

## Statistical evaluation of field-portable X-ray fluorescence soil preparation methods<sup>☆</sup>

Mark B. Bernick<sup>a,\*</sup>, Donna Getty<sup>a</sup>, George Prince<sup>b</sup>, Mark Sprenger<sup>b</sup>

<sup>a</sup> Roy F. Weston-REAC, GSA Raritan Depot, 2890 Woodbridge Avenue, Edison, NJ 08837, USA

<sup>b</sup> US EPA/Environmental Response Team, Raritan Depot-Building 18, MS-101, 2890 Woodbridge Avenue, Edison, NJ 08837, USA

Received 13 January 1995; accepted in revised form 1 February 1995

---

### Abstract

Traditional laboratory X-ray fluorescence (XRF) analysis has utilized sample preparation methods because the fixed nature of the instruments and absence of a surface probe precluded in-situ soil and sediment analysis. Currently, several field-portable XRF (FPXRF) instruments feature surface probes providing the option of in-situ soil and sediment analysis. In-situ FPXRF analysis can be a cost-effective near-real-time method to increase sampling densities due to the simplicity of the sample preparation. The following is a comprehensive statistical evaluation using current US EPA quality assurance guidelines for lead (Pb) data from a battery breakage site and zinc (Zn) and Pb data from a scrap metal site.

*Keywords:* X-ray fluorescence; Soil; Metals; In-situ XRF

---

### 1. Introduction

The US EPA/ERT has been using FPXRF spectrometers to characterize Superfund and hazardous waste sites. The Outokumpu Electronics Inc. (OEI) model X-MET 880<sup>®</sup> and the Spectrace Instruments model Spectrace 9000 FPXRF analyzers have surface probes and have been utilized by the ERT. In-situ FPXRF analysis can be a cost-effective method to increase sampling densities due to the simplicity of the sample preparation method. This improves the reliability of decisions based on spatial models delineating the extent of contamination [1]. Previous work has indicated that in-situ and prepared methods produced similar data sets [2, 3]. The FPXRF sample preparation and analysis methodologies, and QA/QC procedures discussed are

---

<sup>☆</sup>Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

\* Corresponding author. Tel.: +1 908-321-4200. Fax: +1-908-494-4021.

detailed in the paper entitled “Results of Field-Portable X-Ray Fluorescence Analysis of Metal Contaminants in Soil and Sediment” [4].

## 2. Data QA/QC

The FPXRF method detection limit (MDL) is defined as three times the calculated standard deviation value of the mean for each target element. Precision is defined as the coefficient of variation (COV) and should be  $\leq 20\%$  for the data to be considered adequately precise.

XRF data are accepted as US EPA QA1 and QA2, according to OSWER Directive 9360.4-01, “Quality Assurance/Quality Control Guidance for Removal Activities – Sampling QA/QC Plan and Data Validation Procedures”, April 1990. Both in-situ and prepared sample methods can be used with QA1 and QA2 objectives. QA1 is a screening objective. QA2 is a verification objective that requires confirmation of a minimum of 10% of the XRF samples by US EPA-approved laboratory (AA or ICP) methods. The regression analysis of AA versus XRF data sets must have a coefficient of determination ( $R^2$ ) of 0.7 or greater to meet QA2 objectives [5].

## 3. Statistical methods

### 3.1. Paired-difference *t*-test evaluation

The goal of the paired difference *t*-test is to determine if the mean difference of two populations of paired results is different from zero at an alpha level (significance level) equal to 0.05 [6]. If the data are from a normal distribution (bell curve), a two-tailed *t*-test is run on the differences. If the data are not from a normal distribution, a nonparametric Wilcoxon rank sum test is run on the differences. From either of these tests, a *p*-value is generated in the output. *p*-values are the lowest level at which the null hypothesis can be rejected. If the *p*-value is less than the significance level, the null hypothesis is rejected and it is determined that there are significant differences between the data sets. If the *p*-value is greater than the significance level, it is determined that there is no significant difference between the data sets. This does not mean that the data sets are equal, rather that they are not significantly different from each other. Paired difference *t*-test evaluations were performed on paired in-situ and prepared FPXRF sample data (Table 1).

