United States
Environmental Protection
Agency

Office of Research and
Development
Washington, D.C. 20460

EPA/600/R-97/150
March 1998

EPA

Environmental Technology Verification Report

Field Portable X-ray Fluorescence Analyzer

Niton XL Spectrum Analyzer

SITE
SUPERFUND INNOVATIVE
TECHNOLOGY EVALUATION

ETV
Environmental Technology
Verification Program

058CMB98

EPA/600/R-97/150
March 1998

Environmental Technology Verification Report

Field Portable X-ray Fluorescence Analyzer

Niton XL Spectrum Analyzer

U.S. Environmental Protection Agency
Office of Research and Development

Notice

The information in this document has been funded wholly or in part by the U.S. Environmental Protection Agency (EPA) under Contract No. 68-CO-0047 to PRC Environmental Management, Inc. This work supports the Superfund Innovative Technology Evaluation Program administered by the National Risk Management Research Laboratory, Cincinnati, Ohio. This demonstration was conducted under the Monitoring and Measurement Technologies Program which is managed by the National Exposure Research Laboratory–Environmental Sciences Division, Las Vegas, Nevada. It has been subjected to the Agency's peer and administrative review, and has been approved for publication as an EPA document. Mention of corporation names, trade names, or commercial products does not constitute endorsement or recommendation for use of specific products.

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
Office of Research and Development
Washington, D.C. 20460

ENVIRONMENTAL TECHNOLOGY
VERIFICATION PROGRAM

ENVIRONMENTAL TECHNOLOGY VERIFICATION PROGRAM
VERIFICATION STATEMENT

TECHNOLOGY TYPE:	**FIELD PORTABLE X-RAY FLUORESCENCE ANALYZER**
APPLICATION:	**MEASUREMENT OF METALS IN SOIL**
TECHNOLOGY NAME:	**XL SPECTRUM ANALYZER**
COMPANY:	**NITON CORPORATION**
ADDRESS:	**74 LOOMIS ST.** **P.O. BOX 368** **BEDFORD, MA 01730-0368**
PHONE:	**(617) 275-9275**

The U.S. Environmental Protection Agency (EPA) has created a program to facilitate the deployment of innovative technologies through performance verification and information dissemination. The goal of the Environmental Technology Verification (ETV) Program is to further environmental protection by substantially accelerating the acceptance and use of improved and more cost-effective technologies. The ETV Program is intended to assist and inform those involved in the design, distribution, permitting, and purchase of environmental technologies. This document summarizes the results of a demonstration of the Niton XL Spectrum Analyzer.

PROGRAM OPERATION

The EPA, in partnership with recognized testing organizations, objectively and systematically evaluates the performance of innovative technologies. Together, with the full participation of the technology developer, they develop plans, conduct tests, collect and analyze data, and report findings. The evaluations are conducted according to a rigorous demonstration plan and established protocols for quality assurance. The EPA's National Exposure Research Laboratory, which conducts demonstrations of field characterization and monitoring technologies, selected PRC Environmental Management, Inc., as the testing organization for the performance verification of field portable X-ray fluorescence (FPXRF) analyzers.

DEMONSTRATION DESCRIPTION

In April 1995, the performance of seven FPXRF analyzers was determined under field conditions. Each analyzer was independently evaluated by comparing field analysis results to those obtained using approved reference methods. Standard reference materials (SRM) and performance evaluation (PE) samples also were used to independently assess the accuracy and comparability of each instrument.

The demonstration was designed to detect and measure a series of inorganic analytes in soil. The primary target analytes were arsenic, barium, chromium, copper, lead, and zinc; nickel, iron, cadmium, and antimony were secondary analytes. The demonstration sites were located in Iowa (the RV Hopkins site) and Washington (the ASARCO site). These sites were chosen because they exhibit a wide range of concentrations for most of the target metals and are located in different climatological regions of the United States; combined, they exhibit three distinct soil types: sand, clay, and loam. The conditions at these sites are representative of those environments under which the technology would be expected to operate. Details of the demonstration, including a data summary and discussion of results, may be found in the report entitled "Environmental Technology Verification Report, Field Portable X-ray Fluorescence Analyzer, Niton XL Spectrum Analyzer." The EPA document number for this report is EPA/600/R-97/150.

The EPA Method 6200 was tested and validated using the data derived from this demonstration. This method may be used to support the general application of FPXRF for environmental analysis.

TECHNOLOGY DESCRIPTION

This analyzer operates on the principle of energy dispersive X-ray fluorescence spectroscopy where the characteristic components of the excited X-ray spectrum are analyzed directly by an energy proportional response in an X-ray detector. Energy dispersion affords a highly efficient, full-spectrum measurement that enables the use of low intensity excitation sources (such as radioisotopes) and compact battery-powered, field-portable electronics. FPXRF instruments are designed to provide rapid analysis of metals in soil. This information allows investigation and remediation decisions to be made on-site and reduces the number of samples that need to be submitted for laboratory analysis. In the operation of these instruments, the user must be aware that FPXRF analyzers do not respond well to chromium and that detection limits may be 5 to 10 times greater than conventional laboratory methods. As with all field collection programs, a portion of the samples should be sent to a laboratory for confirmatory analyses.

The Niton XL Spectrum Analyzer was originally designed to produce quantitative data for lead in painted surfaces. This demonstration found that it could also provide quantitative data for metals contamination in soil. It is a hand-held instrument, weighing less than 3 pounds, and can be battery powered up to 8 hours. It uses a single radioactive source (cadmium-109) and a silicon pin-diode detector for the analysis of metals in soil using relatively short count times (60 seconds). The single radioactive source limits the number of analytes that can be detected. For the purposes of this demonstration, the XL Spectrum Analyzer's "SOILAIR" software was configured to report concentrations of arsenic, chromium, copper, lead, and zinc in soil samples. The XL Spectrum Analyzer was initially calibrated by the developer using the Compton normalization method to quantitate metals. The XL Spectrum Analyzer can conduct *in situ* measurements or measure samples in cups. The cost of the Niton XL Spectrum Analyzer was $11,990 at the time of the demonstration, or it could be leased for $2,200 per month.

VERIFICATION OF PERFORMANCE

The performance characteristics of the Niton XL Spectrum Analyzer include the following:

- **Detection limits:** Precision-based detection limits were determined by collecting 10 replicate measurements on site-specific soil samples with metals concentrations 2 to 5 times the expected MDLs. The results were 130 milligrams per kilogram (mg/kg) or less for all of the reported target analytes except chromium, which was determined to be 900 mg/kg.
- **Throughput:** Average throughput was 20 - 25 analyses per hour using a live count time of 60 seconds. This rate only represents the analysis time since different personnel were used to prepare the samples.
- **Drift:** This was evaluated using the results of an analysis of an SRM calibration check sample which contained quantifiable levels of arsenic, copper, lead, zinc, and iron. Over the course of the demonstration, this sample was analyzed approximately 100 times. The mean recovery for these analytes was between 85 and 140 percent. The drift RSD for the mean recovery of these analytes was less than 8 percent.
- **Completeness:** The XL Spectrum Analyzer produced results for 1,258 of the 1,260 samples for a completeness of 99.8 percent. The two lost data points were a consequence of operator error.
- **Blank results:** More than 100 lithium carbonate blanks were analyzed during the demonstration. None of the reported analytes were observed above the method detection limits.
- **Precision:** The goal of the demonstration was to achieve relative standard deviations (RSD) less than 20 percent at analyte concentrations of 5 to 10 times the method detection limits. The RSD value for arsenic was

9.2 percent, 13.2 percent for copper, 6.5 percent for lead, and 11.2 percent for zinc. Chromium was not reported due in part to the short 60 live-second count time.

- **Accuracy:** Intramethod accuracy was assessed using site-specific soil PE samples and soil SRMs. The data showed that 18 of 28 or 64.2 percent of the PE sample analytes had recoveries within the quantitative acceptance range of 80 - 120 percent. For the soil SRMs, 11 of 16 (68.7 percent) of the results were within the 80 - 120 percent recovery range.
- **Comparability:** This demonstration showed that the XL Spectrum Analyzer produced data that exhibited a $\log_{10}$-$\log_{10}$ linear correlation to the reference data. The coefficient of determination (r^2) which is a measure of the degree of correlation between the reference and field data was 0.82 for arsenic, 0.50 for chromium, 0.92 for copper, 0.96 for lead, and 0.89 for zinc.
- **Data quality levels:** Using the demonstration derived precision RSD results and the coefficient of determination as the primary qualifiers, the XL Spectrum Analyzer produced definitive level data for lead and data of quantitative screening level for arsenic, copper, and zinc. Since a precision RSD value was not determined for chromium, no data quality level can be assigned.

The results of the demonstration show that the Niton XL Spectrum Analyzer can provide useful, cost-effective data for environmental problem-solving and decision-making. Undoubtedly, it will be employed in a variety of applications, ranging from serving as a complement to data generated in a fixed analytical laboratory to generating data that will stand alone in the decision-making process. As with any technology selection, the user must determine what is appropriate for the application and the project data quality objectives.

Gary J. Foley, Ph.D.
Director
National Exposure Research Laboratory
Office of Research and Development

NOTICE: EPA verifications are based on an evaluation of technology performance under specific, predetermined criteria and the appropriate quality assurance procedures. EPA makes no expressed or implied warranties as to the performance of the technology and does not certify that a technology will always, under circumstances other than those tested, operate at the levels verified. The end user is solely responsible for complying with any and all applicable Federal, State, and Local requirements.

Foreword

The U.S. Environmental Protection Agency (EPA) is charged by Congress with protecting the Nation's land, air, and water resources. Under a mandate of national environmental laws, the Agency strives to formulate and implement actions leading to a compatible balance between human activities and the ability of natural systems to support and nurture life. To meet this mandate, the EPA's Office of Research and Development (ORD) provides data and science support that can be used to solve environmental problems and to build the scientific knowledge base needed to manage our ecological resources wisely, to understand how pollutants affect our health, and to prevent or reduce environmental risks.

The National Exposure Research Laboratory (NERL) is the Agency's center for the investigation of technical and management approaches for identifying and quantifying risks to human health and the environment. Goals of the Laboratory's research program are to develop and evaluate technologies for the characterization and monitoring of air, soil, and water; support regulatory and policy decisions; and provide the science support needed to ensure effective implementation of environmental regulations and strategies.

The EPA's Superfund Innovative Technology Evaluation (SITE) Program evaluates technologies for the characterization and remediation of contaminated Superfund and Resource Conservation and Recovery Act (RCRA) corrective action sites. The SITE Program was created to provide reliable cost and performance data to speed the acceptance of innovative characterization and monitoring technologies.

Effective measurement and monitoring technologies are needed to assess the degree of contamination at a site, to provide data which can be used to determine the risk to public health or the environment, to supply the necessary cost and performance data to select the most appropriate technology, and to monitor the success or failure of a remediation process. One component of the SITE Program, the Monitoring and Measurement Technologies Program, demonstrates and evaluates innovative technologies to meet these needs.

Candidate technologies can originate from within the federal government or from the private sector. Through the SITE Program, developers are given the opportunity to conduct a rigorous demonstration of their technology's performance under realistic field conditions. By completing the evaluation and distributing the results, the Agency establishes a baseline for acceptance and use of these technologies. The Monitoring and Measurement Technologies Program is managed by ORD's Environmental Sciences Division in Las Vegas, Nevada.

Gary J. Foley, Ph.D.
Director
National Exposure Research Laboratory
Office of Research and Development

Abstract

In April 1995, the U.S. Environmental Protection Agency (EPA) conducted a demonstration of field portable X-ray fluorescence (FPXRF) analyzers. The primary objectives of this demonstration were (1) to determine how well FPXRF analyzers perform in comparison to a standard reference method, (2) to identify the effects of sample matrix variations on the performance of FPXRF, (3) to determine the logistical and economic resources needed to operate these analyzers, and (4) to test and validate an SW-846 draft method for FPXRF analysis. The demonstration design was subjected to extensive review and comment by the EPA's National Exposure Research Laboratory, EPA Regional and Headquarters Superfund technical staff, the EPA's Office of Solid Waste–Methods Section, and the technology developers.

Two sites were used for this demonstration: RV Hopkins and the ASARCO Tacoma Smelter. RV Hopkins is an active steel drum recycling facility and the site of a former battery recycling operation. It is located in Davenport, Iowa. The ASARCO site is a former copper and lead smelter and is located in Tacoma, Washington. The samples analyzed during this demonstration, were evenly distributed between three distinct soil textures: sand, loam, and clay. In addition, four sample preparation steps were evaluated. The reference methods used to evaluate the comparability of data were EPA SW-846 Methods 3050A and 6010A, "Acid Digestion of Sediments, Sludges, and Soils" and "Inductively Coupled Plasma-Atomic Emission Spectroscopy," respectively.

The FPXRF analyzers tested in this demonstration were designed to provide rapid, real-time analysis of metals concentrations in soil samples. This information will allow investigation and remediation decisions to be made on-site more efficiently and can reduce the number of samples that need to be submitted for confirmatory analysis. Of the seven commercially available analyzers tested, one is manufactured by Niton Corporation (the XL Spectrum Analyzer); two are manufactured by TN Spectrace (the TN 9000 and TN Pb Analyzer); two are manufactured by Metorex Inc. (the X-MET 920-P Analyzer and the X-MET 920-MP Analyzer); one is manufactured by HNU Systems, Inc. (the SEFA-P Analyzer); and one is manufactured by Scitec Corporation (the MAP Spectrum Analyzer). The X-MET 940, a prototype FPXRF analyzer developed by Metorex, was given special consideration and replaced the X-MET 920-P for part of the RV Hopkins sample analyses. This environmental technology verification report (ETVR) presents information relative to the XL Spectrum Analyzer developed by Niton. Separate ETVRs have been published for the other analyzers demonstrated.

No operational downtime was experienced by the Niton analyzer through the 20 days required to conduct this demonstration. Quantitative data was provided by the analyzer on a real-time basis. The XL Spectrum Analyzer was configured to report arsenic, chromium, copper, lead, and zinc. This analyzer used relatively short count times of 60 live-seconds for this demonstration. This relatively short count time resulted in a high sample throughput, averaging between 20 and 25 samples per hour. The XL Spectrum Analyzer provided definitive level data quality (equivalent to reference quality data) for lead, and

quantitative screening level data quality (not equivalent to reference data but correctable with the analysis of confirmatory samples) for arsenic, copper, and zinc. No data quality assessment could be made for chromium since the short count time made the precision and method detection limit measurements problematic.

This study showed that the Niton XL Spectrum Analyzer produced data that exhibit a $\log_{10}$-$\log_{10}$ relationship with the reference data. The analyzer generally exhibited a lower precision compared to the reference methods. The XL Spectrum Analyzer precision RSD was generally between 6 and 14 percent at 5 - 10 times the method detection limit. The analyzer's quantitative results were based on a developer-set calibration using the Compton Ratio method which required the use of well defined site specific calibration standards. Sample homogenization was the single most important factor influencing data comparability. The site and soil texture variables did not show a measurable influence on data comparability.

This demonstration found that the analyzer was generally simple to operate in the field. The operator required no specialized experience or training. Ownership and operation of this analyzer may require specific licensing by state nuclear regulatory agencies. There are special radiation safety training requirements and costs associated with this type of license.

The Niton XL Spectrum Analyzer is an effective tool for field use and can provide rapid, real-time analysis of the metals content of soil samples at hazardous waste sites. The analyzer can quickly identify contaminated areas allowing investigation or remediation decisions to be made more efficiently on-site, and thus reduce the number of samples that need to be submitted for confirmatory analysis.

Table of Contents

Section **Page**

Section **Page**

List of Figures

List of Tables

List of Abbreviations and Acronyms

α	alpha
β	beta
CCB	continuing calibration blank
CCS	calibration check sample
CCV	continuing calibration verification
Cd^{109}	cadmium-109
CI	confidence interval
CLP	Contract Laboratory Program
cm	centimeter
cm^2	centimeter squared
cm^3	cubic centimeter
Co^{57}	cobalt 57
CRM	certified reference material
EPA	Environmental Protection Agency
ERA	Environmental Resource Associates
ESD	Environmental Sciences Division
ETVR	environmental technology verification report
eV	electron volt
FPXRF	field portable X-ray fluorescence
ICAL	initial calibration
ICB	initial calibration blank
ICP-AES	inductively coupled plasma-atomic emission spectroscopy
ICS	interference check standard
ICV	initial calibration verification
IDL	instrument detection limit
IDW	investigation-derived waste
keV	kiloelectron volt
LCS	laboratory control samples
$\log_{10}$	base 10 logarithm
LRL	lower reporting limit
MCA	multichannel analyzer
mCi	millicurie
MDL	method detection limit
mg/kg	milligram per kilogram
mL	milliliter
mm	millimeter
MMTP	Monitoring and Measurement Technologies Program
mrem/hr	millirems per hour

MRI	Midwest Research Institute
NERL-ESD	National Exposure Research Laboratory—Environmental Sciences Division
NIST	National Institute of Standards and Technology
OSW	Office of Solid Waste
PAL	performance acceptance limit
PARCC	precision, accuracy, representativeness, completeness, and comparability
PE	performance evaluation
PI	prediction interval
ppm	part per million
PRC	PRC Environmental Management, Inc.
QA	quality assurance
QAPP	quality assurance project plan
QC	quality control
r	correlation coefficient
r^2	coefficient of determination
RCRA	Resource Conservation and Recovery Act
RPD	relative percent difference
RSD	relative standard deviation
RTC	Resource Technology Corporation
SD	standard deviation
SITE	Superfund Innovative Technology Evaluation
SOP	standard operating procedure
SRM	standard reference material
TC	toxicity characteristic
USGS	United States Geological Survey
XRF	X-ray fluorescence

Acknowledgments

The U.S. Environmental Protection Agency (EPA) wishes to acknowledge the support of all those who helped plan and conduct this demonstration, interpret data, and prepare this report. In particular, for demonstration site access and relevant background information, Tom Aldridge (ASARCO) and Harold Abdo (RV Hopkins); for turnkey implementation of this demonstration, Eric Hess, Patrick Splichal, and Harry Ellis (PRC Environmental Management, Inc.); for editorial and publication support, Suzanne Ladish, Anne Witebsky, Karen Bollinger, and Ed Hubert (PRC Environmental Management, Inc.); for technical and peer review, Paula Hirtz, David Farnam, and Alan Byrnes (PRC Environmental Management, Inc.); for analyzer operation, Nate Meyer (PRC Environmental Management, Inc.); for sample preparation, Scott Schulte, Keith Brown, and Curt Enos (PRC Environmental Management, Inc.); for EPA project management, Stephen Billets, National Exposure Research Laboratory–Environmental Sciences Division; and for peer review, Sam Goforth (independent consultant), John Wallace (Wallace Technologies), Shirley Wasson (National Risk Management Research Laboratory), Brian Schumacher (National Exposure Research Laboratory), and Bill Engelmann (National Exposure Research Laboratory). In addition, we gratefully acknowledge the participation of Oliver Fordham, EPA Office of Solid Waste; Piper Peterson, EPA Region 10; Brian Mitchell, EPA Region 7; and Stephen Shefsky, Niton Corporation.

Section 1
Executive Summary

In April 1995, the U.S. Environmental Protection Agency (EPA) sponsored a demonstration of field portable X-ray fluorescence (FPXRF) analyzers. The primary objectives of this demonstration were to evaluate these analyzers for: (1) their analytical performance relative to standard analytical methods, (2) the influence of sample matrix variations (texture, moisture, heterogeneity, and chemical composition) on performance, (3) the logistical and economic resources needed to operate these technologies in the field, and (4) to test and validate an SW-846 draft method for FPXRF analysis. Secondary objectives for this demonstration were to evaluate FPXRF analyzers for their reliability, ruggedness, cost, range of usefulness, data quality, and ease of operation.

This demonstration was intended to provide users with a reference measure of performance and to act as a guide for the application of this technology. In this demonstration, the reference methods for evaluating the comparability of data were SW-846 Methods 3050A and 6010A, “Acid Digestion of Sediments, Sludges, and Soils” and “Inductively Coupled Plasma-Atomic Emission Spectroscopy (ICP-AES),” respectively.

The EPA requested that PRC Environmental Management, Inc. (PRC) assist in the planning, execution, and reporting of a demonstration of FPXRF analyzers. This demonstration was conducted under the EPA’s Superfund Innovative Technology Evaluation (SITE) Program and managed by the National Exposure Research Laboratory-Environmental Sciences Division (NERL-ESD) under the Monitoring and Measurement Technologies Program (MMTP), Las Vegas, Nevada.

The FPXRF analyzers tested in this demonstration were designed to provide rapid, real-time analysis of metals concentrations in soil samples. This information will allow investigation and remediation decisions to be made on-site more efficiently, and should reduce the number of samples that need to be submitted for confirmatory analysis. Of the seven commercially available analyzers evaluated, one is manufactured by Niton Corporation (the Niton XL Spectrum Analyzer); two are manufactured by Metorex Inc. (the X-MET 920–P Analyzer and the X-MET 920–MP Analyzer); two are manufactured by TN Spectrace (the TN 9000 and the TN Pb Analyzer); one is manufactured by HNU Systems, Inc. (the SEFA-P Analyzer); and one is manufactured by Scitec Corporation (the MAP Spectrum Analyzer). The X-MET 940, a prototype FPXRF analyzer developed by Metorex, was given special consideration and replaced the X-MET 920-P for part of the RV Hopkins sample analyses. This environmental technology verification report (ETVR) presents information relative to the Niton XL Spectrum Analyzer. Separate ETVRs will be published for the other analyzers that were demonstrated.

The target analytes for this demonstration were selected from the Resource Conservation and Recovery Act's (RCRA) Toxicity Characteristic (TC) list, analytes known to have a high aquatic toxicity and likely to produce interferences for the FPXRF analyzers. The primary analytes for these comparisons were arsenic, barium, chromium, copper, lead, and zinc; nickel, iron, cadmium, and antimony were secondary analytes. Because of design considerations, not all analytes were determined by each instrument. For this demonstration, the Niton XL Spectrum Analyzer was configured to report arsenic, chromium, copper, lead, and zinc.

To demonstrate these analyzers, hazardous waste sites in Iowa (the RV Hopkins site) and in the State of Washington (the ASARCO site) were selected. These sites were chosen because they exhibit a wide range of concentrations for most of the target analytes, are located in different climatological regions of the United States, and combined they exhibit three distinct soil textures: sand, loam, and clay.

This demonstration found that the XL Spectrum Analyzer was simple to operate in the field. The developer provided a training course, which encompassed enough FPXRF theory and hands-on analyzer use to allow the operator to manipulate the data collection software and to adjust instrument parameters, such as count times and target analytes. The analyzer did not experience an operational failure resulting in a project down time or data loss during the demonstration. The analyzer was portable, and could operate continuously over a 12-hour work day with appropriate battery changes. The rainy conditions encountered at one of the sites caused no operational problems with the analyzer.

The XL Spectrum Analyzer can determine a large number of analytes including molybdenum, zirconium, strontium, rubidium, lead, arsenic, zinc, copper, nickel, iron, and chromium. For this demonstration, the Niton Analyzer was configured to report the five target analytes noted previously. The analyzer uses a single radioactive source, Cd^{109}, coupled with a silicon-pin diode detector. The type and strength of the source allow this analyzer to produce reliable data at count times as short as 60 live-seconds. The short count times resulted in a sample throughput averaging between 20 and 25 samples per hour.

An evaluation of the results of this demonstration indicates that the analyzer's data and the reference data follow a strong $\log_{10}$-$\log_{10}$ correlation. The XL Spectrum Analyzer produced data meeting definitive level (equivalent to reference data) quality criteria for lead and quantitative screening level (not equivalent to reference data, but correctable with confirmatory analysis) data quality for arsenic, copper, and zinc. The analyzer's performance on chromium could not be evaluated due to the limited precision and detection limit data.

The XL Spectrum Analyzer exhibited a lower precision relative to the reference methods. Field-based method detection limits (MDL) for this analyzer are generally 2 to 3 times higher than the precision-based value. Of the four levels of sample preparation evaluated, the initial sample homogenization had the largest impact on data comparability. Site and soil texture did not appear to affect data comparability.

