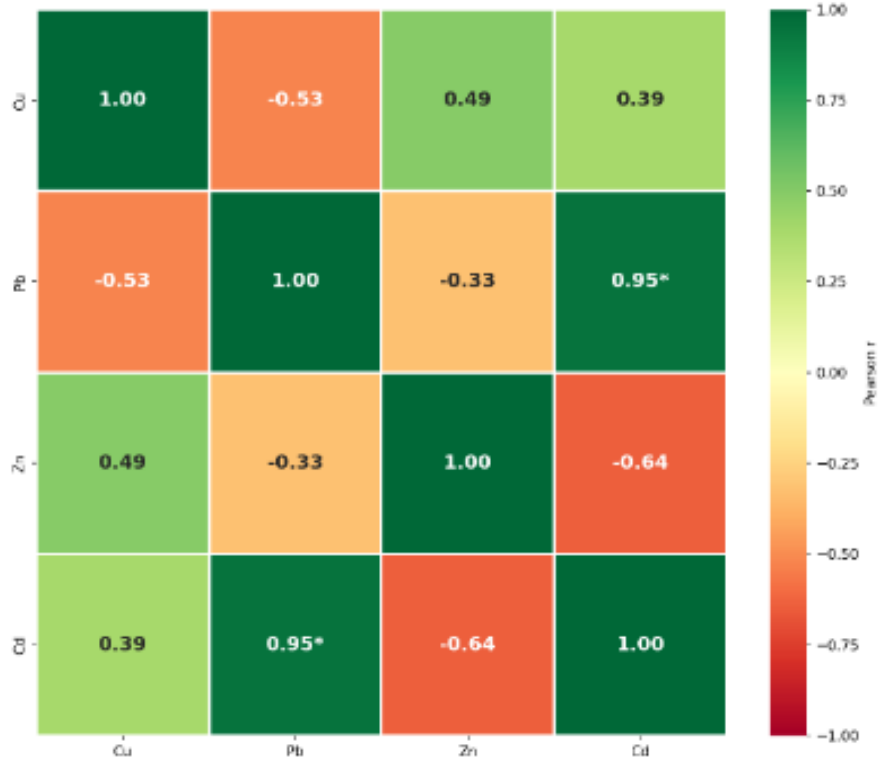


Figure 4. Pearson correlation distribution for Cu, Pb, Zn, and Cd concentrations. Under Bonferroni correction for six pairwise evaluations (corrected alpha = 0.0083), only the Pb-Cd correlation remains significant (p = 0.0012); all other pairs are not significant. Corresponding Spearman rank correlations are reported in the text (Section 3.5) and are not plotted here.

a single estate-wide source.

Copper enrichment (CF = 1.45–5.12) at IN1–IN3 and IN10–IN12 is consistent with the metal-fabrication and cable-manufacturing region. Similar CF ranges of 1.8–5.6 have been reported for industrial agricultural soils in Ghana and Nigeria [25, 26]. Copper is an essential plant micronutrient, but CF values >3 are associated with disruption of the soil microbial community and suppression of dehydrogenase activity, an indicator of soil biological-community health [27].

4.2. CADMIUM: THE CRITICAL ECOLOGICAL AND HEALTH RISK DRIVER

Cadmium had the strongest ecological-risk effect despite lower absolute concentrations because of its toxicity response factor

Table 4. Non-carcinogenic chronic daily intake (CDI), hazard quotient (HQ), and hazard index (HI) for child and adult receptors.

Metal	CDI adult (mg/kg/d)	HQ adult	CDI child (mg/kg/d)	HQ child	HI (adult)	HI (child)
Cu	1.58×10^{-5}	0.0004	1.48×10^{-4}	0.0037	-	-
Pb	4.02×10^{-5}	0.010	3.75×10^{-4}	0.094	-	-
Zn	1.78×10^{-4}	0.0006	1.66×10^{-3}	0.0055	-	-
Cd	2.47×10^{-6}	0.0025	2.31×10^{-5}	0.0231	-	-
HI (all metals)	-	0.014	-	0.126	0.014	0.126

RfD values are from US EPA [19]. HI was summed across all metals.