### 3.2. Linear regression

Linear regression can be a useful tool to evaluate data sets. The coefficient of determination ( $R^2$ ) from linear regression analysis should be 0.7 or greater for the regression analysis to be considered significant, with a value of 1.00 being ideal. The slope from the linear regression analysis can theoretically be any number, with 1.00 being ideal. The slope can be considered a proportionality value, with the ideal slope

Table 1  
Paired in-situ and prepared FPXRF data; paired difference *t*-test evaluation results

Instrument	Waste type	Element	No. of samples	Test used	Calculated <i>p</i> -value	Accept/reject
OEI X-MET 880	Battery breakage	Pb	21	Wilcoxon rank sum	0.0851	Accept
OEI X-MET 880	Scrap metal	Pb	51	Wilcoxon rank sum	0.0087	Reject
Spectrace 9000	Scrap metal	Pb	51	Wilcoxon rank sum	0.0001	Reject
OEI X-MET 880	Scrap metal	Zn	46	Wilcoxon rank sum	0.0774	Accept
Spectrace 9000	Scrap metal	Zn	51	Wilcoxon rank sum	0.0001	Reject

of 1.00 meaning that the FPXRF and AA data are in a 1:1 proportion. Linear regression was run on the FPXRF in-situ FPXRF versus AA, and prepared FPXRF versus AA sample data (Table 2).

### 3.3. Statistical evaluation of regression coefficients

A statistical comparison of slopes (regression coefficients) was performed to compare AA in-situ FPXRF and AA prepared sample FPXRF regression results (Table 3). A methodology was utilized which is similar to that for testing the difference between two population means with the Student's *t*-test [7]. The null hypothesis for this test is  $\beta_1 = \beta_2$ , where  $\beta$  represents the true population regression coefficient. The alternative hypothesis is:  $\beta_1$  does not equal  $\beta_2$ . In all cases, the alpha level (significance level), the probability of rejecting the null hypothesis when it is in fact true, was set equal to 0.05.

A *p*-value is generated in the output of the test. *p*-values are the lowest level at which the null hypothesis can be rejected. If the *p*-value is less than the significance level, the null hypothesis is rejected and it is determined that there is a significant difference between the population regression coefficients. If the *p*-value is greater than the significance level, it is determined that there is no significant difference between the population regression coefficients. This does not mean that the coefficients are equal, rather that they are not significantly different from each other.

## 4. Discussion of results

### 4.1. Paired-difference *t*-test evaluation

All of the paired data sets evaluated were not from a normal distribution. Therefore, the Wilcoxon rank sum test was used (Table 1). The calculated *p*-values were less than the significance level for the scrap metal Spectrace 9000 Pb and Zn, and the OEI

Table 2  
Atomic absorption versus FPXRF data regression analysis, method detection limit and precision results

Instrument	Waste type	Element	Preparation method	MDL <sup>a</sup> (mg/kg) <sup>b</sup>	Precision		Regression statistics	
					Concentration	COV (%) <sup>c</sup>	COD <sup>d</sup>	Slope
OEI X-MET 880	Battery breakage	Pb	P <sup>e</sup>	123	300	13.7	0.97	0.84
			I <sup>f</sup>	123	300	13.7	0.92	0.92
OEI X-MET 880	Scrap metal	Pb	I	93	772	4.0	0.89	1.98
			P	93	772	4.0	0.92	1.78
Spectrace 9000	Scrap metal	Pb	I	123	972	4.4	0.87	1.09
			P	123	972	4.4	0.83	1.04
OEI X-MET 880	Scrap metal	Zn	I	111	556	6.7	0.98	1.95
			P	111	556	6.7	0.96	1.95
Spectrace 9000	Scrap metal	Zn	I	159	783	6.8	0.97	0.78
			P	159	783	6.8	0.98	0.60

<sup>a</sup> Minimum detection limit.

<sup>b</sup> Milligrams per kilogram.

<sup>c</sup> Coefficient of variation in percent.

<sup>d</sup> Coefficient of determination (*R*-square) for the regression.

<sup>e</sup> Prepared samples.

<sup>f</sup> In-situ analysis.

Table 3  
Regression coefficient comparison of atomic absorption (AA)/in-situ FPXRF and AA prepared FPXRF regression results

Instrument	Waste type	Element	Calculated $p$ -value	Accept/reject
OEI X-MET 880	Battery breakage	Pb	$p > 0.50$	Accept
OEI X-MET 880	Scrap metal	Pb	$0.10 < p < 0.20$	Accept
Spectrace 9000	Scrap metal	Pb	$0.20 < p < 0.50$	Accept
OEI X-MET 880	Scrap metal	Zn	$p > 0.50$	Accept
Spectrace 9000	Scrap metal	Zn	$p < 0.001$	Reject

X-MET 880 Pb data sets. The null hypothesis was rejected and it was determined that there was a significant difference between the data sets. The calculated  $p$ -values were greater than the significance level for the scrap metal OEI X-MET 880 Zn and the battery breakage OEI X-MET 880 Pb data sets. It was determined that there was no significant difference between these data sets.

#### 4.2. Linear regression

All of the regression analysis of AA versus XRF data sets have a coefficient of determination ( $R^2$ ) of 0.7 or greater meeting QA2 objectives.

