International Journal of Sustainable Energy and Environmental Research
2026 Vol. 15, No. 2 pp. 18-34
ISSN(e): 2306-6253
ISSN(p): 2312-5764
DOI: 10.18488/13.v15i2.5178

Validation of handheld X-ray fluorescence against wavelength-dispersive X-ray fluorescence for heavy metal quantification in contaminated soils

Mubin Hossain Omio[1+]
Abdus Sattar Mollah[2]

[1,2]*Department of Nuclear Science and Engineering, Military Institute of Science and Technology, Mirpur Cantonment, Dhaka, Bangladesh.*
[1]*Email: omiomho1@gmail.com*
[2]*Email: mosattar54@gmail.com*

(+ Corresponding author)

Article History
Received: 6 May 2026
Revised: 15 June 2026
Accepted: 3 July 2026
Published: 14 September 2026

Keywords
Bland-Altman analysis
Handheld XRF
Heavy metals
Lead
Linear regression
Soil contamination
WDXRF.

ABSTRACT

Handheld X-ray fluorescence (hXRF) technology has revolutionized environmental site assessment by providing rapid, in-situ elemental analysis. Accuracy of these devices in complex, highly contaminated soil matrices remains questionable and validation against laboratory reference methods remains limited. This study evaluates the performance of an Olympus Vanta C Series handheld XRF operated in Alloy Plus mode against laboratory-based wavelength dispersive X-ray fluorescence (WDXRF) using soil samples from a ballistic contamination gradient at the Mirpur Cantonment firing range in Dhaka, Bangladesh. Nine soil samples were collected across a contamination gradient (three sites, three replicates per site). From the study, strong correlations for lead ($r=0.972$; $R2 = 0.945$) and iron ($r = 0.800$) were found with significant systematic bias. In this regard, lead concentrations were overestimated by hXRF about 275.7% with a mean bias = +12.11%, likely due to limited sample size, matrix effects and spectral interferences. Site-specific linear regression correction models were developed and validated using Leave-One-Out Cross-Validation (LOOCV) to mitigate these biases. Empirical correction factors resulted in an RMSE for lead that was reduced from 16.82% to 2.72%, representing an improvement in accuracy of 83.8%. This study concludes that empirical calibration is necessary for quantitative risk assessment, even if uncorrected hXRF data is appropriate for delineating contamination boundaries and relative screening. A tiered decision framework, as proposed, should guide environmental practitioners on the integration of hXRF screening with confirmatory laboratory analysis.

Contribution/Originality: This study contributes the first Bland-Altman-based validation of handheld XRF against WDXRF across an extreme firing-range contamination gradient, quantifying concentration-dependent bias, developing empirical site-specific correction factors that reduced lead measurement error by 83.8%, and proposing a tiered decision framework for field-based contamination screening.

1. INTRODUCTION

Soil contamination due to heavy metals provides a persistent environmental and public health challenge worldwide (Kabata-Pendias, 2011). Toxic heavy metals like lead (Pb), copper (Cu), zinc (Zn) and others accumulate in the environment primarily through anthropogenic activities like industrial waste, mining, agriculture and legacy contamination (Wuana & Okieimen, 2011). Conventional analytical methods like wavelength-dispersive X-ray fluorescence (WDXRF) and ICP-based techniques are highly accurate for quantification, but the requirement of extensive sample preparation and analysis time is a drawback (Borůvka, Vacek, & Jehlička, 2005).

Previously, environmental assessment or surveying required costly laboratory transport of samples. But reducing the need for laboratory testing, handheld X-ray fluorescence (XRF) has revolutionized the in-situ environmental assessment by rapid, non-destructive element analysis directly in the field (Kalnicky & Singhvi, 2001). These instruments complete the assessment within seconds to minutes, utilizing energy dispersive detection systems for analyzing multiple elements (Carr, Zhang, Moles, & Harder, 2008). The fundamental principle for both handheld XRF and WDXRF follows Moseley's law, where characteristic X-ray energy relates to atomic number.

$$E = 13.6eV \cdot (Z - \sigma)^2 \cdot \left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right) \quad (1)$$

Where E is the characteristic X-ray energy, Z is the atomic number, σ is the screening constant, and n_1 and n_2 represent the principal quantum numbers (Beckhoff, Kanngießer, Langhoff, Wedell, & Wolff, 2006; Jenkins, 1999). Despite widespread use, doubt remains regarding the accuracy of complex soil matrices. Wavelength dispersive systems show greater resolution than energy dispersive systems, potentially causing spectral interferences and matrix effects (Potts, Webb, & Watson, 1985; Shackley, 2011). Soil moisture, surface roughness, particle size and heterogeneity are the causes of systematic errors (Rouillon & Taylor, 2016; Tighe, Lockwood, Wilson, & Lisle, 2004).

Handheld XRF has been widely used for rapid soil contamination screening. Conventional validation against laboratory reference methods remains limited, especially across extreme contamination gradients. Previous validation studies for handheld X-ray fluorescence (XRF) exhibit inconsistent results, highlighting the gap in methodological standardization. But some researchers like Weindorf, Zhu, Chakraborty, Bakr, and Huang (2012) found excellent correlation ($R^2 > 0.90$) and substantial systematic bias, especially for lead quantification, while others have identified significant technical difficulties (Weindorf et al., 2012). For example, Shuttleworth, Evans, Hutchinson, and Rothwell (2014) reported a systematic overestimation of lead (Pb), and Rouillon and Taylor (2016) showed that accuracy is highly dependent on matrix-matched calibrations. Despite the existence of USEPA guidelines for handheld XRF validation, their inconsistent implementation across the field continues to produce varied results (United States Environmental Protection Agency, 2007).

A critical knowledge gap exists in the magnitude and concentration dependency of measurement bias, the need to use (Bland & Altman, 1986) analysis rather than correlation alone for quantitative method agreement assessment and the development of validated experimental correction factors to improve field measurement accuracy. Also, evidence-based decision frameworks are lacking to guide users on when handheld XRF screening is sufficient and when laboratory investigation is required for identifying heavy metals and risk assessment in contaminated soil.

This study focuses on the validation of handheld XRF measurements against WDXRF reference standards across a contamination gradient, quantifying systematic bias, developing empirical correction factors, analyzing method agreement using (Bland & Altman, 1986) analysis and establishing a practical decision framework for handheld XRF environmental site surveying.

2. LITERATURE REVIEW

Soil contamination by heavy metals is known to be a serious problem all around the world. Kabata-Pendias (2011) has laid out the geochemical basis for trace elements including lead, copper, and zinc in the soils (Kabata-Pendias, 2011), whereas Wuana and Okieimen (2011) have provided a comprehensive review of anthropogenic sources, which include mining, industrial activities, agricultural practices, and historical pollution together with the remediation techniques (Wuana & Okieimen, 2011). Firing ranges are an acknowledged source of lead and copper pollution with sharp spatial gradients that are different from the agricultural loading. Multivariate statistical tools including PCA allow one to differentiate between geogenic and anthropogenic sources (Borůvka et al., 2005).

Laboratory analytical methods like WDXRF and ICP provide very accurate results and have low detection limits, but complex sample preparation procedures and the nature of laboratory work make them unsuitable for large scale field application. That has motivated the development of handheld X-ray fluorescence (hXRF). The works of Kalnicky

and Singhvi (2001) can be named as one of the first systematic reviews of field portable XRF, and the GIS-based contamination mapping was implemented by Carr et al. (2008). Handheld and laboratory XRF devices have a common physical basis – Moseley's law; however, the detection architecture is quite different (Beckhoff et al., 2006; Jenkins, 1999).