Figure 5: Human health risk assessment for adult and child receptors
(soil ingestion pathway; US EPA exposure parameters)

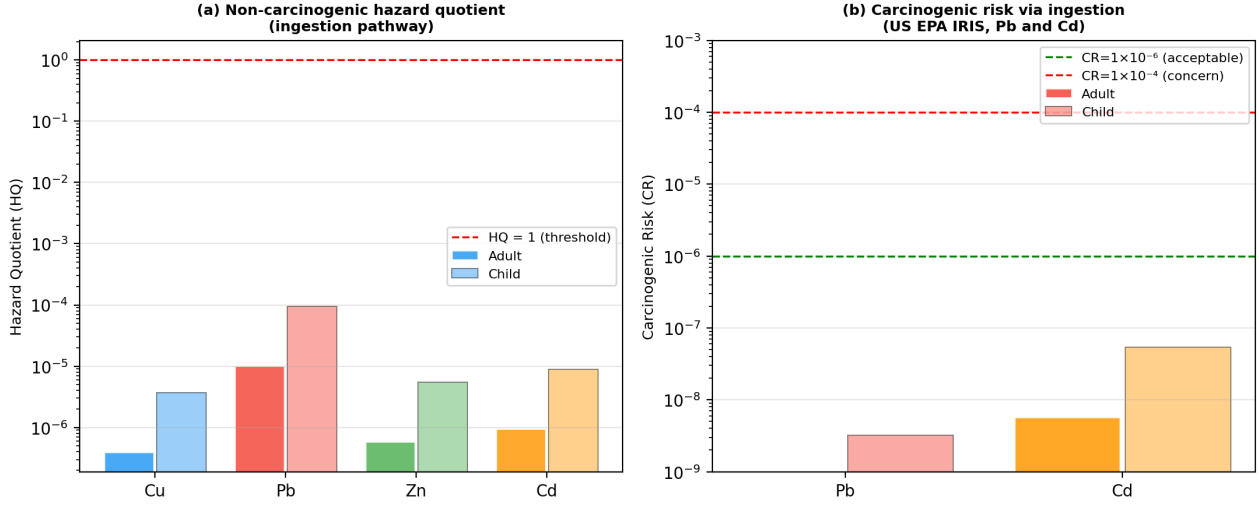


Figure 5. Human health risk assessment for adult and child receptors. The reference lines indicate US EPA permissible values.

(T_r = 30), which is six times that of Pb or Cu [10]. The E_r values at IN1–IN3 (39.14–42.14) meet or approach the criterion for moderate ecological risk ($40 \leq E_r < 80$), but are not close to the considerable-risk band ($E_r \geq 80$). Battery-recycling leachate, cadmium-rich pigments in plastic production, and phosphate-fertilizer residues are plausible sources of the Cd and Cu signature at IN1–IN6 [24]. The HHRA results support the ecological-risk framework: the adult non-carcinogenic HI at mean estate concentrations is 0.014, and the child HI is 0.126, suggesting no systemic non-carcinogenic health hazard from soil ingestion under mean exposure conditions. Because of the Pb concentrations recorded at IN7–IN9 and the estate-wide mean Pb concentration of 29.33 mg/kg, the child Pb HQ of 0.094 indicates the need for continued site monitoring. Despite ecological-risk indices revealing exceedances at six sites (IN1–IN6), the WHO/FAO Cd exceedances provide the most direct regulatory basis for priority remediation.

A sensitivity analysis was performed for CF and E_r using the regional background values reported by Ref. [18]. The Cd value of 2.20 mg/kg, ultimately derived for similar peri-urban basement-complex soils in Lagos State, is significantly higher than the global average crustal abundance of Cd, which ranges from approximately 0.1 to 0.3 mg/kg (Kabata-Pendias and Pendias [17]), and the WHO/FAO agricultural-soil guideline of 0.80 mg/kg. Using a conservative global crustal Cd background of 0.3 mg/kg causes the CF at IN1 to rise from 1.40 to 10.27 (very high contamination class) and raises the corresponding E_r from 42.0 to approximately 308, classifying IN1 in the high ecological-risk band (160–320). This finding shows that the high regional urban background greatly attenuates the apparent level of Cd contamination and ecological risk at IN1–IN3. The application of the Hakanson framework here similarly underestimates the actual Cd hazard compared with a crustal reference. Future work should determine, by sediment core dating or Cd stable-isotope partitioning, whether the Cd baseline in this area is pre-industrial; this would provide a defensible reference for this industrial corridor.