Based on the performance of the XL Spectrum Analyzer, this demonstration found it to be an effective tool for characterizing the concentration of metals in soil samples. As with all of the FPXRF analyzers, unless a user has regulatory approval, confirmatory (reference) sampling and data correction is recommended when using this technology for site characterization or remediation monitoring.

Section 2
Introduction

This environmental technology verification report (ETVR) presents information from the demonstration of the XL Spectrum Analyzer. This analyzer was developed by the Niton Corporation to perform elemental analyses (metals quantitation) in the field. The analyzer uses a silicon pin-diode detector with a radioactive source (cadmium-109 [Cd^{109}]) to detect the metals in the test sample. The analyzer can be operated in either an *in situ* or intrusive mode. The *in situ* mode, commonly called "point-and-shoot," requires the point of measurement on the soil surface be cleared of loose debris and organic matter, the analyzer's probe is then placed in direct contact with the soil surface, and a measurement is taken. In the intrusive mode of operation, a soil sample is physically collected, dried or sieved, and then placed in a sample cup. The sample cup is placed into an analysis chamber on the probe and a measurement is taken.

This section provides general information about the demonstration such as the purpose, objectives, and design. Section 3 presents and discusses the quality of the data produced by the reference methods against which the analyzer was evaluated. Section 4 discusses the XL Spectrum analyzer, capabilities, reliability, throughput, accuracy, precision, comparability to reference methods, and other evaluation factors. Section 5 discusses the potential applications of the analyzer, presents a method for data correction, and suggests a framework for a standard operating procedure (SOP). Section 6 lists references cited in this ETVR.

Demonstration Background, Purpose, and Objectives

The demonstration was conducted under the Monitoring and Measurement Technologies Program (MMTP), a component of the SITE Program. MMTP is managed by NERL-ESD, Las Vegas, Nevada. The goal of the MMTP is to identify and demonstrate new, innovative, and commercially available technologies that can sample, identify, quantify, or monitor changes in contaminants at hazardous waste sites. This includes those technologies that can be used to determine the physical characteristics of a site more economically, efficiently, and safely than conventional technologies. The SITE Program is administered by the National Risk Management Research Laboratory, Cincinnati, Ohio.

The purpose of this demonstration was to provide the information needed to fairly and thoroughly evaluate the performance of FPXRF analyzers to identify and quantify metals in soils. The primary objectives were to evaluate FPXRF analyzers in the following areas: (1) their accuracy and precision relative to conventional analytical methods; (2) the influence of sample matrix variations (texture, moisture, heterogeneity, and chemical composition) on performance; (3) the logistical and economic resources needed to operate these analyzers; and (4) to test and validate an SW-846 draft method for FPXRF analysis.

Secondary objectives for this demonstration were to evaluate FPXRF analyzers for their reliability, ruggedness, cost, range of usefulness, data quality, and ease of operation. The performances of the FPXRF analyzers were not compared against each other. Instead, the performance of each analyzer was independently and individually compared to that of standard analytical methods commonly used in regulatory enforcement or compliance activities. In addition, each analyzer's performance was assessed relative to measurements of standard reference materials (SRM), performance evaluation (PE) samples, and other quality control (QC) samples.

A special request was made by Mr. Oliver Fordham, the demonstration's technical advisor, EPA Office of Solid Waste (OSW), for Midwest Research Institute (MRI) to analyze some of the soil samples to validate the performance of draft Method 3052 "Microwave Assisted Acid Digestion of Ash and Other Siliceous Wastes." Thirty percent of the soil samples were extracted using draft Method 3052 and then analyzed by Method 6010A. The data generated from the draft Method 3052 and Method 6010A analysis were not used for comparative purposes to the FPXRF data in this demonstration.

Reference Methods

To assess the performance of each analyzer, FPXRF data was compared to reference data. The reference methods used for this assessment were EPA SW-846 Methods 3050A/6010A, which are considered the standards for metals analysis in soil for environmental applications. For purposes of these discussions, the term "reference" was substituted for "confirmatory" since the data was used as a baseline for comparison. In accordance with Federal Acquisition Regulations, MRI was awarded a subcontract to analyze soil samples using the reference methods. The award was made based on MRI's costs, ability to meet the demonstration's quality assurance project plan (QAPP) requirements, and as the only commercial laboratory that could perform all the analyses in the required timeframe.

Method 3050A is the standard acid extraction method for determining metals concentrations in soil samples. It is not a total digestion method, and it may not extract all the metals in a soil sample. Method 6010A is the standard method used to analyze Method 3050A extracts. Both of these methods are described in Section 3.

High quality, well documented reference laboratory results were essential for meeting the objectives of the demonstration. For an accurate assessment, the reference methods had to provide a known level of data quality. For all measurement and monitoring activities conducted by the EPA, the Agency requires that data quality parameters be established based on the end uses of the data. Data quality parameters usually include five indicators often known as the PARCC parameters: precision, accuracy, representativeness, completeness, and comparability. In addition, method detection limits (MDLs) are also used to assess data quality.

Reference methods were evaluated using the PARCC parameters to establish the quality of data generated and to ensure that the comparison of FPXRF analyzers to reference data was acceptable. The following narrative provides definitions of each of the PARCC parameters.

Precision refers to the degree of mutual agreement between replicate measurements and provides an estimate of random error. Precision is often expressed in terms of relative standard deviation (RSD) between replicate samples. The term relative percent difference (RPD) is used to provide this estimate of random error between duplicate samples.

Accuracy refers to the difference between a sample result and the reference or true value. Bias, a measure of the departure from perfect accuracy, can be estimated from the reference or true value. Accuracy and bias for the reference laboratory were assessed by evaluating calibration standard linearity, method blank results and the percent recoveries of matrix spike samples, laboratory control samples (LCS), standard reference materials (SRMs), and PE samples.

Representativeness refers to the degree to which data accurately and precisely measures the conditions and characteristics of the parameter of interest. Representativeness for the reference laboratory was ensured by executing consistent sample collection procedures including sample locations, sampling procedures, storage, packaging, shipping, equipment decontamination, and proper laboratory sample handling procedures. Representativeness was ensured by using the appropriate reference method at its optimum capability to provide results that represented the most accurate and precise measurement it was capable of achieving. The combination of the existing method requirements supplemented by the demonstration QAPP provided the guidance to assure optimum performance of the method. Representativeness was assessed by evaluating calibration standards, method blank samples, duplicate samples, and PE samples.

Completeness refers to the amount of data collected from a measurement process compared to the amount that was expected to be obtained. For the reference data, completeness referred to the proportion of valid, acceptable data generated.

Comparability refers to the confidence with which one data set can be compared to another. Data generated from the reference methods should provide comparable data to any other laboratory performing analysis of the same samples with the same analytical methods. Comparability for the reference methods was achieved through the use of standard operating procedures (SOPs), EPA-published guidance, and the demonstration QAPP. QC samples that were used to evaluate comparability include: calibration standards, method blank samples, matrix spike samples, replicate samples, LCSs, SRMs, and PE samples.

Site Selection

PRC conducted a search for suitable demonstration sites between September and November 1994. The following criteria were used to select appropriate sites:

- The site owner had to agree to allow access for the demonstration.
- The site had to have soil contaminated with some or all of the target heavy metals. (Slag, ash, and other deposits of mineralized metals would not be assessed during the demonstration.)
- The site had to be accessible to two-wheel drive vehicles.
- The site had to exhibit one or more of the following soil textures: sand, clay, or loam.
- The site had to exhibit surface soil contamination.
- The sites had to be situated in different climatological environments.

PRC contacted NERL-ESD, regional EPA offices, state environmental agencies, metals fabrication, and smelting contacts to create an initial list of potential demonstration sites. PRC received considerable assistance from the EPA RCRA and Superfund branches in Regions 4, 6, 7, 8, 9, and 10. PRC also contacted the Montana Department of Health and Environment, the Nevada Bureau of Mines and Geology, the Oklahoma Department of Environmental Quality, the Arizona Department of Environmental Quality, the Missouri Department of Natural Resources, the Arizona Bureau of Geology, and the New Mexico Bureau of Mines and Mineral Resources. PRC surveyed its offices in Kansas City, Kansas; Atlanta, Georgia; Denver, Colorado; Dallas, Texas; Albuquerque, New Mexico; Helena, Montana; Chicago, Illinois; Seattle, Washington; and San Francisco, California, for information regarding potential sites. These PRC offices have existing RCRA, Superfund, or Navy environmental contracts that allow access to regional, state, and federal site information. PRC also used the Record of Decision Scan database (Morgan and others 1993) to search for appropriate sites.

PRC screened 46 potential sites based on the site-selection criteria with the assistance of the various contacts listed above. Based on this screening effort, PRC and EPA determined that the RV Hopkins and ASARCO sites met most of the site-selection criteria, and therefore, would be acceptable for the demonstration.

The ASARCO site consists of 67 acres of land adjacent to Commencement Bay. The site is marked by steep slopes leading into the bay, a slag fill that was used to extend the original shoreline, a cooling water pond, and the various buildings associated with the smelting process. Partial facility demolition was conducted in 1987. Most of the buildings were demolished between 1993 and 1994. The only buildings remaining are the Fine Ore Building, the Administrative Building, and a Maintenance Garage.

Past soil sampling results have targeted four general areas of the site as acceptable candidates for the demonstration: the plant administration area, the former cooling pond, the 1987 demolition area, and certain off-site residential areas adjacent to the smelter stack. Previous sampling has shown surficial soils to be more contaminated than subsurface soils. Arsenic, copper, and lead are the predominant contaminants in the local soils. The highest arsenic concentrations were found in the soils around the former arsenic kitchen, along with cadmium and mercury. The soils around the former cooling pond contained the highest copper concentrations and high levels of silver, selenium, barium, and chromium. Lead concentrations are highest northeast of the arsenic plant.

Much of the smelter site is covered with artificial fill material of varying thickness and composition. Two general types of fill are found on-site: granular and slag. The composition of the granular fill material ranges from sand to silt with demolition debris and slag debris mixed throughout. The slag fill is a solid, fractured media restricted to the plant site. The surface soil in the plant administration area has a layer of slag particles on top, ranging from 1 to 3 inches thick. Surficial material in the parking lot area and southwest of the stack is mostly of glacial origin and composed of various mixtures of sand, gravel, and cobbles. The soils around the former cooling pond are fine-grained lacustrine silts and clays. Alluvium upgradient of the former cooling pond has been almost entirely covered with granular fill material. Generally, soils in the arsenic kitchen and stack hill areas are sand mixed with gravel or sandy clay mixed with cobbles. No slag was analyzed as part of this demonstration.

The RV Hopkins site is located in the west end of Davenport, Iowa. The facility occupies approximately 6.7 acres in a heavy industrial/commercial zoned area. Industrial activities in the area of the RV Hopkins property included the manufacture of railroad locomotive engines during the mid-1800's. The RV Hopkins property was a rock quarry during the late 1800's. Aerial surveys beginning in 1929 show that the rock quarry occupied the majority of the site initially, gradually decreasing until it was completely filled by 1982. It was reported that the site was used to dispose of demolition debris, automotive, and scrap metal. The site also has been used by a company that recycled lead acid batteries.

RV Hopkins began operating as a drum reconditioner in 1951 across the street from its current location. In 1964, the site owner reportedly covered the former quarry area of the site with foundry sand. No foundry sand was analyzed as part of this demonstration. RV Hopkins receives between 400 and 600 drums per day for reconditioning, accepting only drums that meet the definition of “empty” according to 40 Code of Federal Regulations 261.7. Most of the drums received at the facility come from the paint, oil, and chemical industries. The surrounding area is reported to be underlain by Devonian-aged Wapsipinicon Limestone, and gray-green shale, lime mud, and sand stringers dating back to the Pennsylvanian age.

The RV Hopkins property is composed of five buildings: the office and warehouse, a warehouse used to store drums of hazardous waste and a waste pile, a manufacturing building, a drum reclamation furnace, and a cutting shed. The office and the warehouse are located on the southwest corner of the site. Areas investigated on each site include the furnace area, the old and new baghouses, the former drum storage area on the north end of the facility, the former landfill, and a drainage ditch. Major contaminants include barium, lead, chromium, and zinc, as well as lesser concentrations of other metals, such as copper and nickel, pesticides, and volatile organic compounds.

Based on historical data, the most concentrated contaminants in the furnace area are chromium, lead, and zinc. The highest concentrations of these elements are at the furnace entrance, as opposed to the furnace exit. The concentrations of lead are higher in the old baghouse than in the new, while the new baghouse exhibits a higher concentration of chromium, as well as high iron, lead, and barium concentrations. The former landfill has concentrations of barium, chromium, lead, nickel, and zinc greater than 1,000 mg/kg. Lead is the most prevalent contaminant in the former drum storage area with lesser concentrations of barium, chromium, and zinc.

Predemonstration Sampling

Predemonstration sampling was conducted at both sites between December 5 and 14, 1994. These sampling events had the following objectives:

- To provide data on, or verify, the extent of surface contamination at each site and to locate optimum sampling areas for the demonstration.
- To allow the developers to analyze samples from the demonstration sites in advance of the demonstration, and if necessary, refine and recalibrate their technologies and revise their operating instructions.
- To evaluate samples for the presence of any unanticipated matrix effects or interferences that might occur during the demonstration.
- To check the quality assurance (QA) and QC procedures of the reference laboratory.

One hundred soil samples were analyzed on each site by the FPXRF analyzers during the predemonstration sampling activities. The samples represented a wide range in the concentration of metals and soil textures. Thirty-nine samples were submitted for reference method analysis using EPA SW-846 Methods 3050A/6010A. Twenty-nine of these samples were split and sent to the developers. Nine field duplicates were collected and submitted for reference method analysis to assess proposed sample homogenization procedures. One purchased PE sample also was submitted to the reference laboratory to provide an initial check of its accuracy.

Additionally, three samples representing low, medium, and high concentrations were collected at each site. These samples were dried, ground, and then analyzed by six independent laboratories before the demonstration began to create site-specific PE samples. These samples were analyzed with laboratory-grade X-ray fluorescence (XRF) analyzers.

Experimental Design

The experimental design of this demonstration was developed to meet the primary and secondary objectives stated above, and was approved by all demonstration participants prior to the start of the demonstration. The design is detailed in the demonstration plan (PRC 1995) and is summarized below.

Approximately 100 soil samples were collected from each of three target soil textures: clay, loam, and sand. This variety of soil textures allowed the examination of the effect of soil texture on data comparability. Splits of these samples were analyzed by all FPXRFs for all sample preparation steps and by the reference methods.

The XL Spectrum Analyzer can be operated in either an *in situ* or intrusive mode. These two modes of analysis involve different measurement and preparation procedures. These procedures allowed for an evaluation of the effects of sample preparation on FPXRF comparability to reference data. For *in situ* analysis, an area 4 inches by 4 inches square was cleared of all vegetation, debris, and gravel larger than 2 millimeters (mm) in diameter. Each analyzer took one *in situ* measurement in each sample area. The data represented FPXRF *in situ* measurements for unprepared soils (*in situ*-unprepared). Replicate measurements were taken at 4 percent of these locations to assess analyzer precision. Figure 2-1 depicts the sample analysis chain for both *in situ* and intrusive analyses.

After the *in situ*-unprepared analysis was complete at a given location, the soil within the 4-inch by 4-inch square was removed to a depth of 1 inch and placed in a plastic bag. This produced a soil sample of approximately 375 grams or 250 cubic centimeters (cm^3). Sample homogenization was monitored by adding 1 to 2 grams of sodium fluorescein salt (which fluoresces when exposed to ultraviolet light) to the sample bag. During the predemonstration, it was determined that sodium fluorescein did not affect the FPXRF or reference method analysis. Sample homogenization took place by kneading the sample and sodium fluorescein salt in a plastic bag for 2 minutes. After this period, the sample preparation technician examined the sample under ultraviolet light to assess the distribution of sodium fluorescein. If the sodium fluorescein salt was not evenly distributed, the homogenization and checking process were repeated until the sodium fluorescein was evenly distributed throughout the sample. This monitoring process assumed that thorough distribution of sodium fluorescein was indicative of good sample homogenization. The effectiveness of this process is discussed later in this section.

The homogenized sample was then spread out inside a 1-inch-deep petri dish. Each FPXRF analyzer took one measurement from this homogenized material. This represented the homogenized sample analysis for the *in situ* analyzers (*in situ*-prepared). This approximated the common practice of sample homogenization in a plastic bag and subsequent sample measurement through the bag. Replicate measurements were also collected from 4 percent of these samples to assess analyzer precision. These replicate measurements were made on the same soils as the unprepared precision measurements.

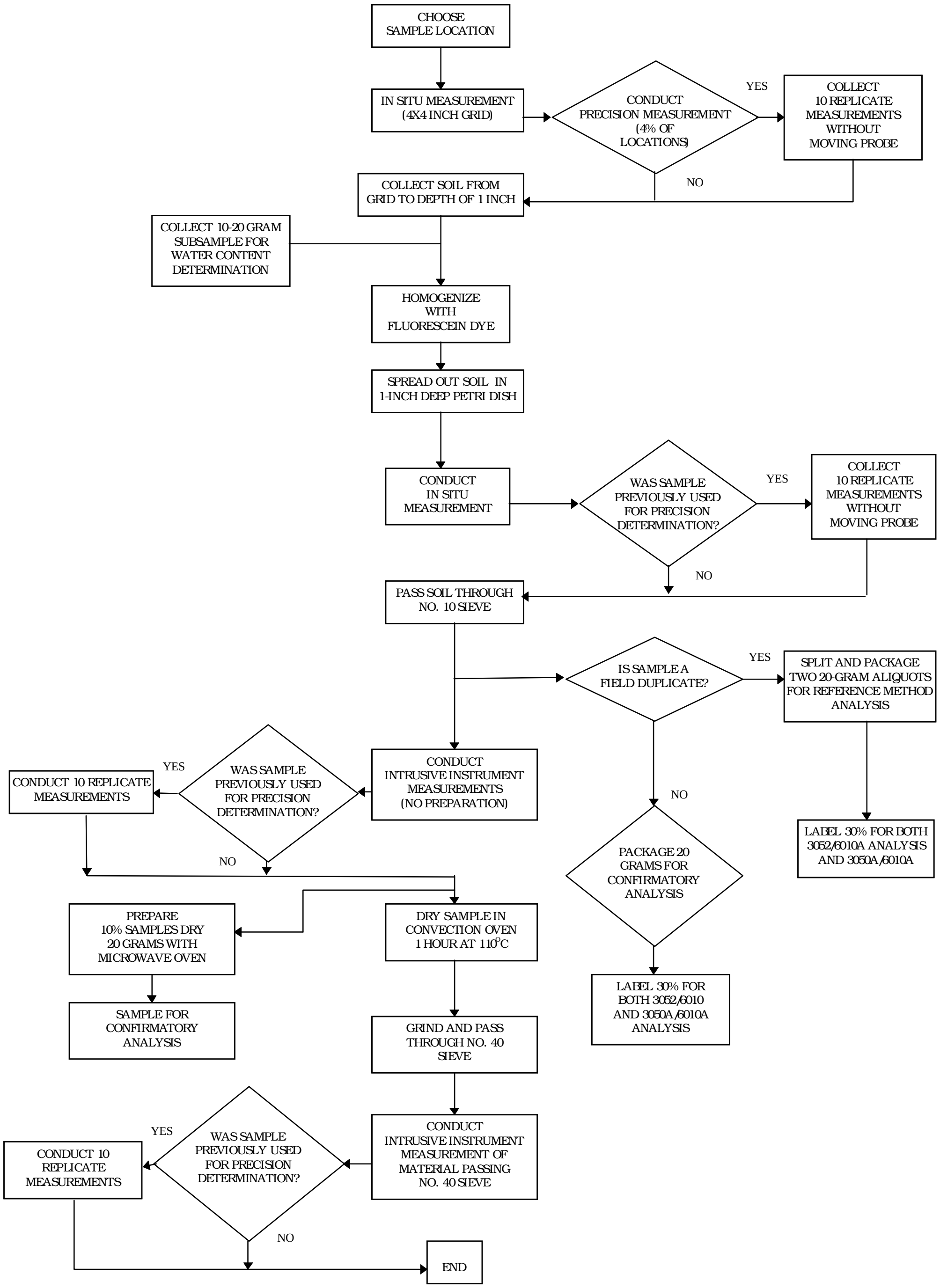


Figure 2-1. Sample Preparation and Analysis: This flowchart depicts the handling procedures for each sample taken during the demonstration.

Following the *in situ*-prepared procedure, the sample material was passed through a No. 10 mesh sieve (2-mm openings) and approximately 10 grams of this material was placed in a sample cup for analysis in an intrusive mode. The same sample cup was used for each FPXRF analyzer. Replicate measurements were collected from 4 percent of these samples to assess analyzer precision. These replicate measurements were made on the same soils as the *in situ*-prepared precision measurements. These data represented FPXRF intrusive mode measurements on soils with no sample preparation (intrusive-unprepared). Sample material from this preparation step was collected and submitted to the reference laboratory for reference method analysis.

Following the intrusive-unprepared analysis, a portion of that soil sample was dried in a convection oven at 110 °C for 1 hour and ground with a mortar and pestle until it passed through a No. 40 stainless-steel sieve (0.425-mm openings). The sample was then analyzed in an intrusive mode. Four percent of these samples underwent replicate measurements to evaluate analyzer precision. These replicate measurements were performed on the same soils as in the intrusive-unprepared precision measurements. This data represented FPXRF intrusive measurements on prepared soils (intrusive-prepared).

Qualitative Factors

There are a number of factors important to data collection that are difficult to quantify and must be evaluated qualitatively. These are considered qualitative factors. One such factor was the amount of training required to operate a given FPXRF analyzer. To assess this factor, PRC operators were trained by the developers on how to operate their respective FPXRF analyzers. All operators met or exceeded the developers' minimum requirements for education and previous experience. Demonstration procedures were designed to simulate routine field conditions as closely as possible. Based on this training and field experience, the operators prepared a subjective evaluation assessing the training and technology operation during the demonstration (Section 4).

Many analytical methods exhibit "operator effects," in which individual differences in sample preparation or operator technique result in a significant effect on the numerical results. To reduce the possible influence of operator effects, a single operator was used to operate each FPXRF analyzer. While this reduced some potential error from the evaluation, it did not allow the analyzers to be evaluated for their susceptibility to operator-induced error. A single operator was used to analyze all of the samples at both sites during this demonstration. Sample preparation variation effects were minimized in the field by using the same personnel to prepare samples. To eliminate the influence of operator effects on the reference method analysis, only one reference laboratory was used to analyze the samples. Based on this design, there could be no quantitative estimate of the "operator effect."

Quantitative Factors

Many factors in this demonstration could be quantified by various means. Examples of quantitative factors evaluated during this demonstration include analyzer performance near regulatory action levels, the effects of sample preparation, effects of microwave sample drying, count times, health and safety considerations, costs, and interferences.

The data developed by the FPXRF analyzers were compared to reference data for the following primary analytes: arsenic, barium, chromium, copper, lead, and zinc; and for the following secondary analytes: nickel, iron, cadmium, and antimony. The specific target analytes determined by the XL Spectrum Analyzer were arsenic, chromium, copper, lead, and zinc.

Evaluations of analyzer data comparability involved examining the effects of each site, soil texture, and sample preparation technique (Table 2-1). Two sites were sampled for this demonstration and therefore, two site variables were examined (RV Hopkins and ASARCO sites). These sites produced samples from three distinct soil textures, and therefore, three soil variables were examined (clays, sands, and loams). Four sample preparation steps were used: (1) *in situ*-unprepared, (2) *in situ*-prepared, (3) intrusive-unprepared, and (4) intrusive-prepared. These variables were nested as follows: each site was divided into RV Hopkins and ASARCO data sets; the RV Hopkins data represented the clay soil texture, and the ASARCO data was divided into sand and loam soil textures; each soil texture was subdivided by the four soil preparations. These variables allowed the examination of particle size and sample homogenization. These effects were believed to have the greatest impact on data comparability.

Table 2-1. Performance and Comparability Variables Evaluated

Variables		
Site Name (315)	**Soil Texture (315)**	**Preparation Step [1,260]**
ASARCO (215)	Sand (100)	in situ-unprepared [100] in situ-prepared [100] intrusive-unprepared [100] intrusive-prepared [100]
	Loam (115)	in situ-unprepared [115] in situ-prepared [115] intrusive-unprepared [115] intrusive-prepared [115]
RV Hopkins (100)	Clay (100)	in situ-unprepared [100] in situ-prepared [100] intrusive-unprepared [100] intrusive-prepared [100]

Notes: ( ) Total number of sample points.
[ ] Total number of measurements taken.