#### 4.3. Statistical evaluation of regression coefficients

Comparison of X-MET 880 battery breakage Pb AA versus in-situ FPXRF, and AA versus prepared sample FPXRF regression results indicated slopes of 0.92 and 0.84, respectively. When applying the Student's  $t$ -test, no significant difference could be found between these two slopes ( $p$ -value  $> 0.50$ ), indicating that they came from the same  $\beta$  population, and that the regression lines can be assumed to be parallel.

Similar results were achieved for scrap metal comparisons for X-MET 880 Pb and Zn and Spectrace 9000 Pb regressions. Results of the X-MET 880 Pb in-situ FPXRF regression (slope = 1.98) versus the prepared sample FPXRF regression (slope = 1.78) also showed no statistical difference between the slopes with  $0.10 < p$ -value  $< 0.20$ . Results of the X-MET 880 Zn in-situ FPXRF regression (slope = 1.98) versus the prepared sample FPXRF regression (slope = 1.78) also showed no statistical difference between the slopes with  $0.10 < p$ -value  $< 0.20$ . Spectrace 9000 comparisons gave the same results as well, with the Pb in-situ FPXRF slope = 1.09, prepared sample FPXRF slope = 1.04, and  $0.20 < p$ -value  $< 0.50$ . In all three cases the null hypothesis could not be rejected, which supports the theory that the true population regression coefficients are in fact the same between AA in-situ and AA prepared sample FPXRF results. Therefore, the pairs of regression lines can be assumed to be parallel in each case.

Results of the scrap metal Spectrace 9000 Zn in-situ FPXRF regression (slope = 0.78) versus the prepared sample FPXRF regression (slope = 0.60) showed a statistical difference between the slopes with  $p$ -value  $< 0.001$ . This indicates that

they came from different  $\beta$  populations and that the regression lines cannot be assumed to be parallel.

#### 4.4. QA/QC

FPXRF Pb detection limits in Table 2 are significantly below typical Pb action levels of 400–2000 milligrams per kilogram (mg/kg); and the Pb and Zn precision is less than 20% COV. All FPXRF Pb and Zn AA versus XRF regression results meet the QA2 data objective coefficient of determination ( $R^2$ ) specification of 0.7 or greater.

### 5. Conclusions

Paired difference *t*-test evaluations illustrate that dissimilarities do exist between the in-situ and prepared FPXRF population distributions. However, in four out of five method regression comparisons, enough evidence did not exist to refute the hypothesis that the true population coefficients are the same and that the regression lines are parallel. Additionally, QA2 data objectives for precision and confirmation analysis regression criteria ( $R^2$ ) were satisfied in all cases.

Paired difference *t*-test evaluations of in-situ and prepared FPXRF data sets provide insight into dissimilar population distributions. However, QA2 criteria do not address population distributions. Rather, QA2 utilizes the regression analysis to evaluate the strength of the linear relationship between FPXRF results relative to the US EPA approved AA results. Therefore, evaluation of the FPXRF method regression relationships appears to provide more relevant information.

Generally, the results indicate that both methods appear comparable when (AA versus FPXRF) regressions are evaluated relative to the confirmation AA analysis. Therefore, in-situ FPXRF analysis should be considered as a viable alternative to prepared sample FPXRF analysis.

### References

- [1] W.H. Cole III, R.E. Enwall, G.A. Raab and C.A. Kuharic, Proc. 2nd Int. Symp. on Field Screening Methods for Hazardous Wastes and Toxic Chemicals, US Environmental Protection Agency, EMSL, Las Vegas, NV, 1991, p. 497.
- [2] M. Bernick, M. Sprenger, G. Prince, L. Kaelin, D. Idler, D. Miller and J. Patel, Proc. 2nd Int. Symp. on Field Screening Methods for Hazardous Wastes and Toxic Chemicals, US Environmental Protection Agency, EMSL, Las Vegas, NV, 1991, p. 603.
- [3] M. Bernick, G. Prince, R. Singhvi, D. Kalnicky and L. Kaclin, Petro-Safe '94 Conf., Book II, p. 195.
- [4] M. Bernick, D. Kalnicky, G. Prince and R. Singhvi, Special Issue on Field and In-Situ Analysis in the Journal of Hazardous Materials, 1995.
- [5] Field Portable X-ray Fluorescence, US Environmental Protection Agency/ERT, Quality Assurance Technical Information Bulletin, Vol. 1, No. 4, May 1991.
- [6] S.D. Schlotzhauer and R.C. Littell, SAS<sup>®</sup> system for elementary statistical analysis, 1987, SAS Institute Inc., Cary, NC, p. 200.
- [7] J.H. Zar, Biostatistical Analysis, Prentice-Hall, Englewood Cliffs, NJ, 2nd edn., 1984, p. 292.