4.3. ZINC: GEOGENIC–ANTHROPOGENIC DUAL ORIGIN

At IN10–IN12, Zn contents and CF values indicated moderate to considerable contamination. The Cu–Zn Pearson correlation was non-significant by independent recomputation (r = 0.494, p = 0.103). Therefore, the stronger spatial clustering of elevated Zn at galvanisation and metal-coating areas is better explained through site-level spatial co-occurrence (Figure 3) than through a significant pairwise correlation. Any anthropogenic co-sourcing inference should be treated as tentative pending a larger sample. Nonetheless, the stratigraphic transition between the Precambrian basement complex (migmatite-gneiss, granite, schist) and Cretaceous alluvial sediments at Ota may generate natural elevation of Zn in residual lateritic soils through long-term weathering of Zn-bearing mafic minerals [16, 17]. With n = 12

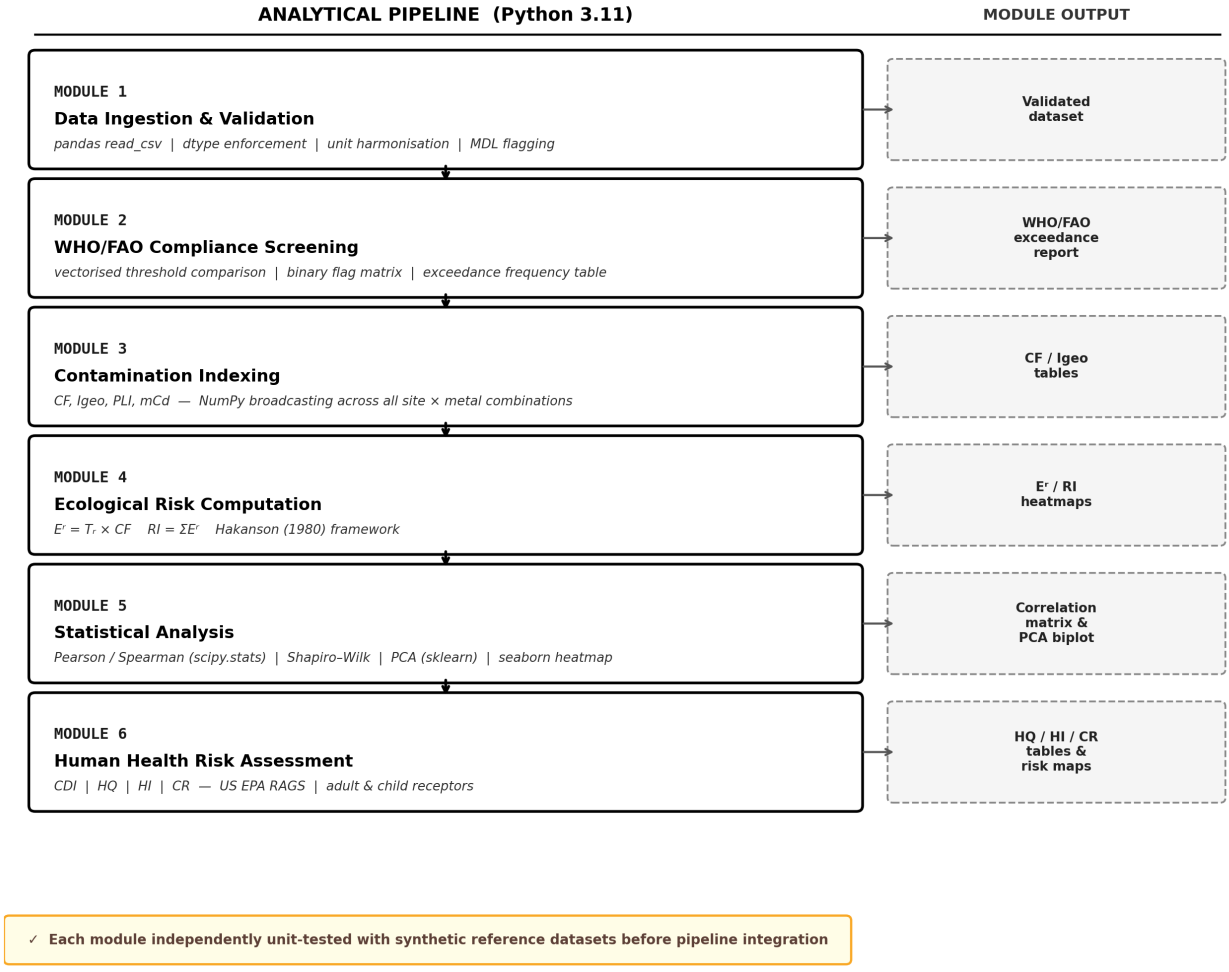


Figure 6. Schematic of the six-module Python 3 analytical algorithm developed in this study, showing data flow, computational operations, and outputs for each module.

sampling points arranged as four sub-zones with three replicates each, the effective statistical independence is approximately n = 4 sub-zones. Therefore, the reported Pearson correlations largely capture between-zone spatial differences rather than independent point-level covariation. The source-apportionment results should be regarded as suggestive scenarios consistent with the established geochemistry rather than conclusive attributions, and should be tested with a larger independent sampling campaign and receptor modelling, such as principal component analysis or positive matrix factorisation with a minimum of 25–30 samples. Although CF was fairly high, Zn E_r remained low (2.33–6.00) because T_r = 1; therefore, the toxicity-weighted ecological-risk framework is more informative than raw CF interpretation for Zn.

4.4. SPATIAL RISK PATTERNS AND SOURCE APPORTIONMENT