A common problem is the lower spectral resolution in energy-dispersive handheld instruments compared to the wavelength-dispersive laboratory instrumentation, making the latter more susceptible to overlapping peaks and matrix effects. The higher resolution of WDXRF over EDDXRF was confirmed by Potts et al. (1985) in relation to trace elements (Potts et al., 1985; Shackley, 2011), who observed matrix effects due to soil moisture, organic matter, and mineral composition affecting handheld instruments Shackley (2011). Tighe et al. (2004) added that soil moisture, surface roughness, and variability in particle sizes were sources of systematic errors in portable XRF (Tighe et al., 2004).

The results of validation tests are mixed. Weindorf et al. (2012) obtained a very high correlation ($R^2 > 0.90$), together with a significant bias (Weindorf et al., 2012), a hint at early recognition of the different nature of correlation and accuracy. Shuttleworth et al. (2014) demonstrated a systematic overestimation of lead concentrations in peatland soils Shuttleworth et al. (2014) and Rouillon and Taylor (2016) concluded that accuracy is dependent on matrix matching in calibration. There is a clear evolution of the field towards accuracy over the emphasis on pattern tracking in early studies. United States Environmental Protection Agency (2007) and its soil sampling SOP (U.S. Environmental Protection Agency Environmental Response Team, 2000) tried to address that problem, but implementation is still inconsistent among studies.

Much of the previous literature also uses correlation and R^2 as the sole means of assessing agreement. It was demonstrated by Bland and Altman (1986) that correlation is inadequate in the presence of constant or proportional bias and that difference vs mean plots along with limits of agreement should be used instead (Altman & Bland, 1983; Bland & Altman, 1986, 1999). This technique is rarely utilized for the hXRF soil validation. In cases when any attempts at correction were made, OLS regression is the typical calibration method employed (Draper & Smith, 1998); leave-one-out cross-validation is suggested for avoiding overfitting in small field samples (Hastie, Tibshirani, & Friedman, 2009; Stone, 1974).

In conclusion, literature sources indicate that hXRF can effectively define the boundaries of contamination; however, the concentration- and matrix-specific bias is always present in hXRF validation and cannot be quantified with the help of correlation. The firing range soils, which include extreme concentrations and Pb-Cu-Sb mixtures are poorly represented in the literature and there are very few cases when both (Bland & Altman, 1999) analysis and cross-validated correction have been applied.

3. MATERIALS AND METHODS

The soil samples are collected from a contamination gradient within the Mirpur Cantonment firing range, with three sampling points (S_1, S_2, S_3). The sampling points are taken in a manner of increasing distance from the primary contamination target to capture the spatial distribution of ballistic residues (Figure 1). This sampling design is taken to analyze the attenuation of heavy metals from high-impact firing zones to peripheral areas. Following the U.S. Environmental Protection Agency (2002) protocol, three replicate samples were prepared from each site (n=9) at a 0-15 cm depth using stainless steel trowels (U.S. Environmental Protection Agency Environmental Response Team, 2000). The collected soil samples were air dried, sieved through 1.18 mm mesh. The cone and quartering technique was followed to achieve homogenization. Sample characteristics and site-level Pb concentrations are summarized in Table 1.

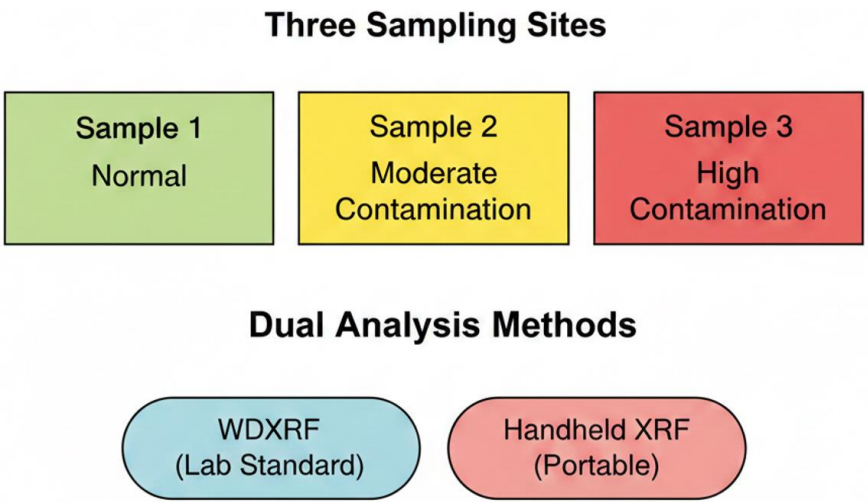

Figure 1. Study design showing three sampling sites with different contamination levels and dual analysis methods (WDXRF and Handheld XRF).

Table 1. Sample characteristics showing site groups, locations, contamination levels, and lead concentrations measured by both methods.

Grp	Sample ID	Pb Level	$WDXRF_{Pb}$ (%)	$Handheld_{Pb}$ (%)
S_1	S_{1_1}	Clean	0	0.08
	S_{1_2}	Clean	0	0.06
	S_{1_3}	Clean	0	0.05
S_2	S_{2_1}	Moderate	12.39	33.62
	S_{2_2}	Moderate	11.63	16.69
	S_{2_3}	Moderate	11.37	17.81
S_3	S_{3_1}	High	27.54	54.51
	S_{3_2}	High	28.67	60.16
	S_{3_3}	High	28.87	46.46

For laboratory analysis, an X-ray fluorescence spectrometer, XRF-1800 wavelength-dispersive XRF from Shimadzu, available in the Department of Nuclear Science and Engineering, Military Institute of Science and Technology, was used for this study. To minimize size effects on X-ray attenuation, approximately 10 grams of soil per sample were ground to <75 micrometers using a mortar and pestle. The samples were then pressed into 32-mm-diameter pellets at 15-ton pressure with boric acid backing. WDXRF measurements were standardized using established protocols, with calibration verified using certified reference materials. Detection limits ranged from 0.001–0.1%, depending on the element and matrix.

Handheld XRF of the Olympus Vanta C series available in the Department of Nuclear Science and Engineering, Military Institute of Science and Technology has been used for this study. This instrument was operated in Alloy Plus mode instead of Geochem mode, as preliminary testing has indicated superior performance for heavy metal concentrations, which is typical for contaminated industrial soils. Alloy Plus mode provides enhanced accuracy for lead, copper and other metals of concern in this study (Olympus Scientific Solutions Americas Inc, 2019). At 40 kV and 35 kV excitation energies, measurements have been taken with automatic beam switching to optimize sensitivity across the elemental range. Analysis time was 50 seconds per sample with triplicate measurements assessing precision. Certified reference material was used for quality control every five samples.

To evaluate the agreement between handheld XRF and WDXRF measurements, statistical analysis was performed to quantify systematic bias and develop empirical correction models. All analyses were performed using Python 3.9 with SciPy, Scikit-learn and Statsmodels libraries (Pedregosa et al., 2011; Seabold & Perktold, 2010; Virtanen et al., 2020). Following conventional practice in analytical chemistry and environmental science, statistical

significance was defined as $\alpha = 0.05$ for all tests, representing a 5% probability of Type I error (Miller, Miller, & Miller, 2018).