Of greatest interest to users is analyzer performance near action levels. For this reason, samples were approximately distributed as follows: 25 percent in the 0 - 100 mg/kg range, 50 percent in the 100 - 1,000 mg/kg range, and 25 percent in the greater than 1,000 mg/kg range. The lower range tested analyzer performance near MDLs; the middle range tested analyzer performance in the range of many action levels for inorganic contaminants; and the higher range tested analyzer performance on grossly contaminated soils. All samples collected for the demonstration were split between the FPXRF analyzers and reference laboratory for analysis. Metal concentrations measured using the reference methods were considered to represent the "true" concentrations in each sample. Where duplicate samples existed, concentrations for the duplicates were averaged and the average concentration was considered to represent the true value for the sample pair. This procedure was specified in the demonstration plan. If one or both samples in a duplicate pair exhibited a nondetect for a particular target analyte, that pair of data was not used in the statistical evaluation of that analyte. The reference methods reported measurable concentrations of target analytes in all of the samples analyzed.

In addition to the quantitative factors discussed above, the common FPXRF sample preparation technique of microwave was evaluated. Sample temperatures during this procedure can be high enough to melt some mineral fractions in the sample or combust organic matter. Several metals that present environmental hazards can volatilize at elevated temperatures. Arsenic sublimes at 188 °C, within the potential temperature range achieved during microwave drying. To assess this potential effect, 10 percent of the homogenized, crushed, oven-dried, and sieved samples were split and heated in a microwave oven on high for 3 minutes. This time was chosen to approximate common microwave drying times used in the field. These split samples were then submitted for reference analysis. The reference data for these samples were compared to the corresponding reference data produced from the convection oven-dried sample. These data showed the effects of the microwave drying variable on analyte concentration. This was a minor variable and was only evaluated for the reference laboratory in an attempt to identify any potential effect on data comparability.

Another quantitative variable evaluated was the count time used to acquire data. During the formal sample quantitation and precision measurement phase of the demonstration, the count times were set by the developers and remained constant throughout the demonstration. Count times can be tailored to produce the best results for specific target analytes. The developers, however, selected count times that produced the best compromise of results for the entire suite of target analytes. To allow a preliminary assessment of the effect of count times, select soil samples were analyzed in replicate using count times longer and shorter than those set by the developers. This allowed the evaluation of the effects of count times on analyzer performance. Since sample throughput can be affected by adjusting count times, operators used only the developer-specified count times throughout the demonstration.

An important health and safety issue during the demonstration was the effectiveness of radioactivity shielding of each FPXRF analyzer. Occasional radiation readings were quantitatively made with a gamma ray detector near each analyzer to assess the potential for exposure to radiation.

A compilation of the costs associated with the use of each FPXRF analyzer was another important evaluation factor. Cost includes analyzer purchase or rental, expendable supplies, such as liquid nitrogen and sample cups, and nonexpendable costs, such as labor, licensing agreements for the radioactive sources, operator training costs, and disposal of investigation-derived waste (IDW). This information is provided to assist the user in preparing a project cost analysis associated with the use of this instrument.

Factors that could have affected the quantitative evaluations included interference effects and matrix effects. Some of these effects and the procedures used to evaluate their influence during this demonstration are summarized below:

- Heterogeneity: For *in situ*-unprepared measurements, heterogeneity was partially controlled by restricting measurements within a 4-by-4-inch area. For measurements after the initial point-and-shoot preparation, heterogeneity was minimized by sample homogenization. This effect was evaluated through the sample preparation data.

- Particle Size: The effect of particle size was evaluated using the two intrusive sample preparation procedures. Theoretically, precision and accuracy should increase as particle size decreases and becomes uniform.

- Moisture Content: It has been suggested that major shifts in sample moisture content can affect a sample's relative fluorescence. This effect could not be evaluated as thoroughly as planned because of the small difference in sample moisture content observed at the two sites.

- Overlapping Spectra of Elements: Interferences result from overlapping spectra of metals that emit X-rays with similar energy levels. The reference method analysis provided data on the concentration of potential interferants in each sample.

Evaluation of Analyzer Performance

Metals concentrations measured by each analyzer were compared to the corresponding reference laboratory data, and to other QA/QC sample results. These comparisons were conducted independently for each target analyte. These measurements were used to determine an analyzer's accuracy, data quality level, method precision, and comparability to reference methods. PE and SRM samples were used to assess analyzer accuracy. Relative standard deviations (RSD) on replicate measurements were used to determine analyzer precision. These data were also used to help determine the data quality of each FPXRF analyzer's output. The data comparability and quality determination was primarily based on a comparison of the analyzer's data and the reference data. Linear regression and a matched pairs t-test were the statistical tools used to assess comparability and data quality.

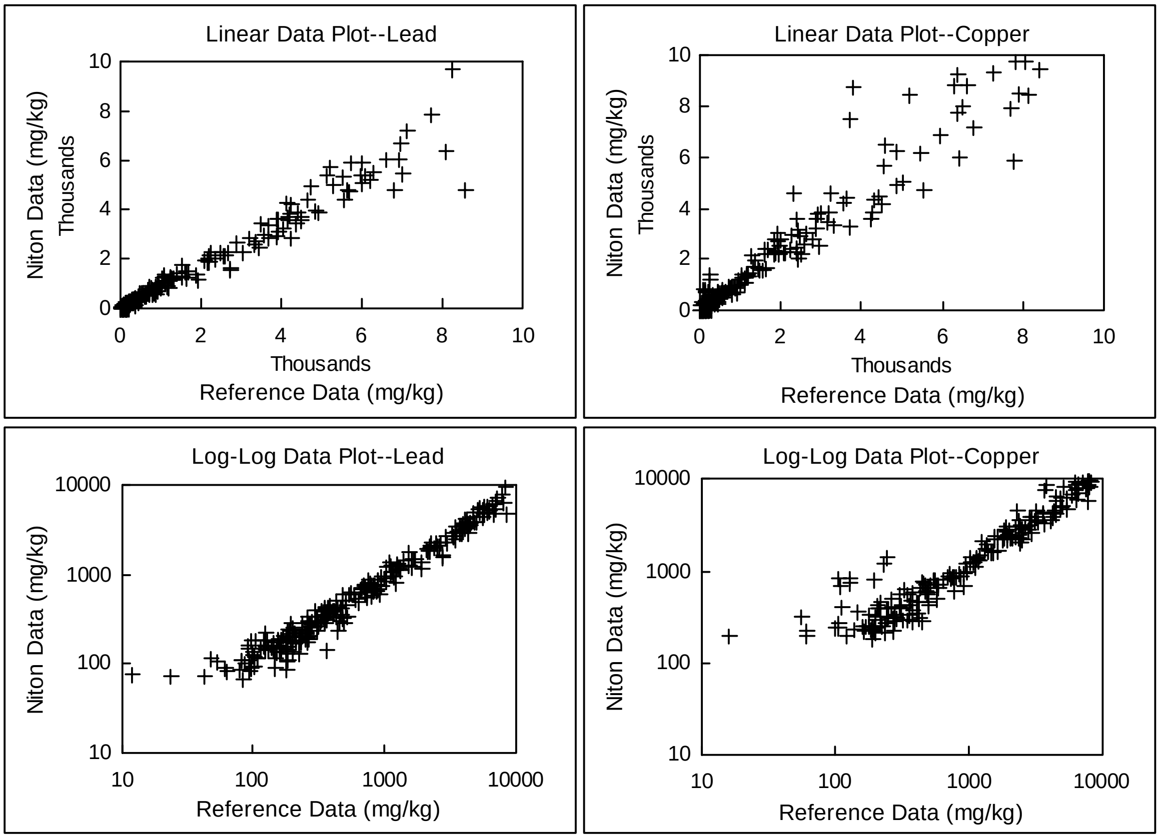
A principal goal of this demonstration was the comparison of FPXRF data and the reference laboratory data. EPA SW-846 Methods 3050A/6010A were selected as the reference methods because they represent the regulatory standard against which FPXRF is generally compared. In comparing the FPXRF data and reference data, it is important to recognize that, while similar, the process by which the data are obtained is not identical. While there is significant overlap in the nature of the analysis, there are also major differences. These differences, or "perspectives," allow the user to characterize the same sample in slightly different ways. Both have a role in site characterization and remediation. It is important to consider these differences and the measurement error intrinsic to each method when comparing the FPXRF method against a reference analytical method.

The reference laboratory methods involve wet chemical analysis and partial acid digestion of approximately 1 to 2 grams of sample (approximately 0.25 cubic centimeters (cm^3) depending on sample bulk density). The digestion process extracts the most acid-soluble portion of the sample. Since the digestion is not complete, the less acid-soluble components are not digested and are not included in the analysis. These components may include the coarser-grained quartz, feldspar, lithic components, and certain metal complexes. In contrast, FPXRF analyzers generally produce X-ray excitation in an area of approximately 3 cm^2 to a depth of approximately 2.5 centimeters (cm). This equates to a sample volume of approximately 7.5 cm^3. X-rays returning to the detector are derived from <u>all</u> matrix material including the larger-grained quartz, feldspar, lithic minerals, metal complexes, and organics. Because the FPXRF method analyzes all material, it represents a total element analysis in contrast to the reference methods, which may only represent a select or partial analysis. This difference can result in FPXRF concentrations that are higher than the corresponding reference data when metals are contained within nonacid soluble complex. It is important to note that if metals are contained in nonacid soluble complexes, a difference between the FPXRF analyzers and the reference methods is not necessarily due to error in the FPXRF result but rather to differences in the sample preparation procedures.

The comparison of FPXRF data and the reference data employs a linear regression as the primary statistical tool. Linear regression analysis intrinsically contains assumptions and conditions that must be valid for the data set. Three important assumptions involve: (1) the linearity of the relationship, (2) the confidence interval and constant error variance, and (3) an insignificant measurement error for the independent variable (reference data).

The first assumption requires that the independent variable (reference data) and the dependent variable (FPXRF data) are linearly related and are not related by some curvilinear or more complex relationship. This linearity condition applies to either the raw data or mathematical transformations of the raw data. Figure 2-2 illustrates that FPXRF data and reference data are, in fact, related linearly and that this assumption is correct.

The second assumption requires that the error be normally distributed, the sum to equal zero, be independent, and exhibit a constant error variance for the data set. Figure 2-2 illustrates that for raw data, this assumption is not correct (at higher concentrations the scatter around the regression line increases), but that for the logarithmic transformation (shown as a log-log plot) of the data, this assumption is valid (the scatter around the regression line is relatively uniform over the entire concentration range). The change in error distribution (scatter) evident in the untransformed data results in the disproportionate influence of large data values compared with small data values on the regression analysis.


Linear Data Plot--Lead

Linear Data Plot--Copper

Log-Log Data Plot--Lead

Log-Log Data Plot--Copper

Figure 2-2. Linear and Log-log Data Plots: These graphs illustrate the linear relationship between the FPXRF data and the reference data. The linear data plots illustrate the concentration dependence of this relationship with increased scatter at higher concentrations. The log-log plots eliminate this concentration effect. Scatter is relatively constant over the entire plot.

The use of least squares linear regression has certain limitations. Least squares regression provides a linear equation, which minimizes the squares of the differences between the dependent variable and the regression line. For data sets produced in this demonstration, the variance was proportional to the magnitude of the measurements. That is, a measurement of 100 parts per million (ppm) may exhibit a 10 percent variance of 10 ppm, while a 1,000 ppm measurement exhibits a 10 percent variance of 100 ppm. For data sets with a large range in values, the largest measurements in a data set exert disproportionate influence on the regression analysis because the least squares regression must account for the variance associated with the higher valued measurements. This can result in an equation that has minimized error for high values, but almost neglects error for low values because their influence in minimizing dependent variable error is small or negligible. In some cases, the resulting equations, biased by high-value data, may lead to inappropriate conclusions concerning data quality. The range of the data examined for the analyzers spanned between 1 and 5 orders of magnitude (e.g., 10 - 100,000 ppm) for the target analytes. This wide range in values and the associated wide range in variance (influenced by concentration) created the potential for this problem to occur in the demonstration data set. To provide a correlation that was equally influenced by both high and low values, logarithms ($\log_{10}$) of the dependent and independent variables were used, thus, scaling the concentration measurements and providing equal weight in the least squares regression analysis to both small and large values (Figure 2-2). All statistical evaluations were carried out on $\log_{10}$ transformed data.

The third assumption, requiring an insignificant measurement error in the reference data, was not true for all analytes. The consequences of measurement error varied depending on whether the error is caused by the reference methods or the FPXRF method. If the error is random or if the error for the reference methods is small compared to the total regression error, then conventional regression analysis can be performed and the error becomes a part of the random error term of the regression model. This error (based on the $\log_{10}$ transformed data) is shown in the regression summary tables in Section 4 as the "standard error." In this case, deviations from perfect comparability can be tied to an analyzer's performance. If the error for the reference methods is large compared to the total error for the correlation of the FPXRF and the reference data, then deviations from perfect comparability might be due in part to measurement error in the reference methods.

It is a reasonable assumption that any measurement errors in either the reference or FPXRF methods are independent of each other. This assumption applies to either the raw data or the $\log_{10}$ transformed data. Given this assumption, the total regression error is approximately the sum of the measurement error associated with the reference methods and the measurement error associated with the FPXRF method. The reference methods' precision is a measure of independent variable error, and the mean square error expressed in the regression analysis is a relative measure of the total regression error that was determined during the regression analysis. Precision data for the reference methods, obtained from RPD analyses on the duplicate samples from each site, for each analyte, indicated the error for the reference methods was less than 10 percent of the total regression error for the target analytes. Subsequently, 90 percent of the total measurement error can be attributed to measurement error associated with the analyzers. Based on this analysis, the reference data did allow unambiguous resolution of data quality determination.

The comparison of the reference data to the FPXRF data is referred to as intermethod comparison. All reference and QA/QC data were generated using an EPA-approved definitive level analytical method. If the data obtained by an analyzer were statistically similar to the reference methods, the analyzer was considered capable of producing definitive level data. As the statistical significance of the comparability decreased, an analyzer was considered to produce data of a correspondingly lower quality. Table 2-2 defines the criteria that determined the analyzer's level of data quality (EPA 1993).

Results from this demonstration were used to assign analyzer data into one of three data quality levels as follows: (1) definitive, (2) quantitative screening, and (3) qualitative screening. The first two data quality levels are defined in EPA guidance (1993). The qualitative screening level criteria were defined in the demonstration plan (PRC 1995) to further differentiate screening level data.

Definitive level data are considered the highest level of quality. These data are usually generated by using rigorous analytical methods, such as approved EPA or ASTM methods. The data is analyte-specific with confirmation of analyte identity and concentration. In addition, either analytical or total measurement error must be determined. Definitive data may be generated in the field, as long as the project QA/QC requirements are satisfied.

Table 2-2. Criteria for Characterizing Data Quality

Data Quality Level	Statistical Parameter[a,b]
Definitive Level	r^2 = 0.85 to 1.0. The precision (RSD) must be less than or equal to 10 percent and the inferential statistics must indicate that the two data sets are statistically similar.
Quantitative Screening Level	r^2 = 0.70 to 1.0. The precision (RSD) must be less than 20 percent, but the inferential statistics indicate that the data sets are statistically different.
Qualitative Screening	r^2 = less than 0.70. The precision (RSD) is greater than 20 percent. The data must have less than a 10 percent false negative rate.

Notes:

[a] The statistical tests and parameters are discussed later in the "Intermethod Comparison" subsection in Section 4.

[b] The regression parameters apply to either raw or $\log_{10}$ transformed data sets. The precision criteria apply to only the raw data.

r^2 Coefficient of determination.

RSD Relative standard deviation.

Quantitative screening data provide confirmed analyte identification and quantification, although the quantification may be relatively imprecise. It is commonly recommended that at least 10 percent of screening level data be confirmed using analytical methods and QA/QC procedures and criteria associated with definitive data. The quality of unconfirmed screening data cannot be determined.

Qualitative screening level data indicates the presence or absence of contaminants in a sample matrix, but does not provide reliable concentration estimates. The data may be compound-specific or specific to classes of contaminants. Generally, confirmatory sampling is not required if an analyzer's operation is verified with one or more check samples.

At the time of this demonstration, an approved EPA method for FPXRF did not exist. As part of this demonstration, PRC prepared draft Method 6200 "Field Portable X-Ray Fluorescence Spectrometry for the Determination of Elemental Concentrations in Soil and Sediment." The draft method has been submitted for inclusion in Update 4 of SW-846 scheduled for approval in 1997. For purposes of this demonstration, the absence of an EPA-approved final method did not preclude the analyzers' data from being considered definitive. The main criterion for data quality level determination was based on the comparability of each analyzer's data to the data produced by the reference methods, as well as analyzer-specific criteria such as precision.

The comparability data set for the XL Spectrum Analyzer consisted of 1,260 matched pairs produced from a total of 315 soil samples. These samples were analyzed by the reference method, and by the XL Spectrum Analyzer four times, using each of the four sample preparation steps. This data set was analyzed as a whole and then subdivided and analyzed with respect to each of the variables listed in Table 2-1. This nesting of variables allowed the independent assessment of the potential influence of each variable on comparability.

Seventy of the 315 samples submitted to the reference laboratory were split and reported as field duplicates to assess the sample homogenization process. Thirty-three of the 315 samples were also split and microwave-dried; then submitted for reference method analysis to assess the effect of microwave drying. Of the 315 samples submitted for reference method analysis, 215 were collected from the ASARCO site and 100 were collected from the RV Hopkins site. Approximately twice as many samples were collected at the ASARCO site because two of the target soil textures (sands and loams) were found there. Only one target soil texture (clay) was found at the RV Hopkins site.

Evaluation of the influence of the site and soil variables was limited to the examination of the lead and zinc data. These were the only primary analytes that exhibited a wide distribution of concentrations across all sites and soil textures. The effects of sample preparation were evaluated for all target analytes. If the evaluation of the influence of a given variable did not result in a better correlation, as exhibited by a higher coefficient of determination (r^2) and smaller standard error of the estimate (using $\log_{10}$ transformed data), then the influence was considered to be insignificant. However, if the correlation worsened, the cause was examined and explained. If the correlation improved, resulting in a higher r^2 value and reduced standard error of the estimate, then the impact of the variable was considered significant. For example, if the r^2 and standard error of the estimate for a given target analyte improved when the data set was divided into the four sample preparation steps, the sample preparation variable was determined to be significant. Once this was determined, the variables of site and soil texture were evaluated for each of the four sample preparations steps. If the site or soil texture variable improved the regression parameters for a given soil preparation, then that variable was also considered significant.

After the significant variables were identified, the impact of analyte concentration was examined. This was accomplished by dividing each variable's $\log_{10}$ transformed data set into three concentration ranges: 0 - 100 mg/kg; 100 - 1,000 mg/kg; and greater than 1,000 mg/kg. Then, linear regression analysis was conducted on the three data sets. If this did not result in improved r^2 values and reduced standard errors of the estimate, the relationship between the analyzer's $\log_{10}$ transformed data and the $\log_{10}$ transformed reference data was considered linear over the entire range of concentrations encountered during the demonstration. This would mean that there was no concentration effect.

Numerous statistical tests have been designed to evaluate the significance of differences between two populations. In comparing the performance of the FPXRF analyzers against the reference methods, the linear regression comparison and the paired t-test were considered the optimal statistical tests. The paired t-test provides a classic test for comparing two populations, but is limited to analysis of the average or mean difference between those populations. Linear regression analysis provides information not only about how the two populations compare on average, but also about how they compare over ranges of values. This statistical analysis provides information about the structure of the relationship; that is, whether the methods differ at high or low concentrations or both. It also indicates whether the FPXRF data is biased or shifted relative to the reference data.

Linear regression provides an equation that represents a line (Equation 2-1). Five linear regression parameters were considered when assessing the level of data quality produced by the FPXRF analyzers. This assessment was made on the $\log_{10}$ transformed data sets. The five parameters were the y-intercept, the slope of the regression line, standard error of the estimate, the correlation coefficient (r), and r^2. In linear regression analysis, the r provides a measure of the degree or strength of the correlation between the dependent variable ($\log_{10}$ transformed FPXRF data), and the independent variable ($\log_{10}$ transformed reference data). The r^2 provides a measure of the fraction of total variation which is accounted for by the regression relation (Havlick and Crain 1988). That is, it is a measure of the scatter about a regression line and, thus, is a measure of the strength of the linear association.

$$Y = mX + b \tag{2-1}$$

where

> b is the y-intercept of the regression line, m is the slope of the regression line, and Y and X are the $\log_{10}$ transformed dependent and independent variables, respectively.

Values for r vary from 1 to -1, with either extreme indicating a perfect positive or negative correlation between the independent and dependent variables. A positive correlation coefficient indicates that as the independent variable increases, the dependent variable also increases. A negative correlation coefficient indicates an inverse relationship, as the independent variable increases the dependent variable decreases. An r^2 of 1.0 indicates that the linear equation explains all the variation between the FPXRF and reference data. As the r^2 departs from 1.0 or -1 and approaches zero, there is more unexplained variation, due to such influences as lack of perfect association with the dependent variable ($\log_{10}$ transformed FPXRF data), or the influence of other independent variables.

If the regression correlation exhibited an r^2 between 0.85 and 1.0, the FPXRF data was considered to have met the first requirement for definitive level data classification (Table 2-2). The second criteria, precision RSD was then examined and required to be equal or less than 10 percent to retain the definitive data quality level assignment. If either or both of these criteria are not satisfied, certain inferential parameters were then evaluated. As a starting point, the regression line's y-intercept and slope were examined. A slope of 1.0 and a y-intercept of 0.0 would mean that the results of the FPXRF analyzer matched those of the reference laboratory ($\log_{10}$ FPXRF=$\log_{10}$ reference). Theoretically, the more the slope and y-intercept differ from the values of 1.0 and 0.0, respectively, the less accurate the FPXRF analyzer. However, a slope or y-intercept can differ slightly from these values without that difference being statistically significant. To determine whether such differences were statistically significant, the Z test statistics for parallelism and for a common intercept was used at the 95 percent confidence level for the comparison (Equations 2-2 and 2-3) (Kleinbaum and Kupper 1978). This process was used to assign a data quality level for each analyte.

The matched pairs t-test was also used to evaluate whether the two sets of $\log_{10}$ transformed data were significantly different. The paired t-test compares data sets, which are composed of matched pairs of data. The significance of the relationship between two matched-pairs sets of data can be determined by comparing the calculated t-statistic with the critical t-value determined from a standard t-distribution table at the desired level of significance and degrees of freedom. To meet definitive level data quality requirements, both the slope and y-intercept had to be statistically the same as their ideal values, as defined in the demonstration plan, and the data had to be statistically similar as measured by the t-test. Log_{10} transformed data meeting these criteria were considered statistically equivalent to the $\log_{10}$ transformed reference data.

Slope Test for Significant Differences **(2-2)**

$$Z = \frac{m - 1}{\sqrt{SE_m + 0}}$$

where

> m is the slope of the regression line, SE is the standard error of the slope, and Z is the normal deviate test statistic.

Y-intercept Test for Significant Differences **(2-3)**

$$Z = \frac{b - 0}{\sqrt{SE_b - 0}}$$

where

b is the y-intercept of the regression line, SE is the standard error of the slope, and Z is the normal deviate test statistic.

If the r^2 was between 0.70 and 1, the precision RSD was less than 20 percent, and the slope or intercept were not statistically equivalent to their ideal values, the analyzer was considered to produce quantitative screening level data quality (Table 2-2). However, the linear regression was deemed sufficiently significant that bias could be identified and corrected. Results in this case could be mathematically corrected if 10 - 20 percent of the samples are sent to a reference laboratory. Reference laboratory analysis results for a percentage of the samples would provide a basis for determining a correction factor.

Data placed in the qualitative screening level category exhibit r^2 values less than 0.70. These data either were not statistically similar to the reference data based on inferential statistics or had a precision RSD greater than 20 percent. An analyzer producing data at this level is considered capable of detecting the presence or lack of contamination, above its detection limit, with at least a 90 percent accuracy rate, but is not considered suitable for reporting of concentrations.