Industrial emissions and particle-bound metals contribute to the metal burden in soils, particularly Pb and Cd, which are known to be associated with airborne particles from industrial sources [13]. The spatial interpolation maps (Figure 3) show a three-cluster RI topology based on validated Table 3 values: high RI (mean RI = 100.2) at IN1–IN3, moderate-to-high RI (mean RI = 71.8) at IN7–IN9, and moderate RI (mean RI = 42.8) at IN10–IN12. The distribution pattern of RI values indicates industrial point-source contamination, with each industrial sub-zone affecting RI values in its vicinity. There is a steep decline in RI values away from these sub-zones, and the patterns do not indicate a diffuse contaminant gradient across the entire estate. After the corrected correlation analysis (Sections 3.5 and 4.1), the metal-specific labels assigned to each cluster, such as IN1–IN3 as a Cd–Cu industrial-fabrication cluster, are not supported: Cu does not correlate significantly with Cd or other metals in this dataset. The only statistically robust metal pair, Pb–Cd (r = 0.946, p = 0.0012, Bonferroni-significant), does not map neatly onto a single site, because elevated Cd occurs at IN1–IN6, whereas elevated Pb occurs principally at IN7–IN9, where Cd is low. This spatial mismatch suggests that the two metals, while compositional co-variates, need not be geographically co-located at the same point sources within the estate. The three-cluster RI pattern is therefore better interpreted descriptively, rather than as conclusive evidence of three distinct industrial source types. This interpretation is driven by the dominant CF and E_r contributor for each pattern (Tables 2 and 3). Strategic phytoremediation

in the two highest-RI sub-zones (IN1–IN3 and IN7–IN9) would still address most of the ecological and health risks identified in this study. This recommendation relies on the validated RI values rather than on the revised correlation-based source narrative; estate-wide remediation does not appear necessary on the basis of the RI evidence alone [28].