To analyse the strength and direction of linear relationships between WDXRF and handheld XRF measurements, Pearson correlation coefficients (r) were determined for each element. The Pearson correlation coefficient is defined by:

$$r = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{[\sum (x_i - \overline{x})^2 \cdot \sum (y_i - \overline{y})^2]}} \quad (2)$$

Where x_i represents the WDXRF concentrations, y_i represents the handheld XRF concentrations, $\overline{x}$ and $\overline{y}$ are their respective means and the summation is over all n samples (Pearson, 1896). It ranges from −1 to +1, with 0 indicating no linear relationship. Using a two-tailed t-test with n-2 degrees of freedom the statistical significance of correlation was assessed (Rodgers & Nicewander, 1988). The test statistic is:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \quad (3)$$

Which follows Student's t-distribution under the null hypothesis of no correlation. Correlation strength was interpreted using established guidelines: $|r| < 0.3$ as weak, $0.3 \leq |r| \leq 0.7$ as moderated and $|r| \geq 0.7$ as strong (Cohen, 1988).

To establish calibration relationships between the two measurement methods, ordinary least squares (OLS) linear regression models were developed. WDXRF concentration was defined as the dependent variable (y) and handheld XRF concentration as the independent variable (x), following the principle that the reference method should be predicted from the test method:

$$C_{WDXRF} = \beta_0 + \beta_1 \cdot C_{hXRF} + \epsilon \quad (4)$$

Where C_{WDXRF} is the predicted WDXRF concentration, C_{hXRF} is the determined handheld XRF concentration, β_0 is the intercept, β_1 is the slope and $\varepsilon$is the residual error (Altman & Bland, 1983). The regression coefficients were determined by Draper and Smith (1998).

$$\beta_1 = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sum (x_i - \overline{x})^2} \quad (5)$$

$$\beta_0 = \overline{y} - \beta_1 \cdot \overline{x} \quad (6)$$

The coefficient of determination (R^2) demonstrates the proportion of variance in WDXRF measurements explained by handheld XRF measurements.

$$R^2 = 1 - \left(\frac{S_{res}}{S_{tot}}\right) = 1 - \frac{\sum (y_i - \dot{y})^2}{\sum (y_i - \ddot{\overline{y}})^2} \quad (7)$$

Where $S_{res}$is the residual sum of the squares and S_{tot} is the total sum of the squares (Nagelkerke, 1991). This R^2 value ranges from 0 to 1; the higher the value the more fitted model.

Using root mean square error (RMSE) and mean absolute error (MAE), prediction accuracy was measured. RMSE is defined as:

$$RMSE = \sqrt{\frac{(y_i - \dot{y}_i)^2}{n}} \quad (8)$$

Where $y_i$is the observed WDXRF values, $\dot{y}_i$is the values predicted by the regression model and n is the number of samples (Willmott & Matsuura, 2005). RMSE is sensitive to large errors due to the squaring operation. MAE provides a metric less sensitive to outliers (Willmott & Matsuura, 2005).

$$MAE = \frac{y_i - \dot{y}_i}{n} \quad (9)$$

Relative error metrics were also calculated to evaluate the comparison with different concentration ranges.

$$\text{Relative Bias (\%)} \frac{\overline{y_{hXRF}}}{\overline{y_{WDXRF}}} \cdot 100 \qquad (10)$$

To assess the model's robustness and prevent overfitting, Leave-one-out cross-validation (LOOCV) was deployed. This is particularly important for the given limited sample size (Hastie et al., 2009). In this process, each sample is iteratively held back, and a model is trained on the remaining n-1 samples. Then the prediction error is calculated for the held-back sample. This process is iterated n times, and the overall prediction error is the average of all the iterations (Stone, 1974). The LOOCV R^2 is determined as:

$$R^2_{CV} = 1 - \frac{\sum (y_i - \dot{y}_i)^2}{\sum (y_i - \bar{\dot{y}}_i)^2} \qquad (11)$$

Where $\dot{y}_i$ is the predicted value for sample i when it is removed from the training set? A substantial difference between R^2 and R^2_{CV} depicts overfitting. Similarly, LOOCV RMSE was determined using:

$$RMSE_{cv} = \sqrt{\frac{\sum (y_i - \dot{y}_i)^2}{n}} \qquad (12)$$

Using Bland-Altman analysis, Method agreement was determined. This is the recommended approach for assessing the agreement between two measurement methods. Bland-Altman analysis technique involves plotting the difference between methods (handheld XRF - WDXRF) against their mean, allowing visualization of systematic bias and assessment of proportional bias (concentration-dependent error) (Bland & Altman, 1986, 1999). For each element, the mean difference (bias) was determined as:

$$Bias = \frac{1}{n} \cdot \sum \left(y_{i,hXRF} - y_{i,WDXRF} \right) \qquad (13)$$

Where $y_{i,hXRF}$ and $y_{i,WDXRF}$ represents the paired measurements from handheld XRF and WDXRF, respectively. A positive bias means that handheld XRF systematically overestimates concentrations relative to WDXRF, while negative bias indicates underestimation.

The standard deviation of the differences (SD) was determined as:

$$SD = \sqrt{\left[\frac{1}{n-1} \times \sum \{ \left(y_{i,hXRF} - y_{i,WDXRF} \right) - Bias \}^2} \qquad (14)$$

The 95% limits of agreement (LOA) were determined as Bland and Altman (1986).

$$UpperLOA = Bias + 1.96 \cdot SD \qquad (15)$$

$$LowerLOA = Bias - 1.96 \cdot SD \qquad (16)$$

These limits define the range of the two methods that are expected to fall within 95% of differences, assuming normal distribution of differences. Wide limits suggest substantial measurement variability, while narrow limits demonstrate good agreement. Proportional bias was assessed by linear regression of the differences against the mean concentrations (Bland & Altman, 1999). A significant non-zero slope depicts that the bias changes with concentration level, meaning that matrix effects or calibration issues vary across the measurement range.

To determine whether systematic bias exists between methods, paired t-tests were performed. The null hypothesis was that the mean difference between paired measurements equals zero ($H_0: \mu_d = 0$). The test hypothesis is:

$$t = \frac{\overline{d} - 0}{\frac{SD}{\sqrt{n}}} \qquad (17)$$

Where $\overline{d}$ is the mean difference, SD is the standard deviation of the differences and n is the sample size (Gosset, 1908). With n-1 degrees of freedom, the test statistic follows a Student's t-distribution. Statistical significance was evaluated at $\alpha = 0.05$, which is two-tailed.

To test statistically significant differences in elemental concentrations among the three sampling sites, one-way analysis of variance (ANOVA) was performed. The null hypothesis was that all group means are equal ($H_0: \mu_1 = \mu_2 = \mu_3$). The F-statistic is calculated as:

$$F = \frac{MS_{between}}{MS_{within}} = \frac{\frac{SS_{between}}{k-1}}{\frac{SS_{within}}{n-k}} \quad (18)$$

Where $MS_{between}$ represents the mean square value of between groups, MS_{within} represents the mean square value within groups, n is the total sample size and k is the number of sample groups (Fisher, 1925). The F-statistic is followed by an F-distribution with (k-1, n-k) degrees of freedom. Using Tukey (1949) Honest Significant Difference (HSD) test, post-hoc pairwise comparisons were performed when ANOVA indicated significant differences (Tukey, 1949), controlling the family-wise error rate at $\alpha = 0.05$.

Using the Shapiro and Wilk (1965) test and visual inspection of Q-Q plots, normality of residuals was assessed (Shapiro & Wilk, 1965). Residual plots were used to evaluate homoscedasticity (constant variance of residuals). For Bland and Altman (1999) analysis, normality of the differences was verified. When assumptions were violated, non-parametric alternatives (Spearman correlation, Wilcoxon signed-rank test) were considered, though parametric tests remained robust for the sample sizes examined (Norman, 2010). Residual diagnostic analysis was used to assure the validity of the linear regression model. Normality of residuals was verified using (Shapiro & Wilk, 1965) tests and visual inspection of Normal Q-Q plots. Homoscedasticity was evaluated via Residuals vs. Fitted plots, utilizing LOWESS smoothing via the statsmodels library to identify any non-linear bias.