MDLs for the analyzers were determined in two ways. One approach followed standard SW-846 protocol. In this approach, standard deviations (SD) from precision measurements for samples exhibiting contamination 5 to 10 times the estimated detection levels of the analyzers were multiplied by 3. The result represents the precision-based MDL for the analyzers.

In a second approach, MDLs were determined by analysis of the low concentration outliers on the $\log_{10}$ transformed FPXRF and $\log_{10}$ transformed reference method data cross plots. These cross plots for all analytes characteristically exhibited a region below the MDL where the linearity of the relationship disintegrated. Above the MDL, the FPXRF concentrations increased linearly with increasing reference method values. Effectively, the linear correlation between the two methods abruptly changes to no correlation below the MDL. An MDL value is assigned at two SDs above the concentration where this linear relationship disintegrates. This MDL represented a field- or performance-based MDL.

Deviations from the Demonstration Plan

Seven deviations were made from the demonstration plan during the on-site activities. The first dealt with determining the moisture content of samples. The demonstration plan stated that a portion of the original sample would be used for determining moisture content. Instead, a small portion of soil was collected immediately adjacent to the original sample location and was used for determining moisture content. This was done to conserve sample volume needed for the reference laboratory. The moisture content sample was not put through the homogenizing and sieving steps prior to drying.

The second deviation dealt with the sample drying procedures for moisture content determination. The demonstration plan required that the moisture content samples would be dried in a convection oven at 150 °C for 2 hours. Through visual observation, it was found that the samples were completely dried in 1 hour with samples heated to only 110 °C. Therefore, to conserve time, and to reduce the potential volatilization of metals, the samples for moisture content determination were dried in a convection oven at 110 °C for 1 hour.

The third deviation involved an assessment of analyzer drift due to changes in temperature. The demonstration plan indicated that at each site, each analyzer would measure the same SRM or PE sample at 2-hour intervals during at least one day of field operation. However, since ambient air temperature did not fluctuate more than 20 °F on any day throughout the demonstration, potential analyzer drift due to changes in temperature was not assessed.

The fourth deviation involved the drying of samples with a microwave. Instead of microwaving the samples on high for 5 minutes, as described in the demonstration plan, the samples were microwaved on high for only 3 minutes. This modification was made because the plastic weigh boats, which contained the samples, were melting and burning when left in the microwave for 5 minutes. In addition, many of the samples were melting to form a slag. PRC found (through visual observation) that the samples were completely dry after only 3 minutes of microwaving. This interval is still within common microwave drying times used in the field.

An analysis of the microwaved samples showed that this process had a significant impact on the analytical results. The mean RPD for the microwaved and nonmicrowaved raw data were significantly different at a 95 percent confidence level. This suggests that the microwave drying process somehow increases error and sample concentration variability. This difference may be due to the extreme heat and drying having an effect on the reference methods' extraction efficiency for target analytes. For the evaluation of the effects of microwave drying, there were 736 matched pairs of data where both element measurements were positive. Of these pairs, 471 exhibited RPDs less than 10 percent. This 10 percent level is within the acceptable precision limits for the reference laboratory as defined in the demonstration QAPP. Pairs exhibiting RPDs greater than 10 percent totaled 265. RPDs greater than 10 percent may have causes other than analysis-induced error. Of these 265, 96 pairs indicated an increase in metals concentration with microwaving, and 169 pairs indicated reductions in the concentration of metals. The RPDs for the microwaved samples were 2 to 3 times worse than the RPDs from the field duplicates. This further supports the hypothesis that microwave drying increases variability.

The fifth deviation involved reducing the percentage of analyzer precision measuring points. The demonstration plan called for 10 percent of the samples to be used for assessment of analyzer precision. Due to the time required to complete analysis of an analyzer precision sample, only 4 percent of the samples were used to assess analyzer precision. This reduction in samples was approved by the EPA technical advisor and the PRC field demonstration team leader. This eliminated 720 precision measurements and saved between 24 and 240 hours of analysis time. The final precision determinations for this demonstration were based on 48 sets of 10 replicate measurements for each analyzer.

The sixth deviation involved method blanks. Method blanks were to be analyzed each day and were to consist of a lithium carbonate that had been used in all sample preparation steps. Each analyzer had its own method blank samples, provided by the developer. Therefore, at the ASARCO site, each analyzer used its own method blank samples. However, at the RV Hopkins site, each analyzer used lithium carbonate method blanks that were prepared in the field, in addition to its own method blank samples. Both types of method blank analysis never identified method-induced contamination.

The seventh deviation involved assessing the accuracy of each analyzer. Accuracy was to be assessed through FPXRF analysis of 10 to 12 SRM or PE samples. Each analyzer measured a total of 28 SRM or

PE samples. Instead, PE samples were used to evaluate the accuracy of the reference methods, and SRMs were used to evaluate the accuracy of the analyzers. This is because the PE concentrations are based on acid extractable concentrations while SRM concentrations represent total metals concentration. SRM data was used for comparative purposes for the reference methods as were PE data for the FPXRF data.

Sample Homogenization

A key quality issue in this demonstration was ensuring that environmental samples analyzed by the reference laboratory and by each of the FPXRF analyzers were splits from a homogenized sample. To address this issue, sample preparation technicians exercised particular care throughout the field work to ensure that samples were thoroughly homogenized before they were split for analysis. Homogenization was conducted by kneading the soil in a plastic bag for a minimum of 2 minutes. If after this time the samples did not appear to be well homogenized, they were kneaded for an additional 2 minutes. This continued until the samples appeared to be well homogenized.

Sodium fluorescein was used as an indicator of sample homogenization. Approximately one-quarter teaspoon of dry sodium fluorescein powder was added to each sample prior to homogenization. After mixing, the sample was examined under an ultraviolet light to assess the distribution of sodium fluorescein throughout the sample. If the fluorescent dye was evenly dispersed in the sample, homogenization was considered complete. If the dye was not evenly distributed, the mixing was continued and repeatedly checked until the dye was evenly distributed throughout the sample.

To evaluate the homogenization process used in this demonstration, 70 field duplicate sample pairs were analyzed by the reference laboratory. Sample homogenization was critical to this demonstration; it assured that the samples measured by the analyzers were as close as possible to samples analyzed by the reference laboratory. This was essential to the primary objectives of this demonstration, the evaluation of comparability between analyzer results and those of the reference methods.

The homogenization process was evaluated by determining the RPD between paired field duplicate samples. The RPDs for the field duplicate samples reflect the total error for the homogenization process and the analytical method combined (Equation 2-4). When total error from the reference laboratory was determined for the entire data set, the resultant mean RPD total (error) and 95 percent confidence interval was 9.7 ± 1.4, for all metals reported. When only the primary analytes were considered, the RPD total (error) and 95 percent confidence interval was 7.6 ± 1.2, including the secondary analytes in the RPD calculation which produced a mean RPD total (error) and a 95 percent confidence interval of 9.3 ± 1.6.

$$Total\ Measurement\ Error\ =\ \sqrt{[(Sample\ Homogenization\ Error)^2\ +\ (Laboratory\ Error)^2]} \qquad \textbf{(2-4)}$$

Using internal QA/QC data from 27 analyses, it was possible to determine the reference laboratory's method error. The reference analytical method precision, as measured by the 95 percent confidence interval around the mean RPDs (laboratory error) of predigestion duplicate analyses, was 9.3 ± 2.9 for the target analytes.

To determine the error introduced by the sample homogenization alone, the error estimate for the reference methods was subtracted from the total error (Equation 2-5). Based on the data presented above, the laboratory-induced error was less than or approximately equal to the total error. This indicates that the sample homogenization (preparation) process contributed little or no error to the overall sample analysis process.

$$Sample\ Homogenization\ Error\ =\ \sqrt{[(Total\ Measurement\ Error)^2\ -\ (Laboratory\ Error)^2]} \tag{2-5}$$

Although the possibility for poorly homogenized samples exists under any homogenization routine, at the scale of analysis used by this demonstration, the samples were considered to be completely homogenized.

Section 3
Reference Laboratory Results

All soil samples collected from the ASARCO and RV Hopkins sites were submitted to the reference laboratory for trace metals analysis. The results are discussed in this section.

Reference Laboratory Methods

Samples collected during this demonstration were homogenized and split for extraction using EPA SW-846 Method 3050A. This is an acid digestion procedure where 1 to 2 grams of soil are digested on a hot plate with nitric acid, followed by hydrogen peroxide, and then refluxed with hydrochloric acid. One gram of soil was used for extraction of the demonstration samples. The final digestion volume was 100 milliliters (mL). The soil sample extracts were analyzed by Method 6010A.

Method 6010A provides analysis of metals using Inductively Coupled Plasma-Atomic Emission Spectroscopy (ICP-AES). This method requires that a plasma be produced by applying a radio-frequency field to a quartz tube wrapped by a coil or solenoid through which argon gas is flowing. The radio-frequency field creates a changing magnetic field in the flowing gas inside the coil, inducing a circulating eddy current on the argon gas that, in turn, heats it. Plasma is initiated by an ignition source and quickly stabilizes with a core temperature of 9,000 - 10,000 degrees Kelvin.

Soil sample extracts are nebulized, and the aerosol is injected into the plasma. Individual analytes introduced into the plasma absorb energy and are excited to higher energy states. These higher energy states have short lifetimes and the individual elements quickly fall back to their ground energy state by releasing a photon. The energy of the emitted photon is defined by the wavelength of electromagnetic radiation produced. Since many electronic transitions are possible for each individual element, several discrete emissions at different wavelengths are observed. Method 6010A provides one recommended wavelength to monitor for each analyte. Due to complex spectra with similar wavelengths from different elements in environmental samples, Method 6010A requires that interference corrections be applied for quantification of individual analytes.

Normal turnaround times for the analysis of soil samples by EPA SW-846 Methods 3050A/6010A range from 21 to 90 days depending on the complexity of the soil samples and the amount of QC documentation required. Faster turnaround times of 1 - 14 days can be obtained, but at additional cost.

Costs for the analysis of soil samples by EPA SW-846 Methods 3050A/6010A range from $150 to $350 per sample depending on turnaround times and the amount of QC documentation required. A sample turnaround of 28 days, a cost of $150 per sample, and a CLP documentation report for QC were chosen for this demonstration.

Reference Laboratory Quality Control

The reference laboratory, Midwest Research Institute (Kansas City, MO), holds certifications for performing target analyte list metals analysis with the U.S. Army Corps of Engineers-Missouri River Division, the State of California, and the State of Utah. These certifications include on-site laboratory audits, data package review audits, and the analysis of PE samples supplied by the certifying agency. PE samples are supplied at least once per year from each of the certifying agencies. The reference laboratory's results for the PE samples are compared to true value results and certifying agency acceptance limits for the PE samples. Continuation of these certifications hinges upon acceptable results for the audits and the PE samples.

The analysis of soil samples by the reference laboratory was governed by the QC criteria in its SOPs, Method 6010A, and the demonstration QAPP. Table 3-1 provides QAPP QC requirements that were monitored and evaluated for the target analytes. Method 6010A QC guidelines also are included in Table 3-1. Due to the complex spectra derived from the analysis of the demonstration samples, the QAPP QC requirements were applied only to the primary analytes. The QAPP QC requirements also were monitored and evaluated for the secondary analytes and other analytes reported by the reference laboratory. However, corrective actions were not required for the secondary analytes.

Table 3-1. Reference Laboratory Quality Control Parameters[a]

Parameter	Frequency	Reference Method Requirement	QAPP Requirement
Initial Calibration Verification (ICV) Standard	With each initial calibration	±10 percent of true value	±10 percent of true value
Continuing Calibration Verification (CCV) Standard	After analysis of every 10 samples and at the end of analytical run	±10 percent of true value	±10 percent of true value
Initial and Continuing Calibration Blanks (ICB) and (CCB)	With each continuing calibration, after analysis of every 10 samples, and at the end of analytical run	±3 standard deviations of the analyzer background mean	No target analytes at concentrations greater than 2 times the lower reporting limit (LRL)
Interference Check Standard (ICS)	With every initial calibration and after analysis of 20 samples	±20 percent of true value	±20 percent of true value
High Level Calibration Check Standard	With every initial calibration	±5 percent of true value	±10 percent of true value
Method Blanks	With each batch of samples of a similar matrix	No QC requirement specified	No target analytes at concentrations greater than 2 times the LRL
Laboratory Control Samples	With each batch of samples of a similar matrix	No QC requirement specified	80 - 120 percent recovery
Predigestion Matrix Spike Samples	With each batch of samples of a similar matrix	80 - 120 percent recovery	80 - 120 percent recovery
Postdigestion Matrix Spike Samples	With each batch of samples of a similar matrix	75 - 125 percent recovery	80 - 120 percent recovery

Table 3-1. ***Continued***

Parameter	Frequency	Reference Method Requirement	QAPP Requirement
Performance Evaluation Samples	As submitted during demonstration	No QC requirement specified	80 - 120 percent recovery within performance acceptance limits (PAL)
Predigestion Laboratory Duplicate Samples	With each batch of samples of a similar matrix	20 percent relative percent difference (RPD)[b]	20 percent RPD[c]
Postdigestion Laboratory Duplicate Samples	With each batch of samples of a similar matrix	No QC requirement specified	10 percent RPD[c]

Notes:

[a] Quality control parameters were evaluated on the raw reference data.

[b] RPD control limits only pertain to original and laboratory duplicate sample results that were greater than 10 times the instrument detection limit (IDL).

[c] RPD control limits only pertain to original and laboratory duplicate sample results that were greater than or equal to 10 times the LRL.

PRC performed three on-site audits of the reference laboratory during the analysis of predemonstration and demonstration samples. These audits were conducted to observe and evaluate the procedures used by the reference laboratory and to ensure that these procedures adhered to the QAPP QC requirements. Audit findings revealed that the reference laboratory followed the QAPP QC requirements. It was determined that the reference laboratory had problems meeting two of the QAPP QC requirements: method blank results and the high level calibration check standard's percent recovery. Due to these problems, these two QAPP QC requirements were widened. The QC requirement for method blank sample results was changed from no target analytes at concentrations greater than the lower reporting limit (LRL) to two times the LRL. The QC requirement for the high level calibration standard percent recovery was changed from ±5 to ±10 percent of the true value. These changes were approved by the EPA and did not affect the results of the demonstration.

The reference laboratory internally reviewed its data before releasing it. PRC conducted a QC review on the data based on the QAPP QC requirements and corrective actions listed in the demonstration plan.

Quality Control Review of Reference Laboratory Data

The QC data review focused upon the compliance of the data with the QC requirements specified in the demonstration QAPP. The following sections discuss results from the QC review of the reference laboratory data. All QC data evaluations were based on raw data.

Reference Laboratory Sample Receipt, Handling, and Storage Procedures

Demonstration samples were divided into batches of no more than 20 samples per batch prior to delivery to the reference laboratory. A total of 23 batches containing 315 samples and 70 field duplicate samples was submitted to the reference laboratory. The samples were shipped in sealed coolers at ambient temperature under a chain of custody.

Upon receipt of the demonstration samples, the reference laboratory assigned each sample a unique number and logged each into its laboratory tracking system. The samples were then transferred to the reference laboratory's sample storage refrigerators to await sample extraction.

Samples were transferred to the extraction section of laboratory under an internal chain of custody. Upon completion of extraction, the remaining samples were returned to the sample storage refrigerators. Soil sample extracts were refrigerated in the extraction laboratory while awaiting sample analysis.

Sample Holding Times

The maximum allowable holding time from the date of sample collection to the date of extraction and analysis using EPA SW-846 Methods 3050A/6010A is 180 days. Maximum holding times were not exceeded for any samples during this demonstration.

Initial and Continuing Calibrations

Prior to sample analysis, initial calibrations (ICAL) were performed. ICALs for Method 6010A consist of the analysis of three concentrations of each target analyte and a calibration blank. The low concentration standard is the concentration used to verify the LRL of the method. The remaining standards are used to define the linear range of the ICP-AES. The ICAL is used to establish calibration curves for each target analyte. Method 6010A requires an initial calibration verification (ICV) standard to be analyzed with each ICAL. The method control limit for the ICV is ±10 percent. An interference check sample (ICS) and a high level calibration check standard is required to be analyzed with every ICAL to assess the accuracy of the ICAL. The control limits for the ICS and high level calibration check standard were ±20 percent recovery and ±10 percent of the true value, respectively. All ICALs, ICVs, and ICSs met the respective QC requirements for all target analytes.

Continuing calibration verification (CCV) standards and continuing calibration blanks (CCB) were analyzed following the analysis of every 10 samples and at the end of an analytical run. Analysis of the ICS was also required after every group of 20 sample analyses. These QC samples were analyzed to check the validity of the ICAL. The control limits for the CCVs were ±10 percent of the true value. The control limits for CCBs were no target analyte detected at concentrations greater than 2 times the LRL. All CCVs, CCBs, and ICSs met the QAPP requirements for the target analytes with the exception of one CCV where the barium recovery was outside the control limit. Since barium was a primary analyte, the sample batch associated with this CCV was reanalyzed and the resultant barium recovery met the QC criteria.

Detection Limits

The reference laboratory LRLs for the target analytes are listed in Table 3-2. These LRLs were generated through the use of an MDL study of a clean soil matrix. This clean soil matrix was also used for method blank samples and LCSs during the analysis of demonstration samples. The MDL study involved seven analyses of the clean soil matrix spiked with low concentrations of the target analytes. The mean and standard deviation of the response for each target analyte was calculated. The LRL was defined as the mean plus three times the standard deviation of the response for each target analyte included in the method detection limit study. All LRLs listed in Table 3-2 were met and maintained throughout the analysis of the demonstration samples.

The reference laboratory reported soil sample results in units of milligram per kilogram wet weight. All reference laboratory results referred to in this report are wet-weight sample results.

Table 3-2. SW-846 Method 6010A LRLs for Target Analytes

Analyte	LRL (mg/kg)	Analyte	LRL (mg/kg)
Antimony	6.4	Copper*	1.2
Arsenic*	10.6	Iron	600[a]
Barium*	5.0	Lead*	8.4
Cadmium	0.80	Nickel	3.0
Chromium*	2.0	Zinc*	2.0

Notes:

a	LRL elevated due to background interference.
*	Primary analyte.
mg/kg	Milligrams per kilogram.

Method Blank Samples

Method blanks were prepared using a clean soil matrix and acid digestion reagents used in the extraction procedure. A minimum of one method blank sample was analyzed for each of the 23 batches of demonstration samples submitted for reference laboratory analysis. All method blanks provided results for target analytes at concentrations less than 2 times the levels shown in Table 3-2.

Laboratory Control Samples

All LCSs met the QAPP QC requirements for all primary and secondary analytes except those discussed below.

The primary analytes copper and lead were observed outside the QC limits in one of the 23 batches of samples analyzed. Reanalysis of the affected batches was not performed by the reference laboratory. These data were qualified by the reference laboratory. Copper and lead data for all samples included in the affected batches were rejected and not used for demonstration statistical comparisons.

Concentrations of secondary analytes antimony, nickel, and cadmium were observed outside the QC limits in the LCSs. Antimony LCS recoveries were continually outside the control limits, while nickel and cadmium LCS recoveries were only occasionally outside QC limits. Antimony was a problem analyte and appeared to be affected by acid digestion, which can cause recoveries to fall outside control limits. Antimony recoveries ranged from 70 to 80 percent. Since secondary analytes were not subject to the corrective actions listed in the demonstration QAPP, no reanalysis was performed based on the LCS results of the secondary target analytes. These values were qualified by the reference laboratory. All other secondary analyte LCS recoveries fell within the QAPP control limits.

Predigestion Matrix Spike Samples

One predigestion matrix spike sample and duplicate were prepared by the reference laboratory for each batch of demonstration samples submitted for analysis. The predigestion matrix spike duplicate sample was not required by the QAPP, but it is a routine sample prepared by the reference laboratory. This duplicate sample can provide data that indicates if out-of-control recoveries are due to matrix interferences or laboratory errors.

Predigestion spike recovery results for the primary analytes arsenic, barium, chromium, copper, lead, and zinc were outside control limits for at least 1 of the 23 sample batches analyzed by the reference method. These control limit problems were due to either matrix effects or initial spiking concentrations below native analyte concentrations.

Barium, copper, and lead predigestion matrix spike recovery results were outside control limits in sample batches 2, 3, and 5. In all of these cases, the unacceptable recoveries were caused by spiking concentrations that were much lower than native concentrations of the analytes. These samples were re-prepared, spiked with higher concentrations of analytes, reextracted, and reanalyzed. Following this procedure, the spike recoveries fell within control limits upon reanalysis.

One predigestion matrix spike recovery was outside control limits for arsenic. The predigestion matrix spike duplicate sample also was outside of control limits. This sample exhibited an acceptable RPD for the recovery of arsenic in the predigestion matrix spike and duplicate. A matrix interference may have been responsible for the low recovery. This sample was not reanalyzed.

Chromium predigestion matrix spike recoveries were outside control limits in 7 of the 23 batches of samples analyzed. Five of these seven failures exhibited recoveries ranging from 67 to 78 percent, close to the low end of the control limits. These recoveries were similar in the predigestion matrix spike duplicate samples prepared and analyzed in the same batch. This indicates that these five failures were due to matrix interferences. The predigestion matrix spike duplicate samples prepared and analyzed along with the remaining two failures did not agree with the recoveries of the postdigestion matrix spike samples, indicating that these two failures may be due to laboratory error, possibly inaccuracies in sample spiking. These seven predigestion matrix spike samples were not reanalyzed.

The zinc predigestion matrix spike recovery data were outside control limits for four batches of samples analyzed. In three of the spike recovery pairs, recoveries ranged from 70 to 76 percent, close to the lower end of the control limits. The fourth recovery was much less than the lower end of the control limits. All of the predigestion matrix spike duplicate samples provided recoveries that agreed with the recoveries for the predigestion matrix spike sample recoveries indicating that the low recoveries were due to matrix effects. These predigestion matrix spikes and associated samples were not reanalyzed.

The secondary analytes, cadmium, iron, and nickel, had predigestion spike recoveries outside control limits. Cadmium spike recoveries were outside control limits six times. These recoveries ranged from 71 to 85 percent. Iron spike recoveries were outside of control limits once. Nickel spike recoveries were outside control limits four times. These recoveries ranged from 74 to 83 percent. Antimony spike recoveries were always outside control limits. No corrective action was taken for these secondary target analytes.

Demonstration sample results for all target analytes that did not meet the control limits for predigestion matrix spike recovery were qualified by the reference laboratory.

Postdigestion Matrix Spike Samples

All postdigestion matrix spike results were within the control limit of 80 - 120 percent recovery for the primary analytes.

Secondary analytes, antimony, and iron were observed outside the control limits. However, no corrective action was taken for secondary analytes as stated in the demonstration QAPP. All postdigestion spike recoveries for target analytes met the QA/QC requirements of the QAPP and were considered acceptable.

Predigestion Laboratory Duplicate Samples

Predigestion laboratory duplicate RPD results were within the control limit of 20 percent for analyte concentrations greater than 10 times the LRL except for the following instances. RPDs for primary analytes barium, arsenic, lead, chromium, and copper were observed above the control limit in five predigestion laboratory duplicate samples. These samples were reanalyzed according to the corrective actions listed in the QAPP. The reanalysis produced acceptable RPD results for these primary analytes.

RPD results for the secondary analytes antimony, nickel, and cadmium were observed outside the control limit for a number of sample batches. No corrective action was taken for secondary analytes that exceeded the RPD control limit.

Postdigestion Laboratory Duplicate Samples

All primary analyte postdigestion laboratory duplicate RPD results were less than the 10 percent control limit for analyte concentrations greater than 10 times the LRL.

The RPDs for secondary analytes antimony and iron were observed above the 10 percent control limit in two sample batches. No corrective action was taken for secondary target analytes that exceeded the RPD control limit.