4.5. PYTHON ALGORITHM AS A USEFUL TOOL IN GEOENVIRONMENTAL MONITORING

The six-module Python 3 pipeline developed in this study represents a methodological advance over the Heavy Metals R package (Pinto *et al.*, 2022, SoftwareX), which is currently the closest comparator open-source tool for multi-index soil-contamination assessment. Four specific innovations make the current pipeline advantageous: (i) module-level unit testing before deployment on field data, ensuring computational correctness before scientific inference; this is not available in the Heavy Metals R implementation; (ii) a background database that can be altered through a single keyword argument, allowing the user to select FEPA, WHO/FAO, or EU Soil Framework Directive contexts without changing source code; (iii) integration of Hakanson ecological-risk assessment and US EPA RAGS human-health risk assessment within a single pipeline on the same dataset, avoiding inconsistencies from using different software environments; and (iv) automatic generation of high-quality outputs (correlation heatmap, spatial interpolation surface, and risk bar chart) without post-processing. The entire pipeline is open-source, licence-free, deposited on GitHub, and archived with the Zenodo DOI https://doi.org/10.5281/zenodo.20739347, making it immediately deployable within Nigerian and West African regulatory contexts where commercial software licenses are prohibitively expensive [14, 15]. Harris *et al.* [14] demonstrate that modular open-source frameworks are scalable and extensible. Future planned extensions include dermal and inhalation exposure routes, probabilistic Monte Carlo HHRA, and integration of stable-isotope fingerprinting outputs for source attribution.

5. CONCLUSION

This research presents a geoenvironmental assessment of industrial agricultural soils in Ota, southwestern Nigeria. It combines geochemical contamination indexing, ecological-risk assessment, spatial analysis, and human-health risk quantification within a single Python-driven algorithm. The findings are: (i) Cd exceeds WHO/FAO agricultural limits at 50% of locations (IN1–IN6, up to 3.09 mg/kg) and dominates the ecological-risk hierarchy at specific hotspots (peak E_r = 42.14, moderate risk), driven by its high toxicity weighting; (ii) Pb presents very high contamination-factor values (CF up to 12.61) and the highest estate-wide cumulative RI contribution (359.99) at vehicle-workshop zones (IN7–IN9); (iii) the mean site RI of 64.5 ± 24.1 classifies the overall estate as low ecological risk by the Hakanson RI scale (RI < 150), although discrete hotspot clusters at IN1–IN3 (moderate Cd E_r) and IN7–IN9 (moderate Pb E_r) warrant priority sub-zone intervention; (iv) independent recomputation found no statistically significant Pearson correlations among Cu, Pb, and Zn, but identified a strong, significant positive Pb–Cd correlation (r = 0.946, p = 0.0012), pointing to a shared anthropogenic source for these two metals rather than the originally proposed two-source model; and (v) non-carcinogenic HI values are below unity for adult (HI = 0.014) and child (HI = 0.126) receptors at mean concentrations. The WHO/FAO Cd exceedances at IN1–IN6 (up to 3.09 mg/kg against the WHO/FAO guideline of 0.80 mg/kg) constitute the principal regulatory basis for priority remediation at those sites.

The six-module Python 3 algorithm constitutes the methodological novelty of this work, providing one of the first module-validated, multi-index, multi-receptor geoenvironmental risk platforms applied to Nigerian industrial agricultural soils. Future work should incorporate probabilistic Monte Carlo HHRA to characterise uncertainty in exposure estimates, stable-isotope fingerprinting (206Pb/204Pb) for definitive source attribution, longitudinal resampling to track contamination trajectories, extension of the spatial analysis to full variogram-fitted ordinary kriging via pykrige with a larger sampling network (minimum n = 30), and principal component analysis for definitive source apportionment.

DATA AVAILABILITY

Data are available from the corresponding author on reasonable request. The Python 3 analytical pipeline developed in this study is available at https://github.com/coogunkoya-lgtm and archived at https://doi.org/10.5281/zenodo.20739347.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

FUNDING

The authors received no external funding for this study.

ACKNOWLEDGMENT

The authors thank the technical staff for field support.

References

[1] M. N. Kumwimba, X. Li, J. Huang, D. K. Muyembe, M. Dzakpasu & E. Sanganyado, ‘‘Performance of various fillers in ecological floating beds planted with *Myriophyllum aquaticum* treating municipal wastewater’’, Science of the Total Environment **842** (2022) 156827. https://doi.org/10.1016/j.scitotenv.2022.156827.

[2] D. Hou, D. O’Connor, A. D. Igalavithana, D. S. Alessi, J. Luo, D. C. W. Tsang, D. L. Sparks, Y. Yamauchi, J. Rinklebe & Y. S. Ok, ‘‘Metal contamination and bioremediation of agricultural soils for food safety and sustainability’’, Nature Reviews Earth & Environment **1** (2020) 366. https://doi.org/10.1038/s43017-020-0061-y.