4. RESULT AND DISCUSSION

A distinct contamination gradient is confirmed via WDXRF analysis across the sampling sites. The sampling site 1 (S_1), sampling site 2 (S_2) and sampling site 3 (S_3) showed minimal heavy metal contamination, with lead below detection limit (0.00 ± 0.00%), intermediate levels (Pb: 11.80 ± 0.38%, Cu: 1.07 ± 0.02%) of heavy metals and severe contamination (Pb: 28.36 ± 0.56%, Cu: 1.42 ± 0.17%), respectively (Table 2).

Table 2. WDXRF elemental concentrations (mean ± SD) by site group for key elements.

Grp	N	Pb (%)	Cu (%)	Fe (%)	Si (%)	Al (%)	Zn (%)	Mn (%)	Ti (%)	Ca (%)	K (%)
S_1	3	ND	ND	21.79 ± 0.95	47.48 ± 1.08	10.97 ± 0.38	0.12 ± 0.02	0.38 ± 0.06	2.20 ± 0.08	4.33 ± 1.48	11.34 ± 1.02
S_2	3	11.80±0.53	1.07 ± 0.02	19.52 ± 0.18	39.44 ± 0.30	6.62 ± 0.06	0.16 ± 0.01	0.30 ± 0.01	1.69 ± 0.15	10.33 ± 0.12	7.67 ± 0.12
S_3	3	28.36±0.72	1.42 ± 0.18	20.18 ± 0.30	33.66 ± 0.73	7.16 ± 0.31	0.19 ± 0.01	0.21 ± 0.01	1.86 ± 0.44	2.09 ± 0.04	4.16 ± 0.19

The Lead (Pb) and Copper (Cu) plots highlight the distinct contamination gradient at the firing range, while Iron (Fe) and Manganese (Mn) reflect geogenic baseline concentrations and anthropogenic dilution effects, respectively. The Lead (Pb) and Copper (Cu) plots highlight the distinct contamination gradient at the firing range. The wider interquartile range observed in Site 3 (S3) for lead suggests effectively capturing the heterogeneous nature of ballistic contamination, where particulate distribution is less uniform than in background soils. Corresponding handheld XRF measurements by site are presented in Table 3, showing the same qualitative gradient with higher absolute values.

Table 3. Handheld XRF elemental concentrations (mean ± SD) by site group for key elements.

Grp	N	Pb (%)	Cu (%)	Fe (%)	Si (%)	Al (%)	Zn (%)
S_1	3	0.06 ± 0.02	0.06 ± 0.00	31.75 ± 5.27	51.07 ± 0.14	10.93 ± 2.69	0.19 ± 0.04
S_2	3	22.71 ± 9.47	1.33 ± 0.51	23.20 ± 4.73	43.27 ± 12.77	6.38 ± 2.66	0.20 ± 0.10
S_3	3	53.71 ± 6.88	1.81 ± 0.66	21.15 ± 0.37	17.85 ± 6.11	3.49 ± 2.50	0.51 ± 0.53

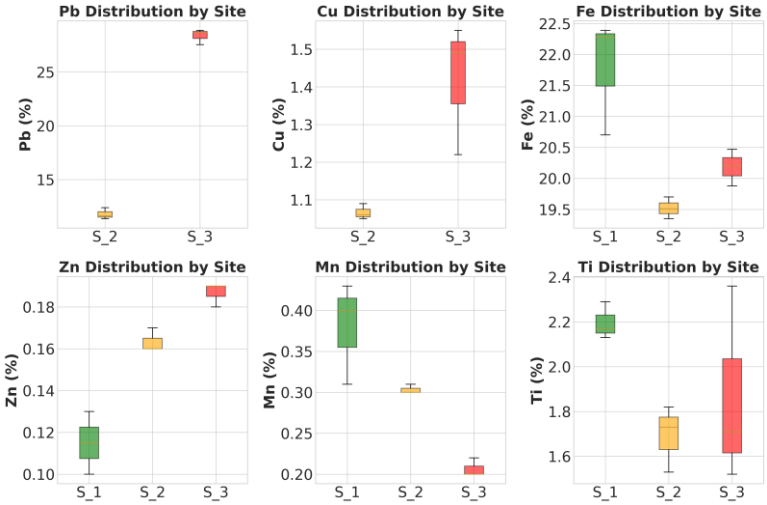

Figure 2. Box-and-whisker plots illustrating the elemental distribution across the three sampling sites (S_1: Green/Normal; S_2: Yellow/Intermediate; S_3: Red/Severe) as determined by WDXRF.

Iron concentrations showed less variation (S_1: 21.79 ± 0.97%, S_2: 19.52 ± 0.18%, S_3: 20.18 ± 0.30%), consistent with natural geogenic sources. Silicon concentrations decreased with increasing contamination (S_1: 47.48 ± 1.08%, S_2: 39.44 ± 0.30%, S_3: 33.66 ± 0.67%), likely reflecting dilution due to anthropogenic additions (Figure 2).

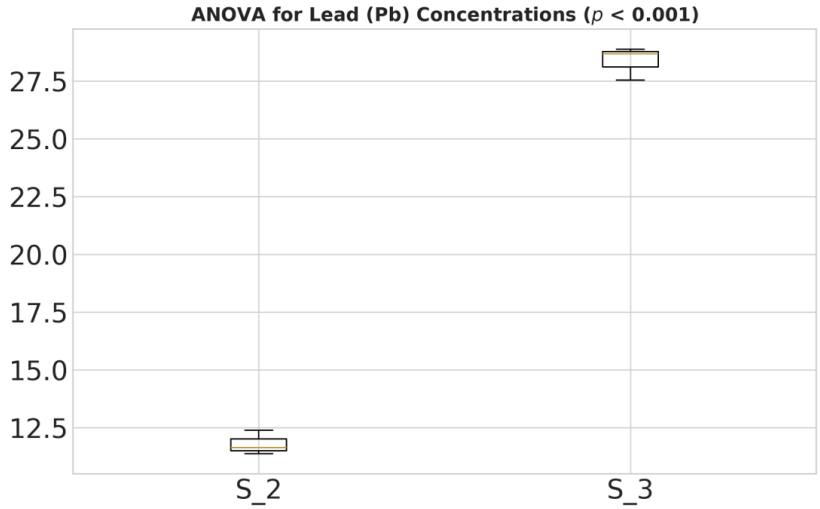

Figure 3. One-way ANOVA results for Lead (Pb) concentration indicating statistically significant differences ($p < 0.001$) among the stratified sampling sites.

ANOVA test demonstrated a statistically significant difference in lead concentrations ($F = 1437.8$, $p < 0.0001$) while validating the stratified sampling design. Figure 3 visually confirms the distinct stratification of the sampling design, with no overlap between the confidence intervals of the Clean (S1), Moderate (S2), and High (S3) groups, validating the distinct contamination zones.

Correlation analysis showed variable agreement between handheld XRF and WDXRF depending on the element. Lead (Pb) showed the strongest correlation ($r = 0.972$, $p < 0.0001$, $n = 9$), indicating handheld XRF reliably identified lead concentration trends despite systematic bias. Also, silicon showed a strong correlation ($r = 0.836$, $p = 0.0049$), while a moderate correlation ($r = 0.800$, $p = 0.0096$) was seen with iron. But copper showed a weaker positive correlation ($r = 0.741$, $p = 0.0566$, $n = 7$), not reaching statistical significance at $\alpha = 0.05$ (Table 4).

Table 4. Method validation statistics showing correlation coefficients, error metrics (RMSE, MAE), bias, and statistical significance for each element.