Performance Evaluation Samples

PE samples were purchased from Environmental Resource Associates (ERA). The PE samples are Priority PollutnT™/Contract Laboratory Program (CLP) QC standards for inorganics in soil. This type of sample is used by the EPA to verify accuracy and laboratory performance. Trace metal values are certified by interlaboratory round robin analyses. ERA lists performance acceptance limits (PAL) for each analyte that represent a 95 percent confidence interval (CI) around the certified value. PALs are generated by peer laboratories in ERA's InterLaB™ program using the same samples that the reference laboratory analyzed and the same analytical methods. The reported value for each analyte in the PE sample must fall within the PAL range for the accuracy to be acceptable. Four PE samples were submitted "double blind" (the reference laboratory was not notified that the samples were QC samples or of the certified values for each element) to the reference laboratory for analysis by EPA SW-846 Methods 3050A/6010A. Reference laboratory results for all target analytes are discussed later in this section.

Four certified reference materials (CRM) purchased from Resource Technology Corporation (RTC) also were used as PE samples to verify the accuracy and performance of the reference laboratory. These four CRMs were actual samples from contaminated sites. They consisted of two soils, one sludge, and one ash CRM. Metal values in the CRMs are certified by round robin analyses of at least 20 laboratories according to the requirements specified by the EPA Cooperative Research and Development Agreement. The certified reference values were determined by EPA SW-846 Methods 3050A/6010A. RTC provides a 95 percent PAL around each reference value in which measurements should fall 19 of 20 times. The reported value from the reference laboratory for each analyte must fall within this PAL for the accuracy to be considered acceptable. As with the four PE samples, the four CRMs were submitted “double blind” to the reference laboratory for analysis by EPA SW-846 Methods 3050A/6010A. The reference laboratory results for the target analytes are discussed later in the Accuracy subsection.

Standard Reference Material Samples

As stated in the demonstration plan (PRC 1995), PE samples also consisted of SRMs. The SRMs consisted of solid matrices such as soil, ash, and sludge. Certified analyte concentrations for SRMs are determined on an analyte by analyte basis by multiple analytical methods including but not limited to ICP-AES, flame atomic absorption spectroscopy, ICP-mass spectrometry, XRF, instrumental neutron activation analysis, hydride generation atomic absorption spectroscopy, and polarography. These certified values represent total analyte concentrations and complete extraction. This is different from the PE samples, CRM samples, and the reference methods, which use acid extraction that allows quantitation of only acid extractable analyte concentrations.

The reference laboratory analyzed 14 SRMs supplied by the National Institute of Standards and Technology (NIST), U.S. Geological Survey (USGS), National Research Council Canada, South African Bureau of Standards, and Commission of the European Communities. The percentage of analyses of SRMs that were within the QAPP-defined control limits of 80 - 120 percent recovery was calculated for each primary and secondary analyte.

Analyses of SRMs were not intended to assess the accuracy of EPA SW-846 Methods 3050A/6010A as were the ERA PE or RTC CRM samples. Comparison of EPA SW-846 Methods 3050A/6010A acid leach data to SRM data cannot be used to establish method validity (Kane and others 1993). This is because SRM values are acquired by analyzing the samples by methods other than the ICP-AES method. In addition, these other methods use sample preparation techniques different from those for EPA SW-846 Methods 3050A/6010A. This is one reason no PALs are published with the SRM certified values. Therefore, the SRMs were not considered an absolute test of the reference laboratory’s accuracy for EPA SW-846 Methods 3050A/6010A.

The SRM sample results were not used to assess method accuracy or to validate the reference methods. This was due to the fact that the reported analyte concentrations for SRMs represent total analyte concentrations. The reference methods are not an analysis of total metals; rather they target the leachable concentrations of metals. This is consistent with the NIST guidance against using SRMs to assess performance on leaching based analytical methods (Kane and others 1993).

Data Review, Validation, and Reporting

Demonstration data were internally reviewed and validated by the reference laboratory. Validation involved the identification and qualification of data affected by QC procedures or samples that did not meet the QC requirements of the QAPP. Validated sample results were reported using both hard copy and electronic disk deliverable formats. QC summary reports were supplied with the hard copy results. This qualified data was identified and discussed in the QC summary reports provided by the reference laboratory.

Demonstration data reported by the reference laboratory contained three types of data qualifiers: C, Q, and M. Type C qualifiers included the following:

- U - the analyte was analyzed for but not detected.

- B - the reported value was obtained from a reading that was less than the LRL but greater than or equal to the IDL.

Type Q qualifiers included the following:

- N - spiked sample recovery was not within control limits.
- * - duplicate analysis was not within control limits.

Type M qualifiers include the following:

- P - analysis performed by ICP-AES (Method 6010).

Quality Assessment of Reference Laboratory Data

An assessment of the reference laboratory data was performed using the PARCC parameters discussed in Section 2. PARCC parameters are used as indicators of data quality and were evaluated using the review of reference laboratory data discussed above. The following sections discuss the data quality for each PARCC parameter. This quality assessment was based on raw reference data and the raw PE sample data.

The quality assessment was limited to an evaluation of the primary analytes. Secondary and other analytes reported by the reference laboratory were not required to meet the QC requirements specified in the QAPP. Discussion of the secondary analytes is presented in the precision, accuracy, and comparability sections for informational purposes only.

Precision

Precision for the reference laboratory data was assessed through an evaluation of the RPD produced from the analysis of predigestion laboratory duplicate samples and postdigestion laboratory duplicate samples. Predigestion laboratory duplicate samples provide an indication of the method precision, while postdigestion laboratory duplicate samples provide an indication of instrument performance. Figure 3-1 provides a graphical summary of the reference method precision data.

The predigestion duplicate RPDs for the primary and secondary analytes fell within the 20 percent control limit, specified in the QAPP, for 17 out of 23 batches of demonstration samples. The six results that exceeded the control limit involved only 11 of the 230 samples evaluated for predigestion duplicate precision (Figure 3-1). This equates to 95 percent of the predigestion duplicate data meeting the QAPP control limits. Six of the analytes exceeding control limits had RPDs less than 30 percent. Three of the analytes exceeding control limits had RPDs between 30 and 40 percent. Two of the analytes exceeding control limits had RPDs greater than 60 percent. These data points are not shown in Figure 3-1. Those instances where the control limits were exceeded are possibly due to nonhomogeneity of the sample or simply to chance, as would be expected with a normal distribution of precision analyses.

The postdigestion duplicate RPDs for the primary and secondary analytes fell within the 10 percent control limit, specified in the QAPP, for 21 out of 23 batches of demonstration samples. The two results that exceeded the control limit involved only 3 of the 230 samples evaluated for postdigestion duplicate precision in the 23 sample batches (Figure 3-1). This equates to 99 percent of the postdigestion duplicate data meeting the QAPP control limits. The RPDs for the three results that exceeded the control limit ranged from 11 to 14 percent.

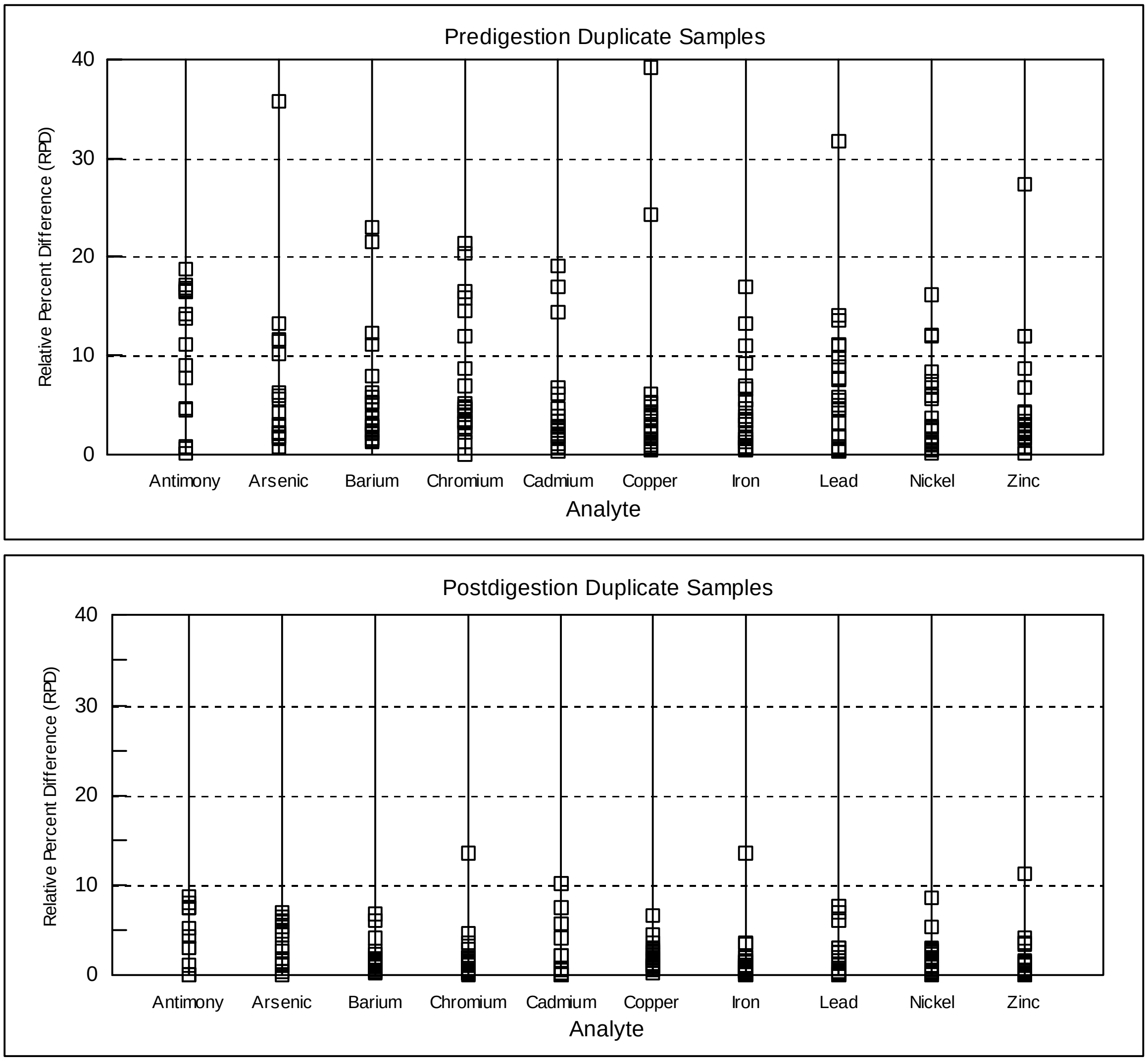


Figure 3-1. Pre- and Postdigestion Duplicate Samples: The top graph illustrates the reference laboratory's performance on analyzing <u>predigestion</u> duplicate samples. Twenty percent RPD represents the <u>predigestion</u> duplicate control limits defined in the demonstration QAPP. Two points were deleted from this top graph: barium at 65 percent RPD and copper at 138 percent RPD. The bottom graph illustrates the reference laboratory's performance on analyzing <u>postdigestion</u> duplicate samples. Ten percent RPD represents the <u>postdigestion</u> duplicate control limits defined in the demonstration QAPP.

Accuracy

Accuracy for the reference laboratory data was assessed through evaluations of the PE samples (including the CRMs), LCSs, method blank sample results, and pre- and postdigestion matrix spike samples. PE samples were used to assess the absolute accuracy of the reference laboratory method as a whole, while LCSs, method blanks, and pre- and postdigestion matrix spike samples were used to assess the accuracy of each batch of demonstration samples.

A total of eight PE and CRM samples was analyzed by the reference laboratory. These included four ERA PE samples and four RTC CRM samples. One of the ERA PE samples was submitted to the reference laboratory in duplicate, thereby producing nine results to validate accuracy. The accuracy data for all primary and secondary analytes are presented in Table 3-3 and displayed in Figure 3-2. Accuracy was assessed over a wide-concentration range for all 10 analytes with concentrations for most analytes spanning one or more orders of magnitude.

Reference laboratory results for all target analytes in the ERA PE samples fell within the PALs. In the case of the RTC CRM PE samples, reference laboratory results for copper in one CRM and zinc in two CRMs fell outside the published acceptance limits. One of the two out-of-range zinc results was only slightly above the upper acceptance limit (811 versus 774 mg/kg). The other out-of-range zinc result and the out-of-range copper result were about three times higher than the certified value and occurred in the same CRM. These two high results skewed the mean percent recovery for copper and zinc shown in Table 3-3. Figure 3-2 shows that the remaining percent recoveries for copper and zinc were all near 100 percent.

Table 3-3 shows that a total of 83 results was obtained for the 10 target analytes. Eighty of the 83 results or 96.4 percent fell within the PALs. Only 3 out of 83 times did the reference method results fall outside PALs. This occurred once for copper and twice for zinc. Based on this high percentage of acceptable results for the ERA and CRM PE samples, the accuracy of the reference methods was considered acceptable.

Table 3-3. Reference Laboratory Accuracy Data for Target Analytes

Analyte	n	Percent Within Acceptance Range	Mean Percent Recovery	Range of Percent Recovery	SD of Percent Recovery	Concentration Range (mg/kg)
Antimony	6	100	104	83 - 125	15	50 - 4,955
Arsenic	8	100	106	90 - 160	22	25 - 397
Barium	9	100	105	83 - 139	21	19 - 586
Cadmium	9	100	84	63 - 93	10	1.2 - 432
Chromium	9	100	91	77 - 101	8	11 - 187
Copper	9	89	123	90 - 332	79	144 - 4,792
Iron	7	100	98	79 - 113	12	6,481 - 28,664
Lead	8	87.5	86	35 - 108	22	52 - 5,194
Nickel	9	100	95	79 - 107	10	13 - 13,279
Zinc	9	78	120	79 - 309	72	76 - 3,021

Notes:		
	n	Number of samples with detectable analyte concentrations.
	SD	Standard deviation.
	mg/kg	Milligrams per kilogram.

LCS percent recoveries for all the primary analytes were acceptable in 21 of the 23 sample batches. Lead recovery was unacceptable in one sample batch and lead results for each sample in that batch were rejected.

Copper recovery was unacceptable in another sample batch, and copper results for each sample in this batch also were rejected. Percent recoveries of the remaining primary analytes in each of these two batches were acceptable. In all, 136 of 138 LCS results or 98.5 percent fell within the control limits.

Method blank samples for all 23 batches of demonstration samples provided results of less than 2 times the LRL for all primary analytes. This method blank control limit was a deviation from the QAPP, which had originally set the control limit at no target analytes at concentrations greater than the LRL. This control limit was widened at the request of the reference laboratory. A number of batches were providing method blank results for target analytes at concentrations greater than the LRL, but less than 2 times the LRL. This alteration was allowed because even at 2 times the LRL, positive results for the method blank samples were still significantly lower than the MDLs for each of the FPXRF analyzers. The results from the method blank samples did not affect the accuracy of the reference data as it was to be used in the demonstration statistical evaluation of FPXRF analyzers.

The percent recovery for the predigestion matrix spike samples fell outside of the 80 - 120 percent control limit specified in the QAPP in several of the 23 batches of demonstration samples. The predigestion matrix spike sample results indicate that the accuracy of specific target analytes in samples from the affected batches may be suspect. These results were qualified by the reference laboratory. These data were not excluded from use for the demonstration statistical comparison. A discussion of the use of this qualified data is included in the "Use of Qualified Data for Statistical Analysis" subsection.

The RPD for the postdigestion matrix spike samples fell within the 80 - 120 percent control limit specified in the QAPP for all 23 batches of demonstration samples.

The QA review of the reference laboratory data indicated that the absolute accuracy of the method was acceptable. Based on professional judgement, it was determined that the small percentage of outliers did not justify rejection of any demonstration sample results from the reference laboratory. The accuracy assessment also indicated that most of the batch summary data were acceptable. Two batches were affected by LCS outliers, and some data were qualified due to predigestion matrix spike recovery outliers. This data was rejected or qualified. Rejected data was not used. Qualified data were used as discussed below.

Representativeness

Representativeness of the analytical data was evaluated through laboratory audits performed during the course of sample analysis and by QC sample analyses, including method blank samples, laboratory duplicate samples, and CRM and PE samples. These QC samples were determined to provide acceptable results. From these evaluations, it was determined that representativeness of the reference data was acceptable.

Completeness

Results were obtained for all soil samples extracted and analyzed by EPA SW-846 Methods 3050A/6010A. Some results were rejected or qualified. Rejected results were deemed incomplete. Qualified results were usable for certain purposes and were deemed as complete.

To calculate completeness, the number of nonrejected results was determined. This number was divided by the total number of results expected, and then multiplied by 100 to express completeness as a percentage. A total of 385 samples was submitted for analysis. Six primary analytes were reported, resulting in an expected 2,310 results. Forty of these were rejected, resulting in 2,270 complete results. Reference laboratory completeness was determined to be 98.3 percent, which exceeded the objective for this demonstration of 95 percent. The reference laboratory's completeness was, therefore, considered acceptable.

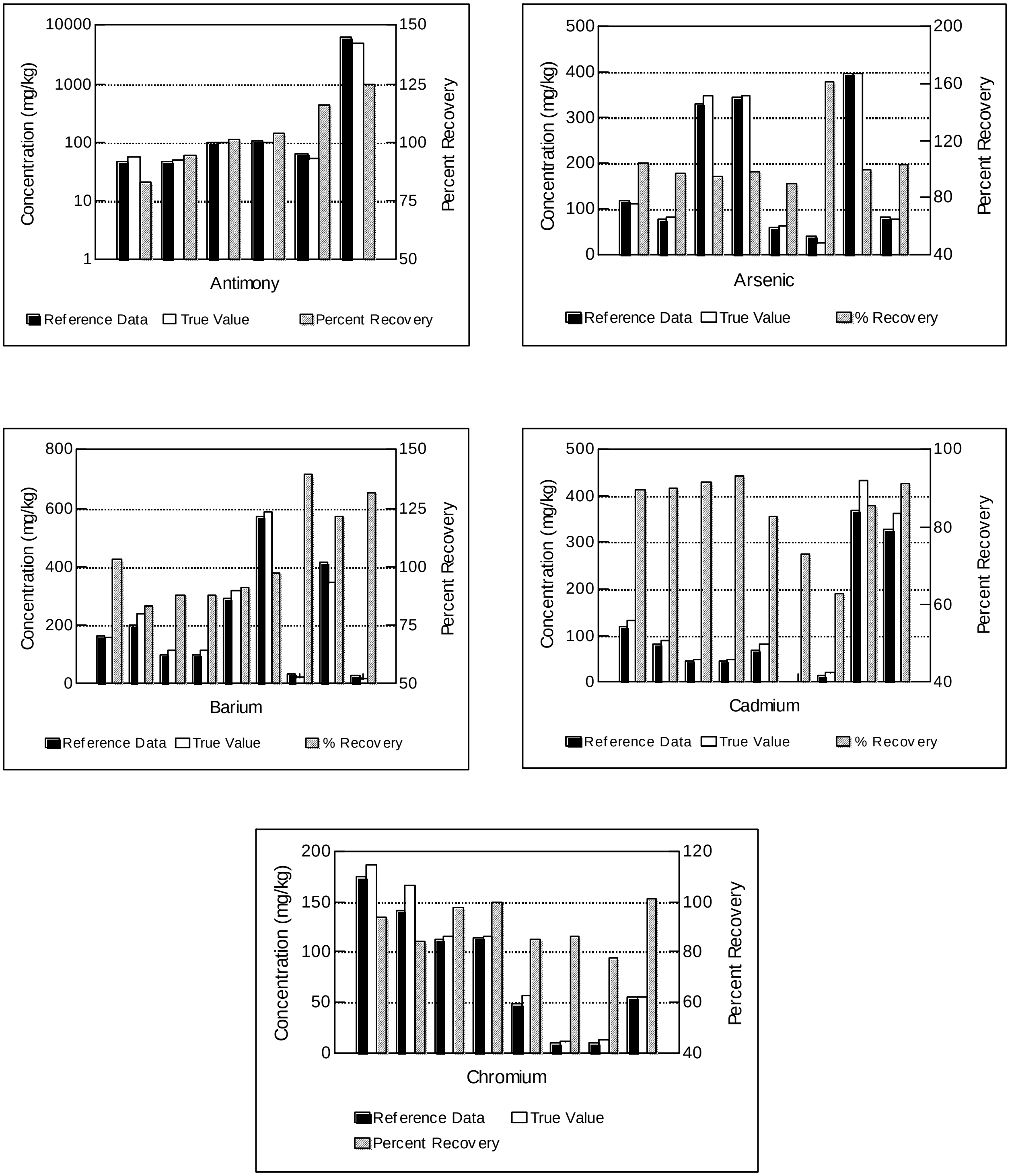


Figure 3-2. Reference Method PE and CRM Results: These graphs illustrate the relationship between the reference data and the true values for the PE or CRM samples. The gray bars represent the percent recovery for the reference data. Each set of three bars (black, white, and gray) represents a single PE or CRM sample. Based on this high percentage of acceptable results for the ERA and CRM PE samples, the accuracy of the reference laboratory method was considered acceptable.

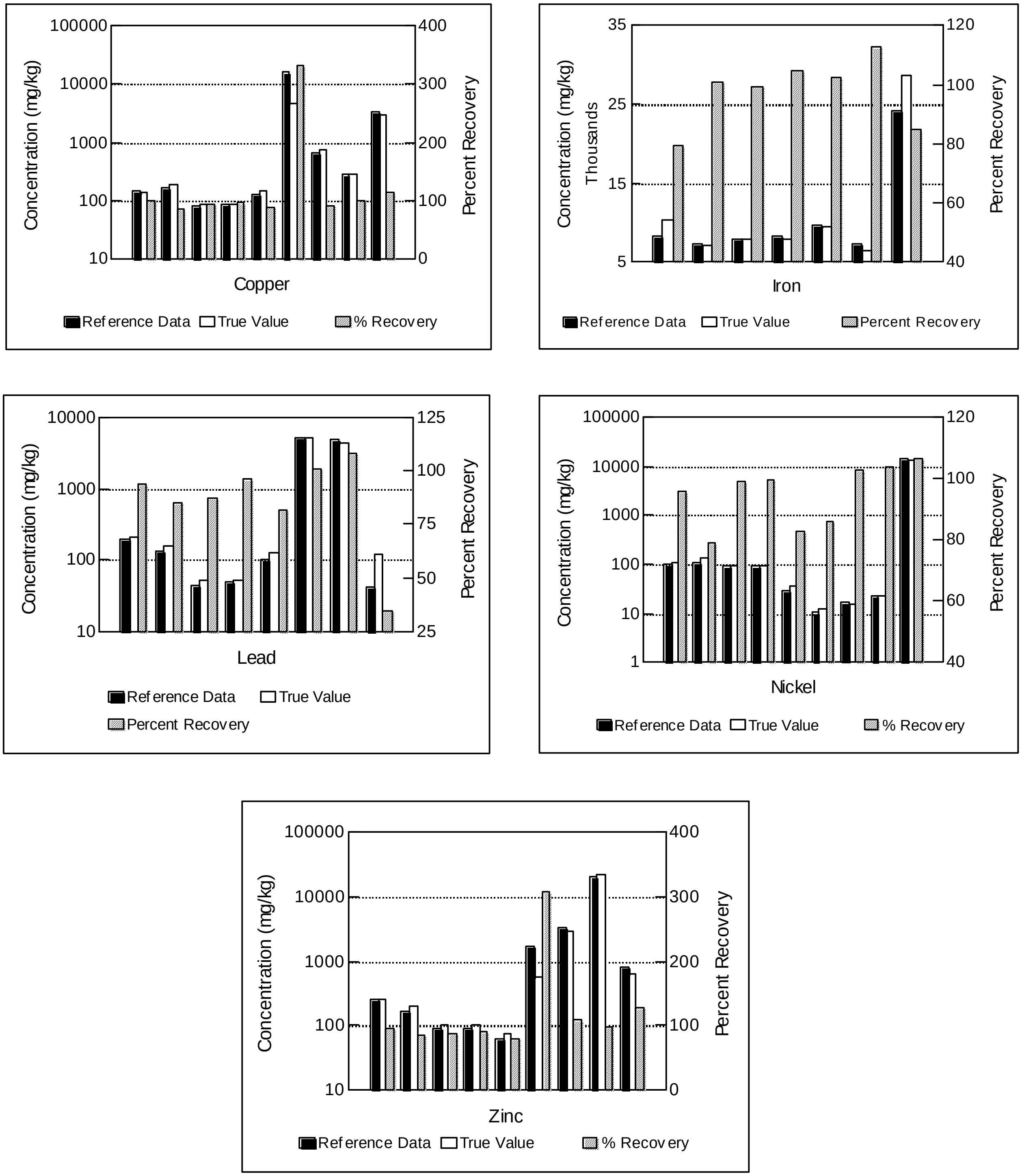


Figure 3-2 (Continued). Reference Method PE and CRM Results: These graphs illustrate the relationship between the reference data and the true values for the PE or CRM samples. The gray bars represent the percent recovery for the reference data. Each set of three bars (black, white, and gray) represents a single PE or CRM sample. Based on this high percentage of acceptable results for the ERA and CRM PE samples, the accuracy of the reference laboratory method was considered acceptable.