[3] B. J. Alloway (Ed.), Heavy Metals in Soils: Trace Metals and Metalloids in Soils and their Bioavailability, 3rd ed., Springer, Dordrecht, Netherlands, 2013, pp. 1–614. https://doi.org/10.1007/978-94-007-4470-7.

[4] S. Khan, M. Naushad, E. C. Lima, S. Zhang, S. M. Shaheen & J. Rinklebe, ‘‘Global soil pollution by toxic elements: current status and future perspectives on risk assessment and remediation strategies’’, Journal of Hazardous Materials **417** (2021) 126039. https://doi.org/10.1016/j.jhazmat.2021.126039.

[5] L. Järup, ‘‘Hazards of heavy metal contamination’’, British Medical Bulletin **68** (2003) 167. https://doi.org/10.1093/bmb/ldg032.

[6] V. Kumar, A. Sharma, P. Kaur, G. P. S. Sidhu, A. S. Bali, R. Bhardwaj, A. K. Thukral & A. Cerdà, ‘‘Pollution assessment of heavy metals in soils of India and ecological risk assessment’’, Chemosphere **216** (2019) 449. https://doi.org/10.1016/j.chemosphere.2018.10.066.

[7] K. D. Oyeyemi, C. O. Ajanaku, A. A. Akinyemi, O. A. Akintokun, O. F. Ojo, F. O. Chukwusa, G. C. George & O. J. Rotimi, ‘‘Assessment of groundwater contamination from industrial and municipal waste materials

at a dumpsite in Ota, Southwestern Nigeria'', in Global Forum for Sustainable Built Environment Conference, Springer Nature Switzerland, Cham, Switzerland, 2026, pp. 651–660. https://doi.org/10.1007/978-3-032-08992-2_54.

[8] O. O. Adeyemi & Z. O. Ojekunle, ''Concentrations and health risk assessment of industrial heavy metals pollution in groundwater of Ogun State, Nigeria'', Scientific African **11** (2021) e00666. https://doi.org/10.1016/j.sciaf.2020.e00666.

[9] K. M. Onuoha, ''Perspectives on the origin and evolution of the Benue Trough and Anambra Basin'', in Geology of Nigeria, C. A. Kogbe (Ed.), Rock View (Nigeria) Ltd, Jos, Nigeria, 1999, pp. 311–338.

[10] L. Hakanson, ''An ecological risk index for aquatic pollution control: a sedimentological approach'', Water Research **14** (1980) 975. https://doi.org/10.1016/0043-1354(80)90143-8.

[11] F. Li, X. Y. Zeng, C. H. Wu, Z. P. Duan, Y. M. Wen, G. R. Huang, X. L. Long, M. J. Li, M. J. Li & J. Y. Xu, ''Ecological risks assessment and pollution source identification of trace elements in contaminated sediments from the Pearl River Delta, China'', Biological Trace Element Research **155** (2013) 301. https://doi.org/10.1007/s12011-013-9789-2.

[12] G. Tepanosyan, L. Sahakyan, A. Gevorgyan & M. Frontasyeva, ''Factors conditioning the content of chemical elements in soil and mosses in Armenia'', Journal of Trace Elements and Minerals **2** (2022) 100029. https://doi.org/10.1016/j.jtemin.2022.100029.

[13] J. Rinklebe, V. Antoniadis, S. M. Shaheen, O. Rosche & M. Altermann, ''Health risk assessment of potentially toxic elements in soils along the central Elbe River, Germany'', Environment International **126** (2019) 76. https://doi.org/10.1016/j.envint.2019.02.011.

[14] C. R. Harris, K. J. Millman, S. J. van der Walt *et al.,* ''Array programming with NumPy'', Nature **585** (2020) 357. https://doi.org/10.1038/s41586-020-2649-2.