Element	N	Pearson r	P-value	RMSE	MAE	Bias (abs)	Bias (%)	t-statistic	t p-value	Significant
Pb	9	0.972	<0.0001	16.818	12.107	12.107	275.7	2.933	0.0189	Yes*
Cu	7	0.741	0.0566	0.568	0.434	0.285	107.1	1.426	0.2039	No
Fe	9	0.8	0.0096	6.884	4.953	4.868	23	2.829	0.0222	Yes*
Si	9	0.836	0.0049	11.683	10.085	-2.799	-9.9	-0.698	0.5049	No
Al	9	0.669	0.0486	3.61	3.012	-1.706	-23.7	-1.517	0.1677	No
Zn	9	0.276	0.4721	0.324	0.184	0.157	246.9	1.56	0.1575	No
Mn	9	0.546	0.1281	0.164	0.123	0.096	34.5	2.051	0.0744	No
Ti	9	0.153	0.6934	0.689	0.505	-0.049	-1.2	-0.202	0.8448	No

Note: N = number of paired handheld XRF–WDXRF measurements; r = Pearson correlation coefficient; RMSE = root mean square error; MAE = mean absolute error. RMSE, MAE and Bias (abs) are expressed in absolute concentration units (weight %), whereas Bias (%) is the relative bias with respect to the corresponding WDXRF value. Bias is calculated as the mean of (handheld XRF − WDXRF); positive values indicate overestimation and negative values indicate underestimation by the handheld instrument relative to WDXRF. The t-statistic and t p-value are obtained from a paired-samples t-test of the null hypothesis of zero mean difference between the two methods; the "Significant?" column therefore refers to the significance of the systematic bias between the two methods and not to the significance of the correlation coefficient. * Denotes a statistically significant systematic bias at $\alpha = 0.05$ (two-tailed t p-value < 0.05), i.e., for Pb and Fe the handheld XRF values differ significantly from the WDXRF reference values. N = 7 for Cu because two sample pairs were excluded in which the concentration reported by one of the methods fell below the corresponding detection limit.

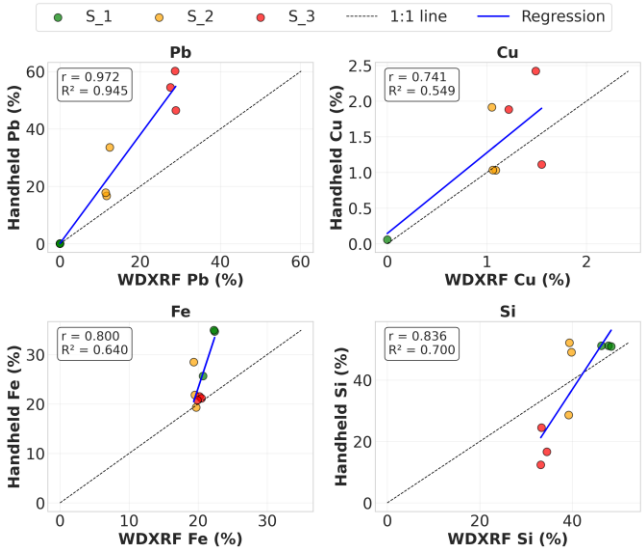

Figure 4. Method validation scatter plots comparing WDXRF and Handheld XRF measurements for Pb, Cu, Fe, and Si, showing 1:1 lines, regression lines, and correlation statistics.

As illustrated in Figure 4, the data points for Lead (Pb) consistently fall below the 1:1 identity line, visually confirming the systematic overestimation by the handheld XRF device prior to correction. Linear regression models provided correction equations for converting portable XRF readings to WDXRF-equivalent values.

$$Pb_{WDXRF} = 0.524 \cdot Pb_{hXRF} + 0.473 \qquad (19)$$

(R^2=0.945, LOOCV R^2=0.920)

This equation indicates handheld XRF systematically overestimates Lead concentration approximately twofold (slope < 1), with small positive interception. Cross-validation results (LOOCV R^2 = 0.920) confirmed model potency despite limited sample size. Regression parameters and correction equations for all elements are given in Table 5. The LOOCV R2 value for lead (-2.33) indicated moderate overfitting, where copper, iron and silicon show severely negative LOOCV R2 values (-169.4, -60.9, and -16.9, respectively), indicating extreme overfitting and poor generalization capability. Negative cross-validation R2 values occur when the model performs worse than simply predicting the mean values.

Table 5. Linear regression parameters (slope, intercept, R^2, LOOCV R^2) and correction equations for key elements.

Element	Slope	Intercept	R^2	LOOCV R^2	Equation	Correction
Pb	0.4992	0.6588	0.9454	-2.3305	WDXRF = 0.499 × Handheld + 0.659	Handheld × 0.499 + 0.659
Cu	0.4852	0.4102	0.5495	-169.447	WDXRF = 0.485 × Handheld + 0.410	Handheld × 0.485 + 0.410
Fe	0.15	16.6919	0.6403	-60.8854	WDXRF = 0.150 × Handheld + 16.692	Handheld × 0.150 + 16.692
Si	0.3044	28.8108	0.6997	-16.8818	WDXRF = 0.304 × Handheld + 28.811	Handheld × 0.304 + 28.811
Zn	0.0537	0.1262	0.0762	-0.9327	WDXRF = 0.054 × Handheld + 0.126	Handheld × 0.054 + 0.126

This is a reflection that means these correlation equations should not be applied to new samples without further validation. Despite the slightly negative LOOCV R^2 of the lead correlation model, this showed the strongest predictive performance in this study (R^2=0.945) and attained significant error reduction (83.8% improvement in RMSE). Larger sample size or more sophisticated modelling may help in enhancing the other correlation models. This finding shows the element-specific nature of correction factors and the need for cross-validation before deploying empirical calibration models in practice.

Element-specific patterns had been seen in the systematic bias analysis. Lead analysis showed substantial positive bias (mean difference = +12.11%, relative bias = +275.7%), with RMSE of 16.82% and MAE of 12.11%. Paired t-test confirmed statistical significance (t = 2.933, p = 0.0189). Iron showed moderate positive bias (+4.87%, +23.0%), which is statistically significant (t = 2.829, p = 0.0222). Silicon showed a small negative bias (-2.80%, -9.9%), which was not statistically significant (p = 0.5049).

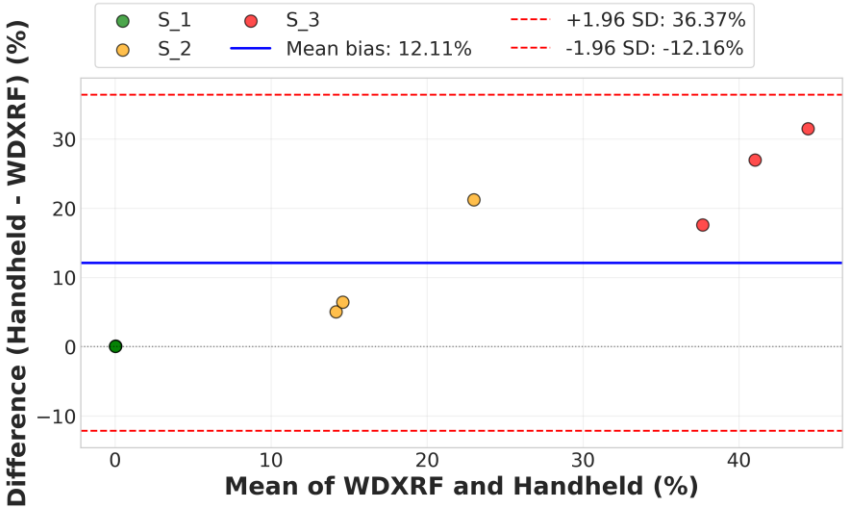

Figure 5. Bland-Altman plot for lead showing mean bias (12.11%) and limits of agreement (±1.96 SD), illustrating systematic positive bias in handheld XRF measurements.