Comparability

Comparability of the reference data was controlled by following laboratory SOPs written for the performance of sample analysis using EPA SW-846 Methods 3050A/6010A. QC criteria defined in the SW-846 methods and the demonstration plan (PRC 1995) were followed to ensure that reference data would provide comparable results to any laboratory reporting results for the same samples.

Reference results indicated that EPA SW-846 Methods 3050A/6010A did not provide comparable results for some analytes in the SRM samples. SRM performance data for target analytes is summarized in Table 3-4 and displayed in Figure 3-3. As with the PEs, the analyte concentrations spanned up to 3 orders of magnitude in the SRMs. The percentage of acceptable (80 - 120 percent recovery) SRM results and mean percent recovery was less than 50 percent for the analytes antimony, barium, chromium, iron, and nickel. The low recoveries for these five analytes reflect the lesser tendency for them to be acid-extracted (Kane and others 1993).

Under contract to the EPA, multiple laboratories analyzed NIST SRMs 2709, 2710, and 2711 by EPA SW-846 Methods 3050A/6010A. A range, median value, and percent leach recovery based on the median value for each detectable element were then published as an addendum to the SRM certificates. These median values are not certified but provide a baseline for comparison to other laboratories analyzing these SRMs by EPA SW-846 Methods 3050A/6010A. Table 3-5 presents the published percent leach recovery for the 10 primary and secondary analytes and the reference laboratory's results for these three NIST SRMs. Table 3-5 shows that the results produced by the reference laboratory were consistent with the published results indicating good comparability to other laboratories using the same analytical methods on the same samples.

Table 3-4. SRM Performance Data for Target Analytes

Analyte	n	Percent Within Acceptance Range	Mean Percent Recovery	Range of Percent Recovery	SD of Percent Recovery	Concentration Range (mg/kg)
Antimony	5	0	22	15 - 37	9	3.8 - 171
Arsenic	11	72	84	67 - 106	10	18 - 626
Barium	8	12	41	21 - 89	21	414 - 1,300
Cadmium	10	50	80	43 - 95	15	2.4 - 72
Chromium	10	0	45	14 - 67	16	36 - 509
Copper	17	88	82	33 - 94	17	35 - 2,950
Iron	7	14	62	23 - 84	25	28,900 - 94,000
Lead	17	82	83	37 - 99	17	19 - 5,532
Nickel	16	19	67	25 - 91	17	14 - 299
Zinc	16	75	81	32 - 93	14	81 - 6,952

Notes:		
	n	Number of SRM samples with detectable analyte concentrations.
	SD	Standard deviation.
	mg/kg	Milligrams per kilogram.

Table 3-5. Leach Percent Recoveries for Select NIST SRMs

Analyte	NIST SRM 2709		NIST SRM 2710		NIST SRM 2711	
	Published Result[a]	Reference Laboratory Result	Published Result[a]	Reference Laboratory Result	Published Result[a]	Reference Laboratory Result
Antimony	–	–	21	–	–	20
Arsenic	–	106	94	87	86	91
Barium	41	37	51	45	28	25
Cadmium	–	–	92	84	96	87
Chromium	61	–	49	–	43	49
Copper	92	85	92	92	88	90
Iron	86	84	80	78	76	66
Lead	69	87	92	96	95	90
Nickel	89	76	71	69	78	70
Zinc	94	78	85	88	89	85

Notes:

[a] Published results found in an addendum to SRM certificates for NIST SRMs 2709, 2710, and 2711.

NIST National Institute of Standards and Technology.

SRM Standard reference materials.

– Analyte not present above the method LRL.

The inability of EPA SW-846 Methods 3050A/6010A to achieve the predetermined 80 - 120 percent recovery requirement indicated that the methods used to determine the certified values for the SRM samples were not comparable to EPA SW-846 Methods 3050A/6010A. Differences in the sample extraction methods and the use of different analytical instruments and techniques for each method were the major factors of this noncomparability. Because of these differences, it was not surprising that the mean percent recovery was less than 100 percent for the target analytes. The lack of comparability of EPA SW-846 Methods 3050A/6010A to the total metals content in the SRMs did not affect the quality of the data generated by the reference laboratory.

The assessment of comparability for the reference data revealed that it should be comparable to other laboratories performing analysis of the same samples using the same extraction and analytical methods, but it may not be comparable to laboratories performing analysis of the same samples using different extraction and analytical methods or by methods producing total analyte concentration data.

Use of Qualified Data for Statistical Analysis

As noted above, the reference laboratory results were reported and validated, qualified, or rejected by approved QC procedures. Data were qualified for predigestion matrix spike recovery and pre- and postdigestion laboratory duplicate RPD control limit outliers. None of the problems were considered sufficiently serious to preclude the use of coded data. Appropriate corrective action identified in the demonstration plan (PRC 1995) was instituted. The result of the corrective action indicated that the poor percent recovery and RPD results were due to matrix effects. Since eliminating the matrix effects would require additional analysis using a different determination method such as atomic absorption spectrometry, or the method of standard addition, the matrix effects were noted and were not corrected.

PARCC parameters for the reference laboratory data were determined to be acceptable. It was expected that any laboratory performing analysis of these samples using EPA SW-846 Methods 3050A/6010A would experience comparable matrix effects. A primary objective of this demonstration was to compare sample results from the FPXRF analyzers to EPA SW-846 Methods 3050A/6010A, the most widely used approved methods for determining metal concentrations in soil samples. The comparison of FPXRF and the reference methods had to take into account certain limitations of both methods, including matrix effects. For these reasons, qualified reference data were used for statistical analysis.

The QC review and QA audit of the reference data indicated more than 98 percent of the data either met the demonstration QAPP objectives or was QC coded for reasons not limiting its use in the data evaluation. Less than 2 percent of the data were rejected based on QAPP criteria. Rejected data were not used for statistical analysis. The reference data were considered as good as or better than other laboratory analyses of samples performed using the same extraction and analytical methods. The reference data met the definitive data quality criteria and was of sufficient quality to support regulatory activities. The reference data were found to be acceptable for comparative purposes with the FPXRF data.

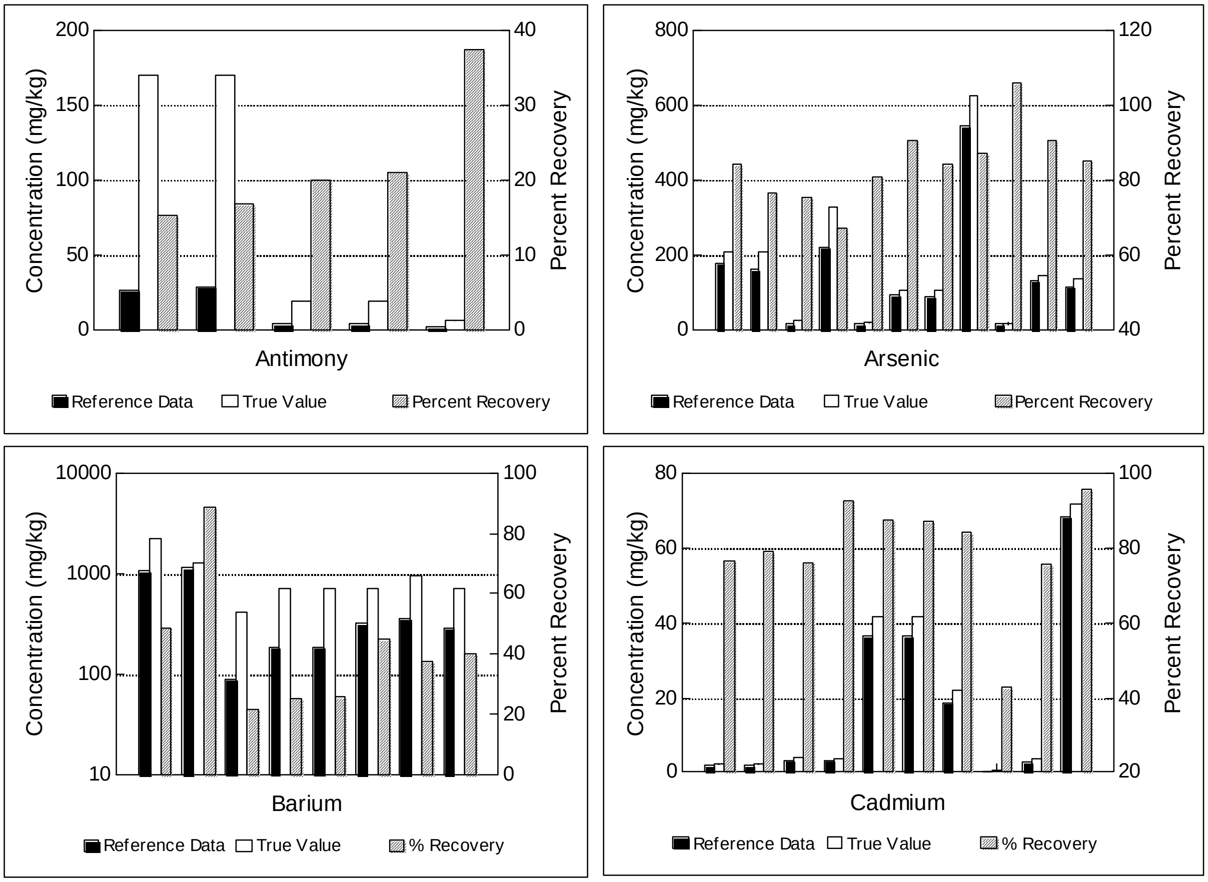


Figure 3-3. Reference Method SRM Results: These graphs illustrate the relationship between the reference data and the true values for the SRM samples. The gray bars represent the percent recovery for the reference data. Each set of three bars (black, white, and gray) represents a single SRM sample.

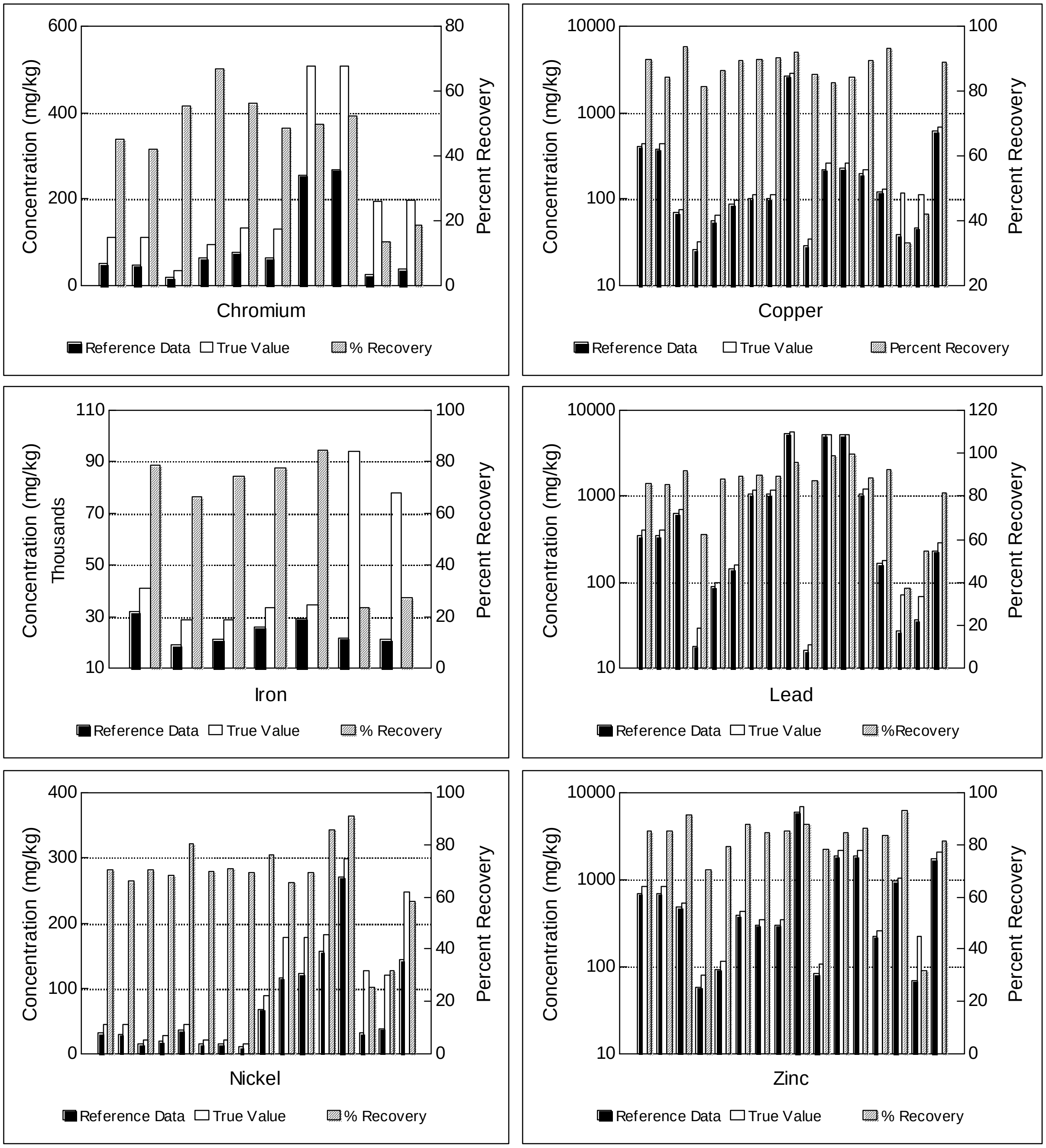


Figure 3-3 (Continued). Reference Method SRM Results: These graphs illustrate the relationship between the reference data and the true values for the SRM samples. The gray bars represent the percent recovery for the reference data. Each set of three bars (black, white, and gray) represents a single SRM sample.

Section 4
XL Spectrum Analyzer

This section provides information on the Niton's XL Spectrum Analyzer including the theory of FPXRF, operational characteristics, performance factors, a data quality assessment, and a comparison of results with those of the reference laboratory.

Theory of FPXRF Analysis

FPXRF analyzers operate on the principle of energy dispersive XRF spectrometry. This is a nondestructive qualitative and quantitative analytical technique that can be used to determine the metals composition in a test sample. By exposing a sample to an X-ray source having an excitation energy close to, but greater than, the binding energy of the inner shell electrons of the target element, electrons are displaced. The electron vacancies that result are filled by electrons cascading in from an outer shell. Electrons in these outer shells have higher potential energy than inner shell electrons, and to fill the vacancies, the outer shell electrons give off energy as they cascade into inner shell (Figure 4-1). This release of energy results in an emission of X-rays that is characteristic of each element. This emission of X-rays is termed XRF.

Because each element has a unique electron shell configuration, each will emit unique X-rays at fixed wavelengths called "characteristic" X-rays. The energy of the X-ray is measured in electron volts (eV). By measuring the peak energies of X-rays emitted by a sample, it is possible to identify and quantify the elemental composition of a sample. A qualitative analysis of the sample can be made by identifying the characteristic X-rays produced by the sample. The intensity of each characteristic X-ray is proportional to the concentration of the source and can be used to quantitate each element.

Three electron shells are generally involved in the emission of characteristic X-rays during FPXRF analysis: the K, L, and M shells. A typical emission pattern, also called an emission spectrum, for a given element has multiple peaks generated from the emission X-rays by the K, L, or M shell electrons. The most commonly measured X-ray emissions are from the K and L shells; only elements with an atomic number of 58 (cerium) or greater have measurable M shell emissions.

Each characteristic X-ray peak or line is defined with the letter K, L, or M, which signifies which shell had the original vacancy and by a subscript alpha (α) or beta (ß), which indicates the next outermost shell from which electrons fell to fill the vacancy and produce the X-ray. For example, a K_α-line is produced by a vacancy in the K shell filled by an L shell electron, whereas a K_β-line is produced by a vacancy in the K shell filled by an M shell electron. The K_α transition is between 7 and 10 times more probable than the K_β transition. The K_α-line is approximately 10 times more intense than the K_β-line for a given element, making the K_α-line analysis the preferred choice for quantitation purposes.

For a given element, the X-rays emitted from L transitions are always less energetic than those emitted from K transitions. Unlike the K-lines, the L-lines (L_α and L_β) for an analyte are of nearly equal intensity. The choice of which one to use for analysis depends on the presence of interfering lines from other analytes.

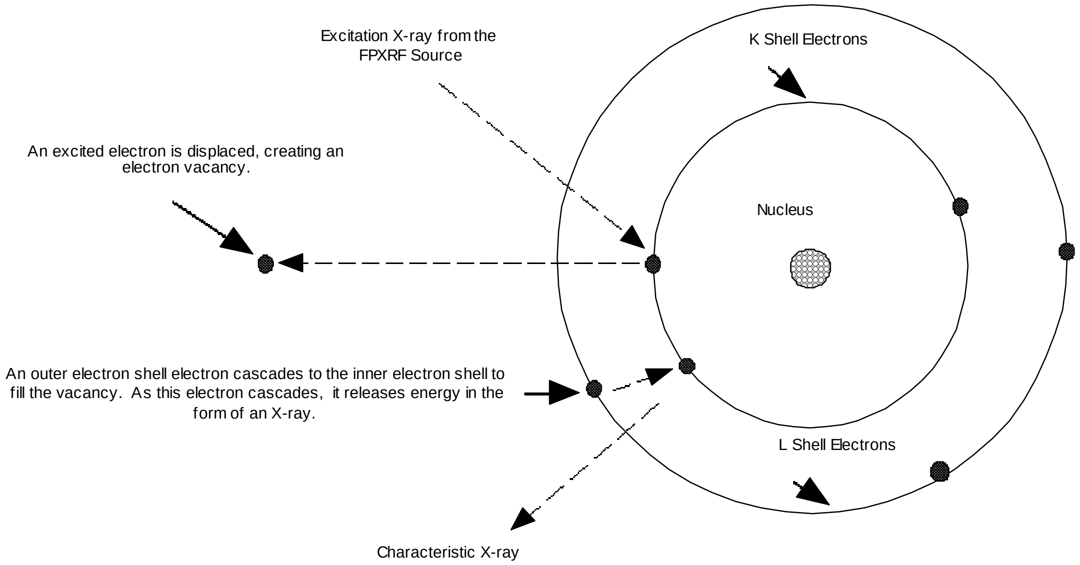


Figure 4-1. Principle of Source Excited X-ray Fluorescence: This figure illustrates the dynamics of source excited X-ray fluorescence.

An X-ray source can excite characteristic X-rays from an analyte only if its energy is greater than the electron binding energies of the target analyte. The electron binding energy, also known as the absorption edge energy, represents the amount of energy an electron has to absorb before it is displaced. The absorption edge energy is somewhat greater than the corresponding line energy. Actually, the K-absorption edge energy is approximately the sum of the K-, L-, and M-line energies of the particular element, and the L- absorption edge energy is approximately the sum of the L- and M-line energies. FPXRF analytical methods are more sensitive to analytes with absorption edge energies close to, but less than, the excitation energy of the source. For example, when using a Cd^{109} source, which has an excitation energy of 22.1 kilo electron volts (keV), an FPXRF analyzer would be more sensitive to zirconium, which has a K-line absorption edge energy of 15.7 keV, than to chromium, which has a K-line absorption edge energy of 5.41 keV.

Background

Niton was given two grants by the EPA through the Small Business Innovative Research Program to develop a lead detector that would be inexpensive, portable, safe, and easy to use. Niton evaluated a number of radioactive sources and chose to use Cd^{109} in the XL Spectrum Analyzer. It found the advantages of the Cd^{109} source over others such as Co^{57} to be lower cost, longer half life, less background interference, and greater safety. Niton developed the XL Spectrum Analyzer to be a hand-held, *in situ*, lead-in-paint instrument. However, it has recently modified the instrument to allow it to be used to determine lead and other metals in soil in both the *in situ* and intrusive mode.

The XL Spectrum Analyzer uses energy dispersive XRF spectrometry to determine the elemental composition of soils and other solid waste materials. The XL Spectrum Analyzer can identify and quantify the concentrations of the elements molybdenum, zirconium, strontium, rubidium, arsenic, zinc, copper, lead, nickel, iron, and chromium. The developer expected iron interference to result in quantitation difficulties for nickel and chromium. Niton is currently considering the development of an iron filter that will block out the iron fluorescence to obtain more accurate readings for nickel and chromium. Because the instrument uses a Cd^{109} source, it cannot analyze for elements such as barium, tin, antimony, silver, and cadmium. The XL Spectrum Analyzer uses a silicon pin-diode detector that achieves a manganese $K\text{-}_{\alpha}$ X-ray resolution of 800 eV (0.80 keV) while operating near ambient temperature. The detector is thermo-electrically cooled using a Peltier effect accessory.

For *in situ* analyses, Niton has developed a metal skid that acts as a protector for the XL Spectrum Analyzer and allows the source-detector window to come into direct contact with the soil surface. For intrusive analyses, a different metal skid is used to hold the XRF sample cup in position during analysis. In either mode, the sample is positioned in front of the source-detector window and sample measurement is initiated by depressing a plunger on the backside of the instrument. This exposes the sample to primary radiation from the Cd^{109} source. Fluorescent and back-scattered radiation reenters the analyzer through the source-detector window and is counted by the silicon pin-diode detector.

During this demonstration, the XL Spectrum Analyzer was operated using a hand-held computer attached to the RS-232 port of the instrument. The computer used a data acquisition and reduction program developed by Niton to record and report multiple element concentrations. The “SOILAIR” program automatically calibrated the XL Spectrum Analyzer at the start of an analysis. This analyzer uses the Compton ratio method to quantitate metals concentrations in samples.

The Compton normalization method for calibration and quantitation is based on the analysis of a single, certified standard and normalization to the Compton peak. The Compton peak is produced from incoherent back-scattering of X-ray radiation for the excitation source and is present in the spectrum of every sample. The Compton peak intensity changes with differing matrices. Generally, matrices dominated by lighter elements produce a larger Compton peak, and those dominated by heavier elements produce a smaller Compton peak. Normalizing to the Compton peak can reduce problems with varying matrix effects among samples. Compton normalization is similar to the use of internal standards in organic analysis.

The certified standard used for this type of calibration and quantitation usually is a NIST SRM, such as 2710 or 2711. The SRM must be in a matrix similar to the test samples and must contain the analytes of interest at concentrations near those expected in the test samples. First, a response factor is determined for each analyte. This factor is calculated by dividing the net peak intensity by the analyte concentration. The net peak intensity is a gross intensity corrected for baseline interference. The concentrations of analytes in samples are then determined by multiplying the baseline corrected analyte signal intensity by the normalization factor and by the response factor. The normalization factor is the quotient of the baseline corrected Compton K_{α} peak intensity of the SRM divided by that of the samples. These calculations are done by the instrument software.

Operational Characteristics

This section discusses equipment and accessories, operation of the analyzer in the field, background of the operator, training requirements, reliability of the analyzer, health and safety concerns, and representative operation costs.

Equipment and Accessories

The XL Spectrum Analyzer comes with all of the accessories necessary for *in situ* and intrusive operation (Table 4-1). A waterproof, unbreakable, plastic carrying case is provided for transportation and storage.