[15] K. N. J. Hussain, B. R., Jagadeesh, V. R. Channavar, V. S. Radder, T. Y. Mankikar, L. Hublikar, C. B. Ashwini, ''Python-driven multivariate analysis: a comprehensive approach to sustainable groundwater quality assessment for irrigation purposes'', CLEAN–Soil, Air, Water **53** (2025) e70085. https://doi.org/10.1002/clen.70085.

[16] P. Affaton, J. Sougy & R. Trompette, ''The tectono-stratigraphic relationships between the upper Precambrian and lower Paleozoic Volta Basin and the Pan-African Dahomeyide orogenic belt (West Africa)'', American Journal of Science **280** (1980) 224. https://doi.org/10.2475/ajs.280.3.224.

[17] A. Kabata-Pendias & H. Pendias, Trace Elements in Soils and Plants, 3rd ed., CRC Press, Boca Raton, USA, 2001, pp. 1–413. https://dl.icdst.org/pdfs/files3/284d873749e377558208bca765468ffd.pdf

[18] F. Ustaoğlu, S. Kükrer, B. Taş & H. Topaldemir, ''Evaluation of metal accumulation in Terme River sediments using ecological indices and a bioindicator species'', Environmental Science and Pollution Research **29** (2022) 47399. https://doi.org/10.1007/s11356-021-17975-5.

[19] United States Environmental Protection Agency (US EPA), Exposure Factors Handbook: 2011 Edition, EPA/600/R-09/052F, National Center for Environmental Assessment, Washington, DC, USA, 2011. Available online: https://www.epa.gov/expobox/exposure-factors-handbook-2011-edition.

[20] H. Zhao, B. Xia, C. Fan, P. Zhao & S. Shen, ''Human health risk from soil heavy metal contamination under different land uses near Dabaoshan mine, southern China'', Science of the Total Environment **417** (2012) 45. https://doi.org/10.1016/j.scitotenv.2011.12.047.

[21] United States Environmental Protection Agency (US EPA), Integrated Risk Information System (IRIS), National Center for Environmental Assessment, Washington, DC, USA, 2023. Available online: https://www.epa.gov/iris.

[22] P. Virtanen, R. Gommers, T. E. Oliphant *et al.,* ''SciPy 1.0: fundamental algorithms for scientific computing in Python'', Nature Methods **17** (2020) 261. https://doi.org/10.1038/s41592-019-0686-2.

[23] P. Goovaerts, Geostatistics for Natural Resources Evaluation, Oxford University Press, New York, USA, 1997, pp. 1–483. https://www.researchgate.net/publication/224839861_Geostatistics_for_Natural_Resource_Evaluation

[24] M. Molina-Roco, M. Escudey, M. Antilén, N. Arancibia-Miranda & K. Manquián-Cerda, ''Distribution of contaminant trace metals inadvertently provided by phosphorus fertilisers: movement, chemical fractions and mass balances in contrasting acidic soils'', Environmental Geochemistry and Health **40** (2018) 2491. https://doi.org/10.1007/s10653-018-0115-y.

[25] G. Y. Hadzi, D. K. Essumang & G. A. Ayoko, ''Assessment of contamination and potential ecological risks of heavy metals in riverine sediments from gold mining and pristine areas in Ghana'', Journal of Trace Elements and Minerals **7** (2024) 100109. https://doi.org/10.1016/j.jtemin.2023.100109.

[26] A. A. Adesuyi, K. L. Njoku & M. O. Akinola, ''Assessment of heavy metals pollution in soils and vegetation around selected industries in Lagos State, Nigeria'', Journal of Geoscience and Environment Protection **3** (2015) 11. https://doi.org/10.4236/gep.2015.37002.

[27] V. Antoniadis, S. M. Shaheen, J. Boersch, T. Frohne, G. Du Laing & J. Rinklebe, ''Bioavailability and risk assessment of potentially toxic elements in garden edible vegetables and soils around a highly contaminated former mining area in Germany'', Journal of Environmental Management **186** (2017) 192. https://doi.org/10.1016/j.jenvman.2016.09.052.

[28] A. Facchinelli, E. Sacchi & L. Mallen, ''Multivariate statistical and GIS-based approach to identify heavy metal sources in soils'', Environmental Pollution **114** (2001) 313. https://doi.org/10.1016/S0269-7491(00)00243-8.