Figure 5 reveals that most data points fall within the 95% limits of agreement. The positive mean bias line (solid line) being significantly above zero further corroborates the systematic overestimation identified in the regression analysis. Bland-Altman analysis for lead showed a mean bias of 12.11% with limits of agreement ranging from -12.16% to +36.37%, indicating wide variability in individual measurement differences. The proportional bias pattern suggested errors increase with concentration, likely due to matrix effects and spectral interferences at higher contamination levels. Application of regression-based correction factors significantly improved measurement accuracy (Table 6, Figure 6). For lead, uncorrected handheld XRF showed RMSE of 16.82% and MAE of 12.11%. RMSE decreased to 2.73% and MAE to 2.17% after applying the linear correction model, which represents an 83.8% improvement in RMSE. Correction reduced RMSE from 0.57% to 0.38% (33.7% improvement) for copper and iron showed RMSE reduction from 6.88% to 4.12% (40.2% improvement). These improvements show that empirical calibration can substantially enhance handheld XRF performance.

Table 6. Correction factor performance showing MAE and RMSE before and after correction, with percent improvement (reduction in RMSE) for Pb, Cu, and Fe.

Element	MAE (Original)	MAE (Corrected)	RMSE (Original)	RMSE (Corrected)	Improvement (%)	Recommended Use
Pb	12.107	2.106	16.818	2.72	83.8	With correction
Cu	0.434	0.265	0.568	0.318	43.9	Screening only
Fe	4.953	0.46	6.884	0.638	90.7	Screening only

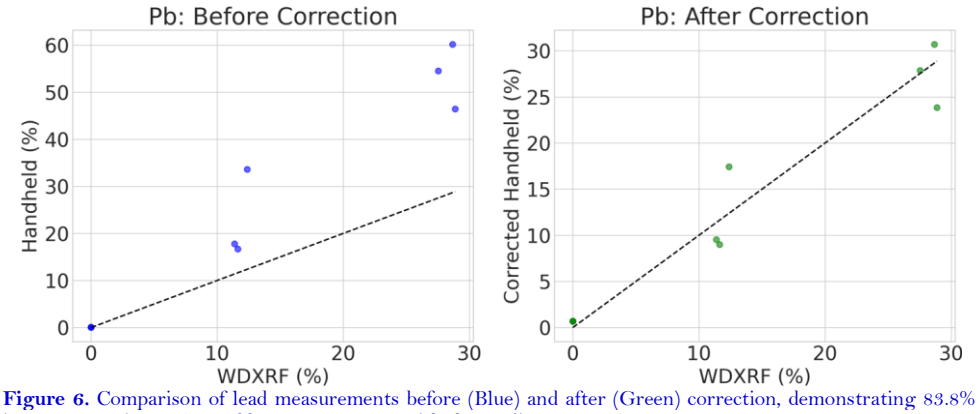

Figure 6. Comparison of lead measurements before (Blue) and after (Green) correction, demonstrating 83.8% improvement in RMSE and better agreement with the 1:1 line.

Model diagnostics (Figure 7) confirmed the suitability of the linear framework. The Residuals vs. Fitted plot exhibited a random distribution around the zero-line with no significant non-linear trends identified by the LOWESS smoothing.

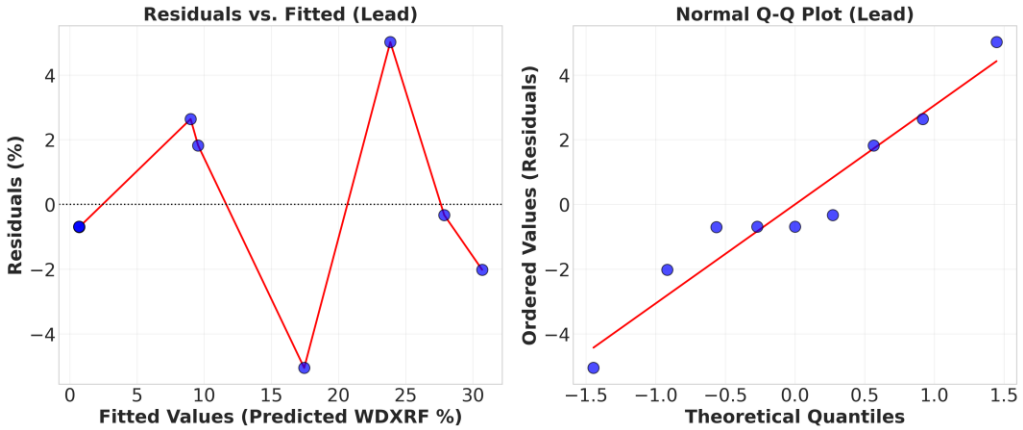

Figure 7. Residual diagnostic plots for the Lead (Pb) regression model: (a) Residuals vs. Fitted values with LOWESS smoothing to check homoscedasticity; (b) Normal Q-Q plot to verify normality of errors.

The systematic positive bias found in this study in handheld XRF lead measurements can happen due to several factors, despite the operation being done in Alloy Plus mode. While Alloy Plus mode works efficiently for metallic element quantification and provides better accuracy than geochem mode for highly contaminated samples, it does not remove all sources of bias. Firstly, complex soil composition causes a matrix effect that influences X-ray attenuation according to the Beer-Lambert law.

$$I = I_0 \cdot e^{-\mu\rho x} \qquad (20)$$

Where I represents transmitted intensity, I_0 is incident intensity, $\mu$is the mass attenuation coefficient, $\rho$is density, and x is the path length (Jenkins, 1999). Effective path length gets changed due to variation in soil density, moisture content, and elemental composition. Elemental interdependencies supporting this interpretation are shown in Figure 8.

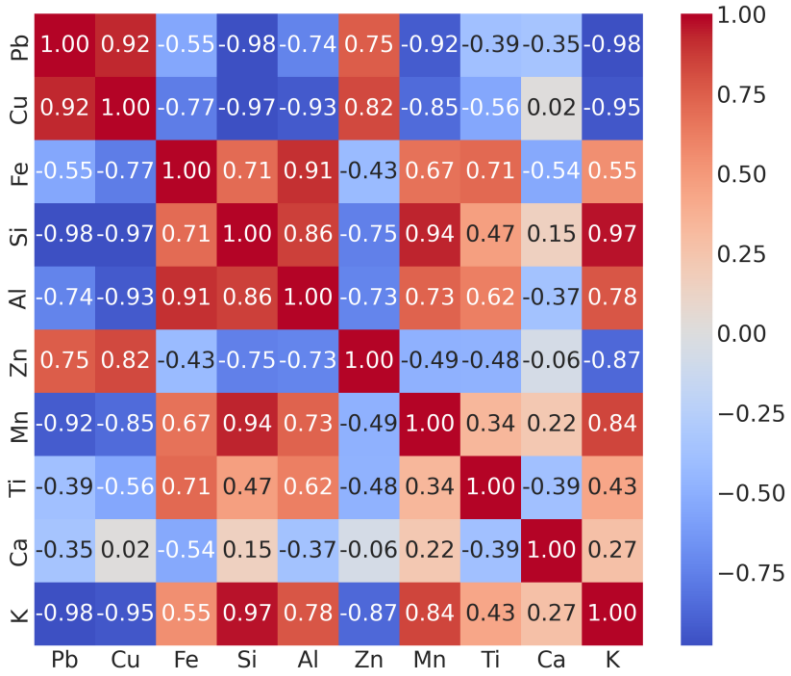

Figure 8. Elemental correlation matrix (WDXRF data) showing strong positive correlations between Pb-Cu (0.92), Si-K (0.97), and negative correlations between Pb-Si (-0.98).