Table 4-1. Analyzer Instrument Specifications

Characteristic	Specification
Resolution	800 eV (Manganese-K_α)
Source	10 millicuries Cd^{109}
Detector	Silicon pin-diode—Peltier cooled
Analyzer Size	4.76 cm x 7.62 cm x 20.95 cm
Analyzer Weight (including batteries)	1.13 kilograms
Probe Operating Temperature	5 to 41 °C
Analyzer's Storage Capacity	500 sets of numerical results and 500 spectra
Power Source	120V AC or Internal Batteries
Operational Checks	2 NIST SRMs, silicon dioxide and Teflon® blanks, pure element check sample kit
Intrusive Operation	Sample and analyzer mount
Computer Interface Operation	RS 232 serial input/output cable, operators manual, application and results software, and training video
Contact: Stephen Shefsky Niton Corporation 74 Loomis Street PO Box 368 Bedford, MA 01730 (617) 275-9275 (617) 275-2397 (FAX)	

The XL Spectrum Analyzer uses Cd^{109} as a sample excitation source. This source has an initial strength of 10 millicuries (mCi). The source exposes the sample to excitation radiation through a 1 by 2 cm window on the backside of the instrument. The X-ray-induced fluorescence from the sample passes back through the window and is intercepted by the silicon pin-diode detector. The detector measures the energy of each X-ray and builds a spectrum of element peaks on a 1,024 multichannel analyzer (MCA), with up to 100 channels visible on the analyzer's liquid crystal display. A spectrum contains the peak lines for all the source-detectable metals present in the sample.

The XL Spectrum Analyzer displays a spectrum with each sample measured for visual identification of the elements present in the sample. Three buttons on the face of the instrument allow the operator to view the full spectrum after each reading or scroll back to previous readings. The XL Spectrum Analyzer contains an audible time signal that beeps at 5 and 20 seconds to assist the operator when timing measurements. The instrument has the capacity to store 500 readings (including spectra). The instrument contains an RS-232 port for downloading the data to a computer. The instrument identifies the date, time, temperature, humidity, test number, and spectra for each measurement recorded.

The developer provided two battery packs and a battery charger with the instrument. The battery packs consisted of eight wrapped nickel metal hydride batteries. Each battery pack was capable of lasting 8 hours. Two battery packs were provided so that one pack could always be charging. The battery charger came with accessories so that it could recharge from a 110-volt electrical outlet or from a car cigarette lighter. A full battery recharge could be accomplished in 3 hours.

Along with the plastic carrying case, the XL Spectrum Analyzer also came with a lightweight canvas waist pack to assist when carrying the instrument in the field. Both the waist pack and the plastic carrying case contained a protective lead plate to shield the operator from the radioactive source should it be damaged or leaking.

Two metal soil testing mounts or skids were provided as separate components to operate the XL Spectrum Analyzer when conducting soil analysis. One skid was designed for *in situ* sampling and functioned as a barrier between the instrument and soil. This skid contained a 1 cm by 2 cm opening that allowed the Cd^{109} source X-rays to penetrate the sample media. A Mylar™ film (0.25 micrometers) was placed over this opening to prevent soil particles from touching the instrument shutter. The second metal skid was used for intrusive sampling and was designed to hold an XRF sample cup against the instrument during use.

Operation of the Analyzer

To obtain numerical results for elements other than lead, the analyzer required an external data processor. A hand-held computer was provided for ease of portability in the field during *in situ* sample analysis. A laptop computer was provided to serve the same purpose while operating in the intrusive mode. A computer interface kit was included with the palm-top computer to allow data transfer from the palm-top to the laptop computer. Once the data was transferred to the laptop, it could then be saved on disk for permanent storage.

Both of the computers used during the demonstration contained a data acquisition and reduction software program “SOILAIR.” This program enabled the computer to read the raw data and calculate concentrations in parts per million for each metal that was detected. Both computers contained a program to print the data collected. The computers also contained a program named “XL” that allowed the XL Spectrum Analyzer to perform a quick data transfer to another computer where the data could be saved on a computer disk for later use. The operator could then zero the instrument memory and begin new readings without losing any data. The instrument came with a computer cable (RS-232) to connect the instrument to a laptop computer.

The instrument requires a 15-minute warmup prior to operation. The unit was initially calibrated by the developer; this calibration must be monitored through the analysis of check samples. The developer recommends the analysis of calibration check samples, such as NIST 2710. The developer’s calibration also accounts for temperature and other environmental conditions such as humidity. Due to the cooling efficiency of detector assembly, instrument temperatures exceeding 80 °F could reduce the detector's resolution. During data acquisition, the developer recommended that no measurement less than 3 times greater than its associated standard deviation be considered a useable result.

The operator for this demonstration had no prior experience operating an FPXRF instrument and found the XL Spectrum Analyzer easy to operate. The operator felt the instrument was conducive for field use because of its small size and light weight. She noted that since the instrument was automatically calibrated, this eliminated operator error in the calibration process and allowed more time for sample analysis. Since the XL Spectrum Analyzer is capable of holding 500 measurements, which is more than can be collected in one day, it was not necessary to interrupt daily activities to download data. The operator noted that with the proper training from Niton, it was easy to download data from the instrument. The instrument does not allow additional readings to be taken if the battery is low, so no data loss occurs due to a low battery.

The operator was required to kneel on the ground or place a brick on the instrument while conducting *in situ* analyses to keep the plunger depressed (which kept the shutter open) and to keep the instrument from moving while collecting a reading. The operator recommended that Niton redesign the soil testing mount so that it would keep the plunger depressed and hold the instrument firmly in place while conducting *in situ* analyses.

Background of the Technology Operator

The operator chosen for analyzing soil samples using the Niton XL Spectrum Analyzer has been an employee of PRC for 6 years. She holds a bachelor's degree in natural resource management and physical science and a master's degree in public administration with an emphasis in environmental policy. She has performed soil and water sampling at hazardous waste sites for more than 5 years while employed at PRC.

Training

The operator received two separate phases of training on the XL Spectrum Analyzer. Because this instrument requires a specific license, the operator was required by Niton to attend a radiation safety course prior to use of the instrument. This training course was taken 2 months prior to the field demonstration so that all the necessary paperwork could be completed and approved to obtain the specific license. This first training course also discussed the operation of the analyzer to determine lead concentrations in paint.

Prior to beginning the field demonstration, the operator received approximately 4 hours of training from Niton on the XL Spectrum Analyzer. A Niton representative observed field use of the instrument during the first 4 days of the demonstration providing additional instructions and suggestions. The training focused on the instrument components, calibration, and operation.

Reliability

During the 20 days of the demonstration, the calibration monitoring never exhibited characteristics indicative of accuracy drift requiring recalibration. This monitoring involved a daily measurement of a NIST SRM.

During the demonstration, two areas of concern were noted by the operator. The first was the fact that the XL Spectrum Analyzer was designed for indoor use and extra care was required to ensure unfavorable weather did not harm the instrument. Because there was frequent light to moderate rain at one site, the operator expended a considerable amount of time waterproofing the instrument by covering it with plastic bags and using an umbrella in the field to provide additional rain protection. Though waterproofing was attempted, on one occasion water did enter the instrument causing water vapor to form on the source-detector window inside the instrument. The window was replaced and measurements resumed. This replacement took less than 10 minutes.

The second difficulty dealt with the portable hand-held computer. The computer cord that attached the palm-top computer to the XL Spectrum Analyzer would often drag on the ground and was easily unhooked when jarred. This caused data to be lost, and it required samples to be reanalyzed. Another problem with the hand-held computer was a feature that caused it to automatically shut off after 60 seconds of idle time. This feature caused a loss of data, on occasion requiring sample reanalysis in the field, because the computer was off when the readings were being collected. The Niton representative explained that the computer hookup was a temporary measure and that ultimately the software and computer would be incorporated into the analyzer.

Health and Safety

The potential for exposure to radiation from the excitation sources is the largest health and safety consideration while using an FPXRF instrument. Radiation was monitored with a radiation survey meter using a pancake probe. Background radiation at the two sites was between 0.006 and 0.012 millirems per hour (mrem/hr). Radiation exposure was monitored in the *in situ* and intrusive mode while the shutters of the instruments were open to obtain a worst-case scenario measurement. The radiation was measured within 5 cm of the shutter while the instrument was analyzing a sample. Radiation exposure also was monitored at a point on the instrument where the operator's hand was located during analysis. This provided a realistic value of potential operator exposure. For example, in the State of Kansas, the permissible occupational exposure is 5,000 mrem/year, which equates to approximately 2 to 3 mrem/hr assuming constant exposure for an entire work year.

While taking *in situ* measurements, the following radiation values were obtained for the XL Spectrum Analyzer with the Cd^{109} source exposed: 0.20 - 0.30 mrem/hr at the shutter, 0.07 to 0.08 mrem/hr at the front of the instrument, 0.02 to 0.03 mrem/hr at the side of the instrument, and 0.02 to 0.03 mrem/hr on top of the instrument where the operator's hand was placed. While collecting intrusive measurements with the XL Spectrum Analyzer, the following radiation values were obtained with the Cd^{109} source exposed: 0.13 to 0.15 mrem/hr at the shutter, 0.03 to 0.04 mrem/hr on top of the instrument where the operator's hand was placed, and background levels under the wooden table where the operator was taking measurements. The source-detector window is pointing down during intrusive measurement. All measured radiation values were less than the permissible 2 to 3 mrem/hr. Although the radiation readings underneath the wooden table were at background levels, it is a safe practice not to sit at the table with one's legs under the instrument while taking measurements. The operator noted that the safety features on the instrument made it difficult to cause an accidental exposure while using the instrument.

Cost

At the time of the demonstration, the XL Spectrum Analyzer with its standard software package cost $11,990 to purchase. This includes all of the equipment necessary for operation of the instrument. An extra battery pack costs $300, while a wrist support is $15. Shipping, handling, and insurance costs $80. An extended 15-month warranty can be purchased for $1,200. Periodic maintenance includes replacement of the Cd^{109} source every 2 years at a cost of $2,200, which includes old source disposal, a leak test, and a certificate of ownership. The Cd^{109} source replacement, and routine maintenance, is another available option that costs $2,600 every 2 years. The XL Spectrum Analyzer also can be rented for $2,200 per month.

The XL Spectrum Analyzer requires a specific radiation license for use. This requires attending a special radiation safety course and completing all the necessary paperwork to obtain the specific radiation license. The radiation safety and operator's training course costs $350 per person, and travel expenses. A specific license for the XL Spectrum Analyzer was required for this demonstration; it cost $500 to obtain the license for ownership and operation of a sealed radioactive source through the State of Kansas. Since the two demonstration sites were in Washington and Iowa, reciprocal agreements were required from both states to operate the instrument in those states. These reciprocal agreements cost $585 for Washington and $700 for Iowa. Operator costs will vary depending on the technical knowledge of the operator. Niton claims the XL Spectrum Analyzer can be used by individuals with no more than a high school education and a minimal amount of technical training, thereby decreasing the cost.

The primary cost benefit of field analysis is the quick access to analytical data. This allows the process dependent on the testing to move efficiently onto the next stage. Costs associated with field analysis are dependent on the scope of the project. Since most of the mobilization costs are fixed, analyzing a large number of samples lowers the per sample cost. This is a key advantage that field analysis has over a conventional laboratory. Furthermore, more samples are usually taken for field analysis since questions raised in the preliminary findings may be resolved completely without the need to return for another sample collection event.

A representative list of costs associated with the Niton XL is presented in Table 4-2. Also included in this table is the measured throughput and the per sample charge of the reference laboratory. Given the special requirements of this demonstration, it was not considered fair to report a per sample cost for the field analysis. However, some estimate can be derived from the data provided in this table.

Table 4-2. Instrument and Field Operation Costs

Item	Amount	
Niton XL Spectrum Analyzer	$ 11,990	Purchase Price
	2,200	Per Month Lease
Replacement Source	2,200	2 Year Lifetime Warranty
Operator Training (Vendor Provided)	350	—
Radiation Safety License (State of Kansas)	500	—
Field Operation Costs		
Supplies and Consumables (Sample cups, window film, sieves, standards)	300 - 500	(Varies, depending on sample load)
Field Chemist (Labor Charge)	100 - 150	Per day
Per diem	80 - 120	Per day
Travel	200 - 500	Per traveler
Sample Throughput	20 - 25	Samples per hour
Cost of Reference Laboratory Analysis	150	Per sample

Performance Factors

The following paragraphs describe performance factors, including detection limits, sample throughput, and drift.

Detection Limits

MDLs, using SW-846 protocols, were determined by collecting 10 replicate measurements on site-specific soil samples with metals concentrations 2 to 5 times the expected MDLs. These data were obtained during the precision evaluation. Based on these findings, a standard deviation was calculated and the MDLs were defined as 3 times the SD for each target analyte. All the precision-based MDLs were calculated for soil samples that had been dried, ground, and placed in a sample cup, the highest degree of sample preparation. The precision-based MDLs for the XL Spectrum Analyzer are shown in Table 4-3.

Table 4-3. Method Detection Limits

Analyte	Precision-based MDL (mg/kg)	Field-based MDL (mg/kg)
Arsenic	120	320
Chromium	900	1,370
Copper	130	365
Lead	75	135
Zinc	115	240

Notes:

ND	Not determined. The sample's nickel concentrations were not reported as positive values.
mg/kg	Milligrams per kilogram.

Based on the demonstration data, most of these MDLs seem reasonable except chromium. None of the reported chromium data below 1,000 mg/kg met the developer's data acceptance criteria. The developer recommends that no measurement less than 3 times its SD should be considered as a quantifiable concentration.

Another method of determining MDLs involved the direct comparison of the FPXRF data and the reference data. When these data sets are plotted against each other, the resultant plots were linear. This method is discussed in greater detail in the "Intermethod Assessment" later in this section. As the plotted line approached zero for either method, there was a point at which the FPXRF data intersects at a reading for a concentration of the reference data. Figure 4-2 illustrates this effect for arsenic. This point was determined by observation and is somewhat subjective; however, a sensitivity analysis showed that even a 25 percent error in identifying this point resulted in intersects only a 10 percent change in the MDL calculation. By determining the mean values of this FPXRF data and subsequently two SDs around this mean, it was possible to determine a field or performance-based MDL for the analyzer. The XL Spectrum Analyzer field-based MDLs are shown in Table 4-3. These field-based MDLs are greater than the precision-based MDLs.

During the comparability study, several matrix interferences were observed. These interferences had a direct effect on MDL estimates. These interferences produced a field method detection limit of approximately 1,400 mg/kg for arsenic, when lead concentrations were 10 times or greater than the corresponding arsenic concentrations. For samples where the arsenic and lead ratios were closer to 1:1 or arsenic was more abundant than lead, the field-based MDL was 320 mg/kg. Both of these MDLs are greater than the 120 mg/kg MDL based on the precision data. A second interference involved a copper and zinc. Analyzer accuracy for zinc quantitation was greatly reduced in the presence of copper concentrations in the tens of thousands of milligrams per kilogram.

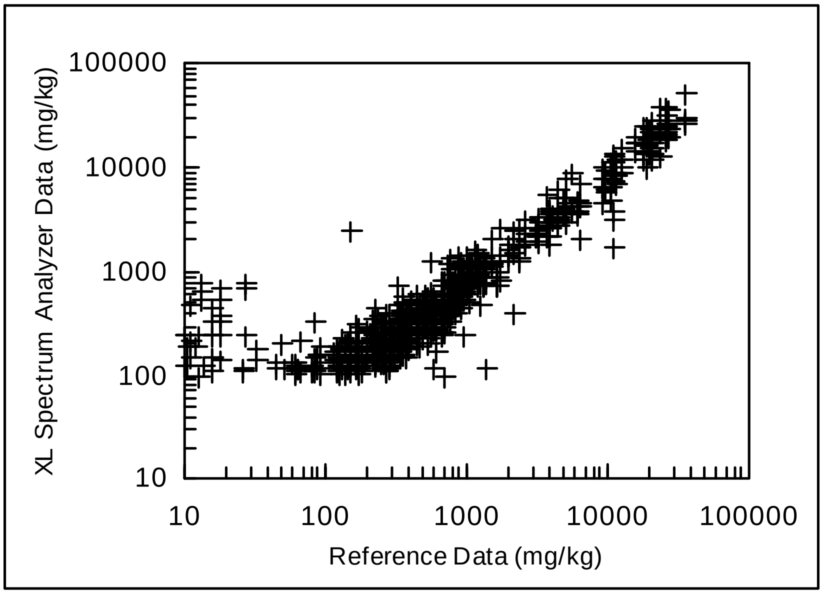


Figure 4-2. Critical Zone for the Determination of a Field-based Method Detection Limit for Arsenic: Between 100 and 200 mg/kg for the reference data the linear relationship between the two data sets changes. This point of change identified the point at which field-based MDLs for the analyzer were determined.

Throughput

The Niton XL Spectrum Analyzer used a Cd^{109} source live count time of 60 seconds. With the additional "dead" time of the detector and the time required to label each sample and store data in between sample measurements, the time required to analyze one soil sample was between 2 and 2.5 minutes. When collecting *in situ* measurements in the field, the throughput was approximately 20 samples per hour. One day at the ASARCO site, it rained frequently, which caused the operator to take extra precautions to waterproof the instrument. On this day, the sample throughput was 11 samples per hour. The throughput for the intrusive measurements rose to 25 samples per hour. The minimum number of samples analyzed in a 10-hour day was 110 samples. This was for *in situ* measurements in the field at the ASARCO site when the operator was slowed due to the frequent rains. The maximum number of samples analyzed in a 10-hour day was 250 samples for intrusive measurements at the RV Hopkins site.

This throughput included an average of 20 QC samples, such as spikes, blanks, and calibration check, that were analyzed daily. The sample analysis time did not include the time required for sample handling and preparation or for data downloading, printing, and documentation. Considerable time was spent preparing the *in situ* homogenized samples and the intrusive samples. Homogenization required approximately 5 minutes per sample (*in situ*-prepared), 20 minutes per sample were required for No. 10 sieving (intrusive-prepared), and 10 minutes per sample were required for grinding and sieving (intrusive-prepared). Approximately 1 hour was spent daily downloading the data to a PC and obtaining a hard copy of the data.

Drift

Drift is a measurement of an analyzer's variability in quantitating a known amount of a standard over time. Drift was evaluated by reviewing results from the periodic analysis of the calibration check sample. No developer claims were made concerning drift.

For the XL Spectrum Analyzer, drift was evaluated by reviewing results from the analysis of NIST SRM 2710. This sample contained quantifiable levels of arsenic, copper, lead, zinc, and iron. The developer recommended that the operator run this SRM as a calibration check one time per 20 samples analyzed. NIST SRM 2710 was analyzed approximately 100 times during the almost 1,300 measurements taken during this demonstration. This data was reduced to RSDs for the target analytes and the percent drift from the true value, or from 100 percent recovery (Figure 4-3). This figure compiles the results from the first analysis of NIST SRM 2710 run each day that measurements were taken by the XL Spectrum Analyzer. The RSD values for all analytes were less than 8 percent. The mean percent recoveries depicted in the figure were between 82 and 137 percent. The high percent recovery for iron and zinc indicated the instrument was biased high for these analytes. Given the low RSD values and percent recoveries near 100 percent would indicate that for the other analytes found in NIST SRM 2710, the XL Spectrum Analyzer showed little drift during the demonstration.

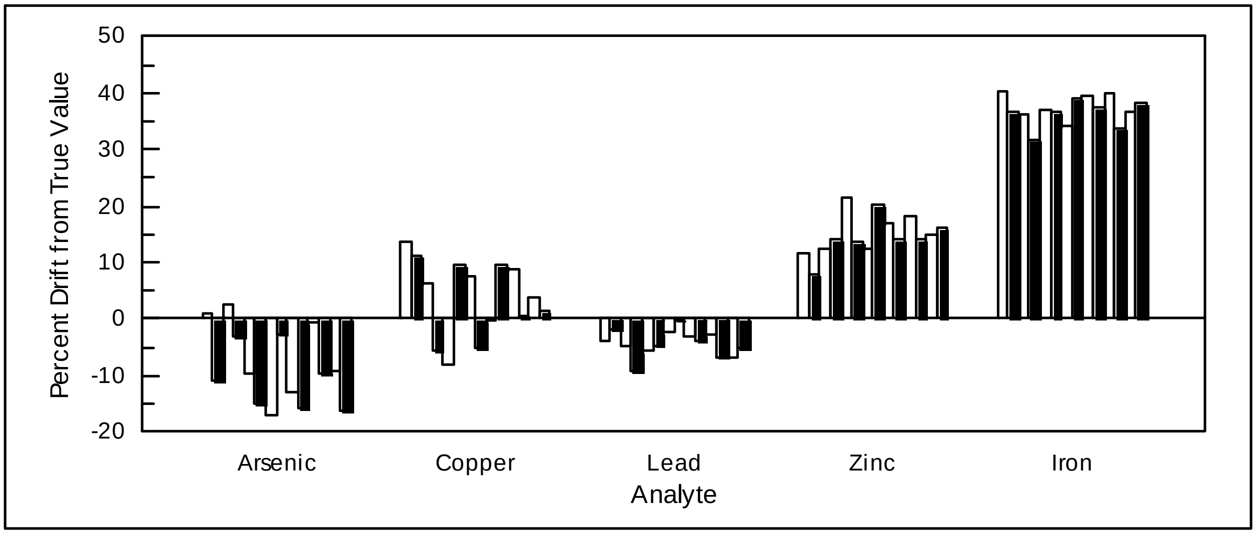


Figure 4-3. Drift Summary: This figure shows the general bias of the analyzer's results in measuring NIST SRM 2710. Each bar represents a different day's analysis of the same sample. The daily fluctuations exhibited for each analyte is a direct representation of drift.

Intramethod Assessment

Intramethod assessment measures of the analyzer's performance include results on instrument blanks, completeness of the data set, intramethod precision, and intramethod accuracy. The following paragraphs discuss these characteristics.

Blanks

NIST SRM 2709 was used as an instrument blank for the XL Spectrum Analyzer on the days when *in situ* field measurements were collected. This SRM contains concentrations of the target analytes below the MDLs for the XL Spectrum Analyzer. During the remainder of the demonstration, a lithium carbonate sample was used as an instrument blank. At the beginning of the demonstration, the blanks were analyzed at a frequency of one per every 10 samples. After the first 2 days of the demonstration, the frequency of blank analysis was changed to one per every 20 samples. This was done at the developer's request. The blanks were used to monitor for contamination of the probe from sources such as residual soil left on the window of the instrument. More than 100 blanks were analyzed during the demonstration. Iron was the only target analyte detected in any of the blanks. It was always detected in NIST SRM 2709 because this SRM had a known iron concentration of 35,000 mg/kg. No iron was detected in the lithium carbonate blanks. The results of the blanks demonstrated there was no problem with cross-contamination from sample to sample or with contamination on the window of the instrument.

Completeness

For this demonstration, completeness refers to the proportion of valid, acceptable data generated. A total of 315 soil samples was analyzed four times (four preparation steps) resulting in 1,260 sample results. The XL Spectrum Analyzer produced results for 1,258 of the 1,260 samples for a completeness of 99.8 percent, above the demonstration target of 95 percent. The two missing results were for *in situ* measurements at the RV Hopkins site. In both cases, the operator analyzed the sample and a reading was collected. However, due to operator oversight, a hard copy of the reading was not printed and could not be retrieved from the computer. This high degree of completeness demonstrated the reliability and ruggedness of this instrument.

Precision

Precision was expressed in terms of the percent RSD between replicate measurements. The precision data for the target analytes detectable by the XL Spectrum Analyzer are shown in Table 4-4. The results reflected in the 5 to 10 times the MDL range reflects the instrument precision generally referred to in analytical methods, such as SW-846.

Table 4-4. Precision Summary

	Mean % RSD Values by Concentration Range			
Analyte	**5 - 10 Times MDL[a] (mg/kg)**	**50 - 500 (mg/kg)**	**500 - 1,000 (mg/kg)**	**>1,000 (mg/kg)**
Arsenic	9.2 (8)	23.4 (20)	9.3 (8)	3.1 (4)
Chromium	ND	ND	37.4 (8)	28.2 (4)
Copper	13.2 (4)	30.7 (28)	12.8 (8)	3.5 (12)
Lead	6.5 (4)	18.4 (20)	4.7 (4)	3.3 (20)
Zinc	11.2 (12)	18.1 (20)	11.5 (8)	ND

Notes:

[a]	The MDLs referred to in this column are the precision-based MDLs shown in Table 4-3.
mg/kg	Milligrams per kilogram.
ND	No data.
( )	Number of samples, including all four preparation methods, each consisting of 10 replicate analyses.