Secondly, spectral interferences from overlapping characteristic X-ray lines may cause overestimation. Lead $L\alpha$(10.55 keV) can overlap with arsenic $k\alpha\alpha$ (10.54 keV) in arsenical soils. Also $L\beta$ lines may interfere with other elements (Shackley, 2011). Lastly, particle size influences effective sampling volume and analysis depth. The handheld XRF measurement depth for lead in soil matrices is around 1-2 mm, whereas WDXRF samples the entire pellet uniformly (Lemière, 2018). Thus, the samples of surface heterogeneity can bias the handheld XRF measurements.

This study demonstrates that handheld XRF can effectively screen for contamination gradients despite systematic bias. The strong correlation for lead exhibits excellent relative accuracy and the instrument correctly ranks samples by contamination level even though the concentrations are biased. This allows the handheld XRF to be identified as suitable for preliminary site assessment, delineating contamination boundaries, and guiding sampling strategies. But uncorrected handheld XRF data may lead to errored results for quantitative risk assessment or regulatory compliance. From the study, we got a 275.7% positive bias for lead, which would significantly overestimate exposure risks and remediation costs. Our correction model provides site-specific calibration enabling more accurate quantification, with an 83.8% reduction in RMSE demonstrating correction effectiveness.

The choice of Alloy Plus mode in this study requires discussion. While handheld XRF instruments typically offer multiple operating modes (e.g., Geochem mode, Alloy mode, Alloy Plus mode, etc), mode selection significantly impacts measurement accuracy and bias characteristics. Alloy Plus mode was selected for this study because

preliminary testing shows effective performance for high-concentration heavy metal measurements, which are typical of industrial contaminated sites. This mode uses more measurement times and calibration algorithms optimized for metallic elements, providing better sensitivity and accuracy for lead, copper, zinc, and other heavy metals (Olympus Scientific Solutions Americas Inc, 2019). However, this mode of selection may influence bias patterns differently than conventional geochem mode, and the correction factors developed in this study are more specific. Future validation studies should compare performance across different operating modes to establish comprehensive guidance for practitioners.

The results of this study align with and extend previous comparative studies. Weindorf et al. (2012) reported R^2 > 0.90 for lead in agricultural soils with moderate contamination in their study, which aligns with our R^2 = 0.945 (Weindorf et al., 2012). However, they observed less severe systematic bias, likely due to more sampling, lower contamination levels and different soil matrices. Shuttleworth et al. (2014) reported systematic lead overestimation in urban soils, particularly at high concentrations, directly supporting our observations. Our bias quantification (275.7%) exceeds their reported values, possibly reflecting more severe contamination gradients (up to 28.9% Pb). Rouillon and Taylor (2016) showed matrix-matched calibration importance; showing factory calibrations can produce biased results. Our site-specific correction models support this conclusion, showing significant accuracy improvements by empirical calibration. A key difference between this study and other validation work is the implementation of rigorous cross-validation (LOOCV) to assess the model generalizability. While high R^2 values mean model quality, our study shows that the R^2 value is insufficient, as cross-validation revealed that only the lead model generalizes adequately while other models show severe overfitting despite good R^2 values. This finding suggests that many published correction factors may be overoptimistic and require re-evaluation using proper cross-validation techniques.

Based on these findings, we propose a tiered decision framework for handheld XRF environmental site assessment (Figure 9). Handheld XRF provides rapid contamination mapping without correction factors for initial screening, which identifies high-risk areas requiring focused investigation. If the handheld XRF measurement crosses a threshold (e.g., >10% Pb), confirmatory WDXRF analysis is mandatory for regulatory decisions.

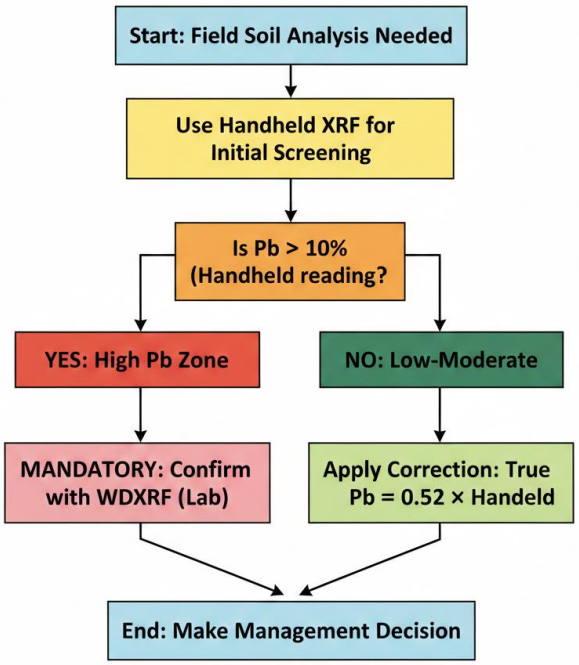

Figure 9. Proposed decision framework for appropriate use of portable XRF in field soil analysis, showing when to use screening, correction factors, or confirmatory WDXRF analysis.

For sites with moderate contamination where rapid decisions are needed, corrected handheld XRF values using site-specific regression models may be acceptable for preliminary risk assessment, though laboratory confirmation remains advisable. This framework balances cost-effectiveness and analytical difficulties, maximizing handheld XRF utility while maintaining data quality standards.

Several limitations required consideration. First, limited sample size (n = 9) limits the power of statistics and regression model generalizability. Although LOOCV partially addresses overfitting concerns, larger sample sets would strengthen conclusions. Second, the study focused on three sites in a specific geographic region; soil mineralogy, texture, and contamination source variations may affect handheld XRF performance differently elsewhere. Third, we examined only one handheld XRF model operated exclusively in Alloy Plus mode; different instruments with varying detector technologies or different operating modes (e.g., Geochem mode) may show different bias patterns. The correction factors developed here are specific to Alloy Plus mode and may not be directly applicable to other modes. Fourth, we analyzed dried, sieved samples rather than in-situ measurements. Field measurements may show additional variability from soil moisture and surface roughness. Finally, our correction models are element-specific and site-specific; extrapolation to other elements or sites should be validated.

5. CONCLUSION

This study provides systematic proof that handheld XRF operated in Alloy Plus mode can be used as an effective preliminary screening tool for heavy metal contamination assessment coupled with appropriate calibration protocols. We evaluated the analytical agreement between handheld XRF and WDXRF for assessing heavy metal contamination in firing range soils. Our findings confirm that while hXRF provides excellent relative accuracy, evidenced by the strong correlation for Lead, it is subjected to significant systematic positive bias when operated in Alloy Plus mode on highly contaminated samples. The observed 275.7% overestimation of lead highlights the risks of relying on raw field data for regulatory compliance or quantitative risk modeling.

However, the research successfully showed that site-specific empirical correction factors for lead can fill the gap between field screening and laboratory precision. By applying linear regression models, we achieved an 83.8% reduction in measurement error for Lead, proving that hXRF can be a reliable quantitative tool for this critical contaminant when properly calibrated against a reference subset. Moreover, models for copper, iron, and silicon showed severely negative cross-validation R^2 values (-169.4, -60.9, -16.9, respectively), indicating that these elements require other calibration approaches, larger training datasets, or more sophisticated modeling techniques (e.g., non-linear regression, machine learning) for reliable field quantification. This finding indicates that correction factor validity is highly element-specific and that cross-validation is absolutely essential before deploying any empirical calibration model in environmental practice.

Future research should examine larger sample sets across different soil types, compare performance across different handheld XRF operating modes, compare with different handheld XRFs with the same mode, test alternative correction algorithms including machine learning approaches, and validate in-situ field measurements against controlled laboratory conditions, using other parameters like humidity, pH, etc to get more clear idea of the reasons for biases. To enhance handheld XRF reliability for environmental applications, investigating moisture correction factors, mode-specific bias characterization and standardized quality assurance protocols are required. Doing all these may enhance the reliability of in-situ nuclear analytical techniques in environmental science.

Funding: This research was funded by the Military Institute of Science and Technology.
Institutional Review Board Statement: Not applicable.
Transparency: The authors state that the manuscript is honest, truthful, and transparent, that no key aspects of the investigation have been omitted, and that any differences from the study as planned have been clarified. This study followed all writing ethics.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: Both authors contributed equally to the conception and design of the study. Both authors have read and agreed to the published version of the manuscript.

REFERENCES

Altman, D. G., & Bland, J. M. (1983). Measurement in medicine: The analysis of method comparison studies. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *32*(3), 307-317. https://doi.org/10.2307/2987937

Beckhoff, B., Kanngießer, B., Langhoff, N., Wedell, R., & Wolff, H. (2006). *Handbook of practical X-ray fluorescence analysis*. Berlin, Germany: Springer-Verlag.

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, *327*(8476), 307-310. https://doi.org/10.1016/S0140-6736(86)90837-8

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, *8*(2), 135-160. https://doi.org/10.1177/096228029900800204

Borůvka, L., Vacek, O., & Jehlička, J. (2005). Principal component analysis as a tool to indicate the origin of potentially toxic elements in soils. *Geoderma*, *128*(3-4), 289-300. https://doi.org/10.1016/j.geoderma.2005.04.010

Carr, R., Zhang, C., Moles, N., & Harder, M. (2008). Identification and mapping of heavy metal pollution in soils of a sports ground in Galway City, Ireland, using a portable XRF analyser and GIS. *Environmental Geochemistry and Health*, *30*(1), 45-52. https://doi.org/10.1007/s10653-007-9106-0

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Draper, N. R., & Smith, H. (1998). *Applied regression analysis* (3rd ed.). New York, United States: John Wiley & Sons.

Fisher, R. A. (1925). *Statistical methods for research workers*. Edinburgh, Scotland: Oliver and Boyd.

Gosset, W. S. (1908). The probable error of a mean. *Biometrika*, *6*(1), 1–25.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer: New York, United States.

Jenkins, R. (1999). *X-Ray fluorescence spectrometry* (2nd ed.). New York: Wiley.

Kabata-Pendias, A. (2011). *Trace elements in soils and plants* (4th ed.). Boca Raton, FL: CRC Press.

Kalnicky, D. J., & Singhvi, R. (2001). Field portable XRF analysis of environmental samples. *Journal of Hazardous Materials*, *83*(1-2), 93-122. https://doi.org/10.1016/S0304-3894(00)00330-7

Lemière, B. (2018). A review of pXRF (field portable X-ray fluorescence) applications for applied geochemistry. *Journal of Geochemical Exploration*, *188*, 350-363. https://doi.org/10.1016/j.gexplo.2018.02.006

Miller, J. N., Miller, J. C., & Miller, R. D. (2018). *Statistics and chemometrics for analytical chemistry* (7th ed.). Harlow, England: Pearson Education Limited.

Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, *78*(3), 691–692. https://doi.org/10.1093/biomet/78.3.691

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education*, *15*(5), 625-632. https://doi.org/10.1007/s10459-010-9222-y

Olympus Scientific Solutions Americas Inc. (2019). *Vanta family X-ray fluorescence analyzer: User's manual*. Waltham, Massachusetts, United States: Olympus Scientific Solutions Americas Inc.

Pearson, K. (1896). Mathematical contributions to the theory of evolution—III. Regression, heredity, and panmixia. *Philosophical Transactions of the Royal Society of London. Series A*, *187*, 253–318.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., . . . Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Potts, P. J., Webb, P. C., & Watson, J. S. (1985). Energy-dispersive X-ray fluorescence analysis of silicate rocks: Comparisons with wavelength-dispersive performance. *Analyst*, *110*(5), 507-513. https://doi.org/10.1039/AN9851000507

Rodgers, J. L., & Nicewander, W. A. (1988). Thirteen ways to look at the correlation coefficient. *The American Statistician*, *42*(1), 59–66. https://doi.org/10.1080/00031305.1988.10475524

Rouillon, M., & Taylor, M. P. (2016). Can field portable X-ray fluorescence (pXRF) produce high quality data for application in environmental contamination research? *Environmental Pollution*, *214*, 255-264. https://doi.org/10.1016/j.envpol.2016.03.055

Seabold, S., & Perktold, J. (2010). *Statsmodels: Econometric and statistical modeling with Python.* Paper presented at the Proceedings of the 9th Python in Science Conference, SciPy, Austin, Texas, United States.

Shackley, M. S. (2011). *X-ray fluorescence spectrometry (XRF) in geoarchaeology.* New York, United States: Springer.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika, 52*(3-4), 591-611. https://doi.org/10.1093/biomet/52.3-4.591

Shuttleworth, E. L., Evans, M. G., Hutchinson, S. M., & Rothwell, J. J. (2014). Assessment of lead contamination in peatlands using field portable XRF. *Water, Air, & Soil Pollution, 225*(2), 1844. https://doi.org/10.1007/s11270-013-1844-2

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society: Series B (Methodological), 36*(2), 111-133. https://doi.org/10.1111/j.2517-6161.1974.tb00994.x

Tighe, M., Lockwood, P., Wilson, S., & Lisle, L. (2004). Comparison of digestion methods for ICP-OES analysis of a wide range of analytes in heavy metal contaminated soil samples with specific reference to arsenic and antimony. *Communications in Soil Science and Plant Analysis, 35*(9–10), 1369–1385. https://doi.org/10.1081/CSS-120037552

Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics, 5*(2), 99-114. https://doi.org/10.2307/3001913

U.S. Environmental Protection Agency. (2002). *Guidance on choosing a sampling design for environmental data collection: For use in developing a quality assurance project plan (EPA QA/G-5S; EPA/240/R-02/005).* Washington, DC: U.S. Environmental Protection Agency.

U.S. Environmental Protection Agency Environmental Response Team. (2000). *Soil sampling (Standard Operating Procedure No. 2012, Rev. 0.0).* New Jersey, United States: U.S. Environmental Protection Agency.

United States Environmental Protection Agency. (2007). *Method 6200: Field portable X-ray fluorescence spectrometry for the determination of elemental concentrations in soil and sediment (Revision 0; SW-846).* Washington, DC, United States: U.S. Environmental Protection Agency.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., . . . SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods, 17*(3), 261-272. https://doi.org/10.1038/s41592-019-0686-2

Weindorf, D. C., Zhu, Y., Chakraborty, S., Bakr, N., & Huang, B. (2012). Use of portable X-ray fluorescence spectrometry for environmental quality assessment of peri-urban agriculture. *Environmental Monitoring and Assessment, 184*(1), 217-227. https://doi.org/10.1007/s10661-011-1961-6

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research, 30*, 79-82. https://doi.org/10.3354/cr030079

Wuana, R. A., & Okieimen, F. E. (2011). Heavy metals in contaminated soils: A review of sources, chemistry, risks and best available strategies for remediation. *International Scholarly Research Notices, 2011*(1), 402647. https://doi.org/10.5402/2011/402647