The XL Spectrum Analyzer performed 10 replicate measurements on 12 soil samples that had analyte concentrations ranging from less than 50 mg/kg to tens of thousands of milligrams per kilogram. Each of the 12 soil samples underwent four different preparation steps described previously. Therefore, there were 48 total precision samples analyzed by the XL Spectrum Analyzer. The replicate measurements were taken using the source count times discussed in the previous section of this report. For each detectable analyte in each precision sample, a mean concentration, SD, and RSD were calculated.

In this demonstration, the analyzer's precision RSD for a given analyte had to be less than or equal to 20 percent to be considered quantitative screening level data and less than or equal to 10 percent to be considered definitive level data. The analyzer's precision data, in the 5 to 10 times MDL range, were below the 10 percent RSD required for definitive level data quality classification for lead and arsenic. Copper and zinc had method precision RSDs greater than 10 percent, but less than 20 percent, placing the results into the quantitative screening level quality category. Table 4-4 shows that chromium precision was greater than 20 percent, placing the chromium results in the qualitative screening level data quality classification based on precision. The lower precision for chromium was expected because chromium is a problematic analyte for FPXRF analysis, especially at 60 live-second count times. Since no precision data was reported for chromium in the 5 to 10 times MDL range, no recommendation regarding a data quality level for chromium could be made.

There was no observable effect of sample preparation on precision. This was expected because the method used to assess precision during this demonstration was measuring analyzer precision, not total method precision. There was a concentration effect on the precision data, precision increased with increasing concentration. Figure 4-4 shows an asymptotic relationship between concentration and precision. In this figure, precision shows little improvement at concentrations greater than 500 mg/kg; however, at concentrations below 500 mg/kg, precision is highly concentration dependent. The precision samples were purposely chosen to span a large concentration range to test the effect of analyte concentration on precision.

Accuracy

Accuracy refers to the degree by which a measured value for a sample agrees with a reference or true value for the same sample. Accuracy was assessed for the XL Spectrum Analyzer by using site-specific PE samples and SRMs. Accuracy was evaluated through a comparison of percent recoveries for each primary and secondary target analyte reported by the XL Spectrum Analyzer. The XL Spectrum Analyzer analyzed six site-specific PE samples and 14 SRMs. The operator knew the samples were PE samples or SRMs, but did not know the true concentration or the acceptance range. These PE samples and SRMs were analyzed in the same way as all other samples.

The site-specific PE samples consisted of three samples from each of the two demonstration sites. These six PE samples were collected during the predemonstration activities and sent to six independent laboratories for analysis by laboratory grade XRF analyzers. The mean measurement for each analyte was used as the true value concentration. The 14 SRMs included 7 soil, 4 stream or river sediment, 2 ash, and 1 sludge SRM. The SRMs were obtained from NIST, USGS, Commission of European Communities, National Research Council-Canada, and the South African Bureau of Standards. The SRMs contained known certified concentrations of certain target analytes.

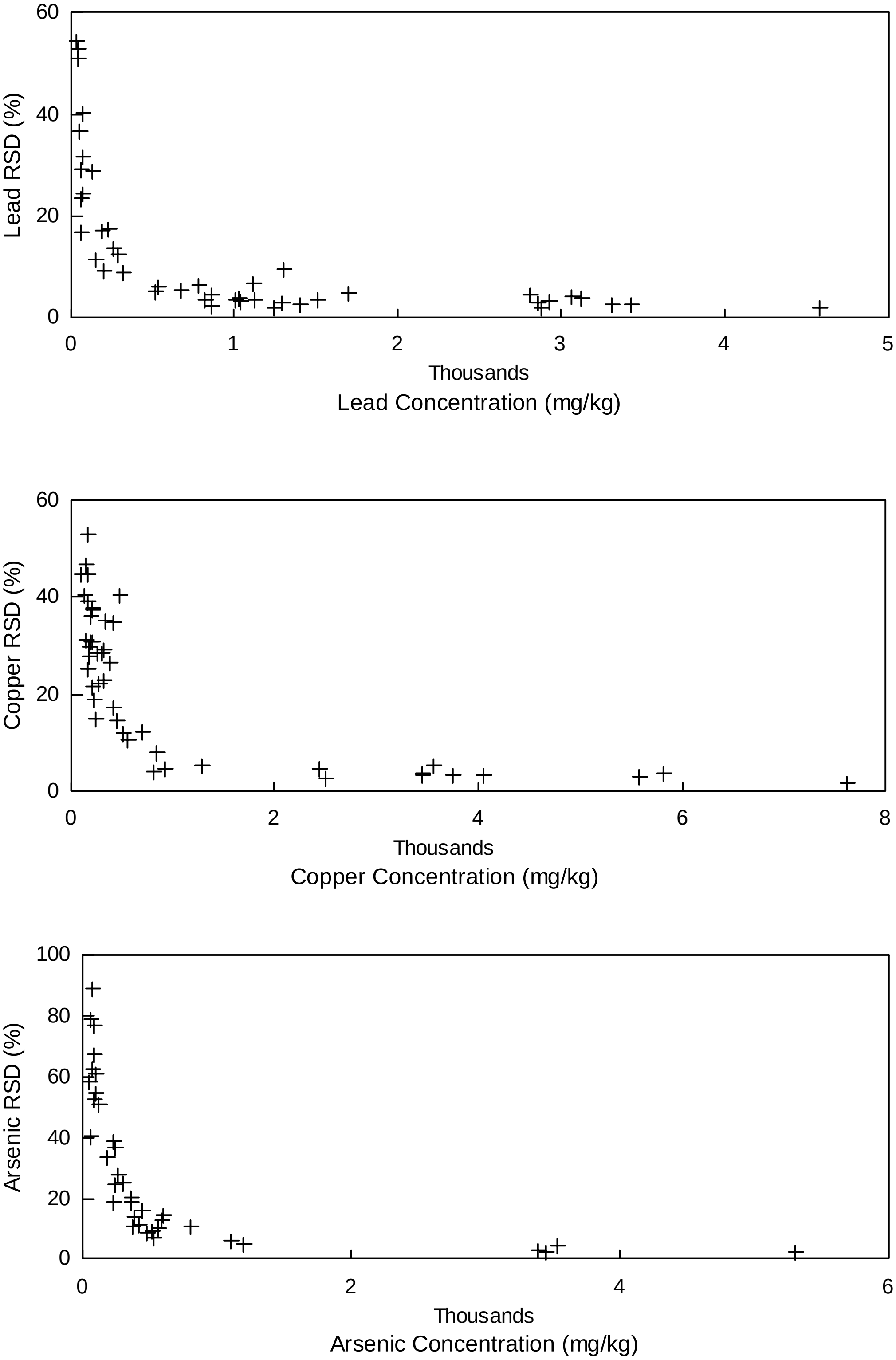


Figure 4-4. Precision vs. Concentration: This graph illustrates the analyzer's precision as a function of analyte concentration.

Site-specific PEs and SRMs did not have published acceptance ranges. As specified in the demonstration plan, an acceptance range of 80 - 120 percent recovery of the true value was used to evaluate accuracy for the six site-specific PEs and 14 SRMs. Table 4-5 summarizes the accuracy data for the primary and secondary target analytes for the XL Spectrum Analyzer. Figures 4-5 and 4-6 show the true value, the measured value, and percent recovery for the individual SRMs and site-specific PEs, respectively. No figure was presented for chromium because only two samples produced detectable concentrations of chromium by the XL Spectrum Analyzer. True value results from the site-specific PEs and SRMs with concentrations less than the precision-based MDLs listed in Table 4-4 also were excluded from the accuracy assessment.

Table 4-5. Accuracy Summary for Site-Specific PE and SRM Results

Analyte	n	Percent Within Acceptance Range	Mean Percent Recovery	Range of Percent Recovery	SD of Percent Recovery	Concentration Range (mg/kg)
Site-Specific Performance Evaluation Samples						
Arsenic	3	100	84	82 - 87	2.7	419 - 22,444
Chromium	2	0	119	79 - 158	NA	939 - 3,800
Copper	5	40	131	86 - 209	48	300 - 7,132
Iron	6	67	119	89 - 173	35	27,320 - 70,495
Lead	6	83	92	76 - 108	11	292 - 14,663
Zinc	6	67	117	89 - 156	23	164 - 4,205
Soil Standard Reference Materials						
Arsenic	2	0	159	159	NA	330 - 626
Copper	2	50	116	105 - 127	NA	131 - 2,950
Iron	3	67	112	97 - 137	22	28,900 - 35,000
Lead	5	80	100	90 - 126	14	101 - 5,532
Zinc	4	100	105	96 - 113	7.3	350 - 6,952
Sediment Standard Reference Materials						
Arsenic	1	0	535	535	NA	211
Copper	3	33	326	111 - 480	192	219 - 452
Iron	1	100	104	104	NA	41,100
Lead	4	100	89	80 - 94	6.5	161 - 5,200
Zinc	4	25	84	71 - 105	18	264 - 2,200
Ash & Sludge Standard Reference Materials						
Arsenic	2	0	176	168 - 184	NA	136 - 145
Copper	1	0	219	219	NA	696
Iron	2	50	118	115 - 122	NA	77,800 - 94,000
Lead	1	0	75	75	NA	286
Zinc	3	33	78	58 - 109	27	210 - 2,122

Notes:		
	n	Number of samples with detectable analytes.
	SD	Standard deviation.
	mg/kg	Milligrams per kilogram.
	NA	Not applicable, standard deviation not calculated for two or fewer results.

Based on the 80 - 120 percent recovery acceptance range, the XL Spectrum Analyzer's accuracy varied from 0 percent for chromium to 100 percent for arsenic in the site-specific PEs. Overall, the XL Spectrum Analyzer produced 19 out of 38 results or 64.3 percent within the 80 - 120 percent recovery acceptance range for all analytes in the six site-specific PE samples. Eight of the 10 results falling outside of the acceptance range were above the upper limit of 120 percent recovery. Table 4-5 also shows that the mean percent recoveries for four of the six analytes in the site-specific PEs were greater than 100 percent. This indicates that, in general, the XL Spectrum Analyzer was producing results that

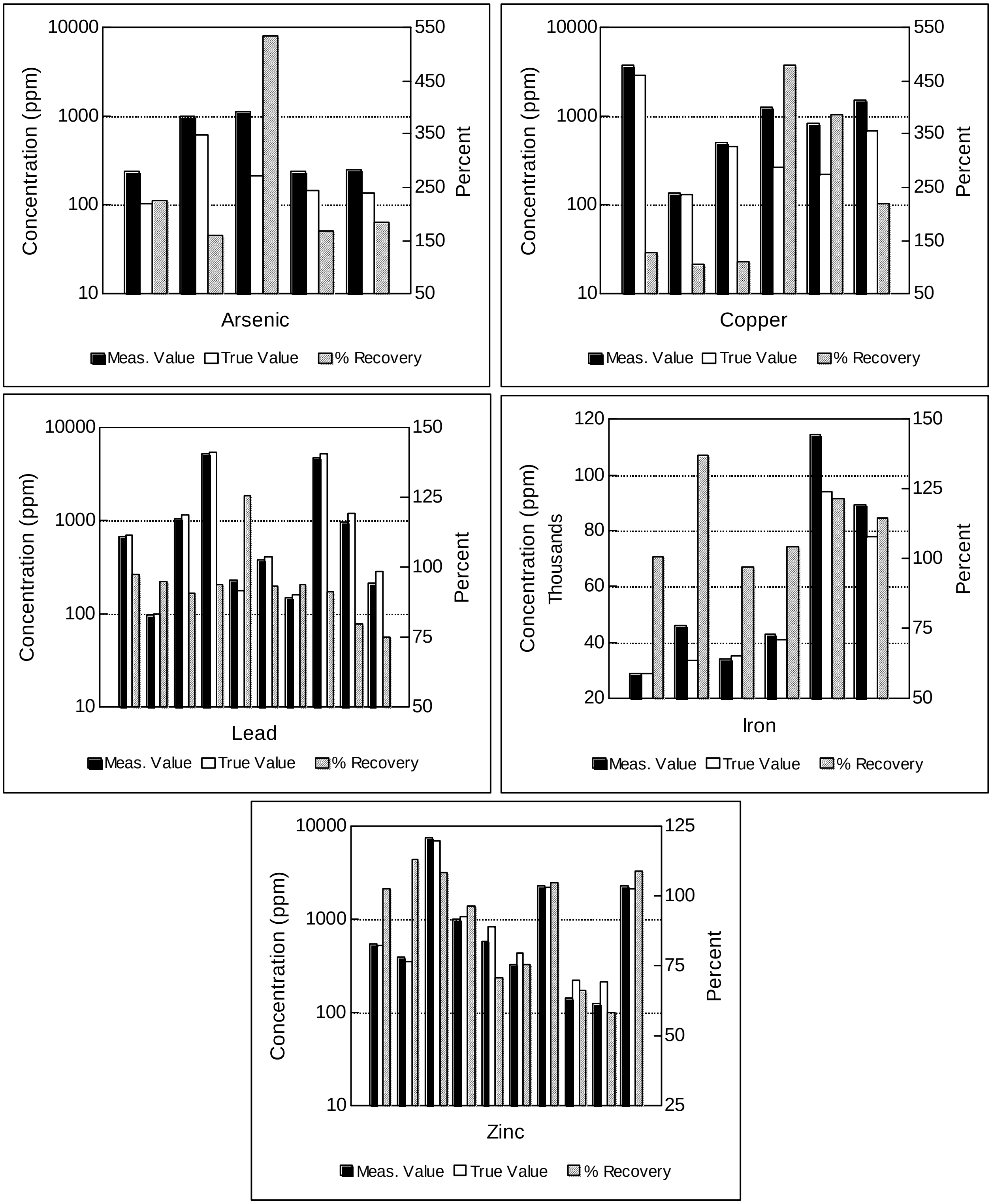


Figure 4-5. SRM Results: These graphs illustrate the relationship between the analyzer's data (measured values) and the true values for the SRMs. The gray bars represent the percent recovery for the analyzer. Each set of three bars (black, white, and gray) represents a single SRM sample.

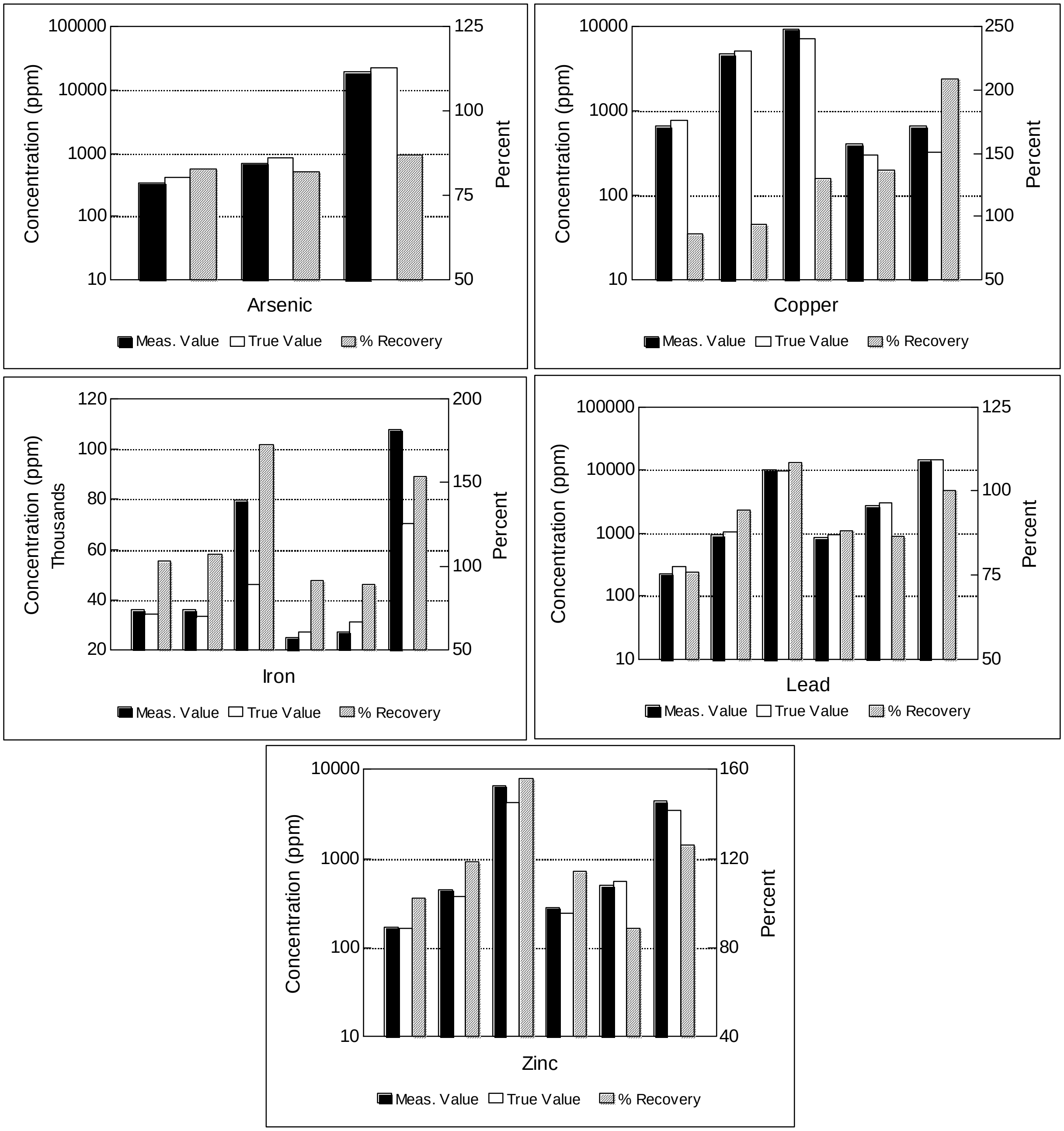


Figure 4-6. Site-specific PE Sample Results: These graphs illustrate the relationship between the analyzer's data (measured values) and the true values for the site-specific PE samples. The gray bars represent the percent recovery for the analyzer. Each set of three bars (black, white, and gray) represents a single site-specific PE sample.

were biased slightly high for chromium, copper, iron, and zinc. The XL Spectrum Analyzer performed well for lead in the site-specific PEs, with percent recoveries ranging from 76 to 108 percent. Its poorer performance for chromium was expected. The developer was not confident in the XL Spectrum Analyzer's ability to quantitate for chromium given the resolution of the detector and interference problems from high iron concentrations found in most of the soil samples used in this demonstration.

Seven of the eight results that were above the upper acceptance limit of 120 percent recovery in the site-specific PE samples occurred in the two high concentration PE samples from each site. These PE samples had multiple analytes at concentrations exceeding 3,000 mg/kg. The results of these PE samples may indicate an interelement interference problem not compensated for by the XL Spectrum Analyzer. It may be an inability for the detector to resolve all the analyte peaks at high concentrations, especially for analytes close together in the spectrum such as chromium and iron or copper and zinc. Lead and arsenic results in these two PE samples fell within the acceptance ranges.

Table 4-5 provides a summary of the accuracy data for the standard reference materials. Overall for all SRMs, the XL Spectrum Analyzer produced 20 out of 38 results within the 80 - 120 percent recovery acceptance range for an accuracy of 51.3 percent. Of the 19 results that fell outside of the acceptance range, 7 results were low and 12 were high. All but one of the low results were for lead and zinc, while all the high results were for arsenic, copper, and iron. This breakdown, in addition to the mean percent recoveries shown in Table 4-5, indicates that the XL Spectrum Analyzer was showing a high bias for arsenic, copper, and iron in all SRMs and a slightly low bias for lead and zinc in the sediment, ash, and sludge SRMs.

A more detailed analysis of the SRM data showed that there was a matrix effect on the XL Spectrum Analyzer's accuracy. The XL Spectrum Analyzer produced 10 out of 16 results or 62.5 percent within the acceptance range for all target analytes in the 7 soil SRMs; 7 out of 13 results or 53.8 percent within the acceptance range for all target analytes in the 4 sediment SRMs; and 2 out of 9 results or 22.2 percent within the acceptance range for all target analytes in the ash and sludge SRMs. This demonstrates that the XL Spectrum Analyzer is more accurate when analyzing SRMs of a soil matrix than sediment, sludge, or ash. This may indicate that the Compton ratio method (see "Background" subsection of this section) of calibration performs better for soil than other matrices.

In general, the XL Spectrum Analyzer displayed similar accuracy for the soil SRMs and the site-specific PEs. It was expected that the XL Spectrum Analyzer would be more accurate for the site-specific PE samples than for the SRMs for two reasons. First, the analytical technique (laboratory-grade XRF) used to determine the true analyte concentrations in the site-specific PEs was similar to the FPXRF technique. As described in Section 3, varying analytical techniques were used to determine the total analyte concentrations in the SRMs. Second, the analyte concentrations were often higher in the site-specific PEs versus the soil SRMs.

As would be expected, the overall XL Spectrum Analyzer accuracy was greatest for lead. It produced 13 out of 16 results or 81.2 percent within the acceptance range. The lowest percent recovery for lead was 75 percent and the highest percent recovery was 126 which is not much different from the 80 to 120 percent acceptance range. The accuracy was similar for copper, iron, and zinc in both the SRMs and PEs. The accuracy for arsenic was vastly different for the SRMs (0 percent) as compared to the PEs (100 percent). This is probably attributable to the much higher concentrations of arsenic in the PEs as compared to the SRMs.

Comparability

Intramethod comparability for the XL Spectrum Analyzer was assessed through the analysis of four ERA PEs and four CRM PEs. This was done to present users with additional information on data comparability relative to different commercially available QC samples. The eight PEs were analyzed in the same way as all other samples. As described in Section 3, these eight PEs had certified analyte values determined by EPA SW-846 Methods 3050A/6010A. Therefore, since these methods do not necessarily determine total metals concentrations in a soil, it was expected that the FPXRF would tend to overestimate analyte concentrations relative to PALs. The ability of the XL Spectrum Analyzer to produce results within the PALs or prediction intervals (PI) and the percent recovery for each of the analytes was used to evaluate the XL Spectrum Analyzer's intramethod comparability. True value analyte concentrations in the ERA and CRM PEs that were below the precision-based MDLs listed in Table 4-4 were excluded from the intramethod comparability assessment. The value "n" in Table 4-6 gives an indication of how many of the four ERA PEs and four CRM PEs actually had analyte concentrations above the precision-based MDLs.

Table 4-6. PE and CRM Results

Analyte	n	Percent Within Acceptance Range	Mean Percent Recovery	Range of Percent Recovery	SD of Percent Recovery	Concentration Range (mg/kg)
ERA Performance Evaluation Samples						
Arsenic	1	0	221	221	NA	349
Copper	3	67	136	92 - 172	41	144 - 196
Iron	4	0	179	145 - 244	45	7,130 - 10,400
Lead	3	0	153	140 - 165	12	128 - 208
Zinc	3	100	108	91 - 121	15	101 - 259
Certified Reference Materials						
Arsenic	1	100	115	115	NA	397
Chromium	1	100	115	115	NA	161,518
Copper	4	50	234	95 - 547	213	279 - 4,792
Iron	3	67	107	26 - 187	80	6,481 - 191,645
Lead	4	50	478	88 - 1,478	669	120 - 144,742
Nickel	1	0	124	123	NA	13,279
Zinc	4	0	756	0 - 1,900	991	546 - 22,217

Notes:
n Number of samples with detectable analytes.
SD Standard deviation.
mg/kg Milligrams per kilogram.
NA Not applicable, analyte not present above the LRL.

The XL Spectrum Analyzer performance data for all primary and secondary target analytes for the eight CRMs and PEs are summarized in Table 4-6 and Figure 4-7. No data is presented for chromium and nickel for the ERA PE samples because all samples had nondetectable concentrations of these two analytes. The measured values, true values, and percent recoveries for all detectable analytes for all PEs are shown in Figure 4-7. No figure is shown for arsenic, chromium, and nickel because there were only one or two detects for these three analytes. For the ERA PEs, the XL Spectrum Analyzer produced 5 out of 14 results or 35.7 percent within the acceptance range. For the CRM PEs, the XL Spectrum Analyzer produced 8 out of 18 results or 44.4 percent within the acceptance range. With the ERA and CRM PEs combined, the XL Spectrum Analyzer produced 13 out of 32 results or 40.6 percent within the acceptance range. Based on the data presented in Table 4-6, the XL Spectrum Analyzer's results were more comparable to the CRM PEs than the ERA PEs, most likely because the analyte concentrations were higher in the CRM PEs than in the ERA PEs.

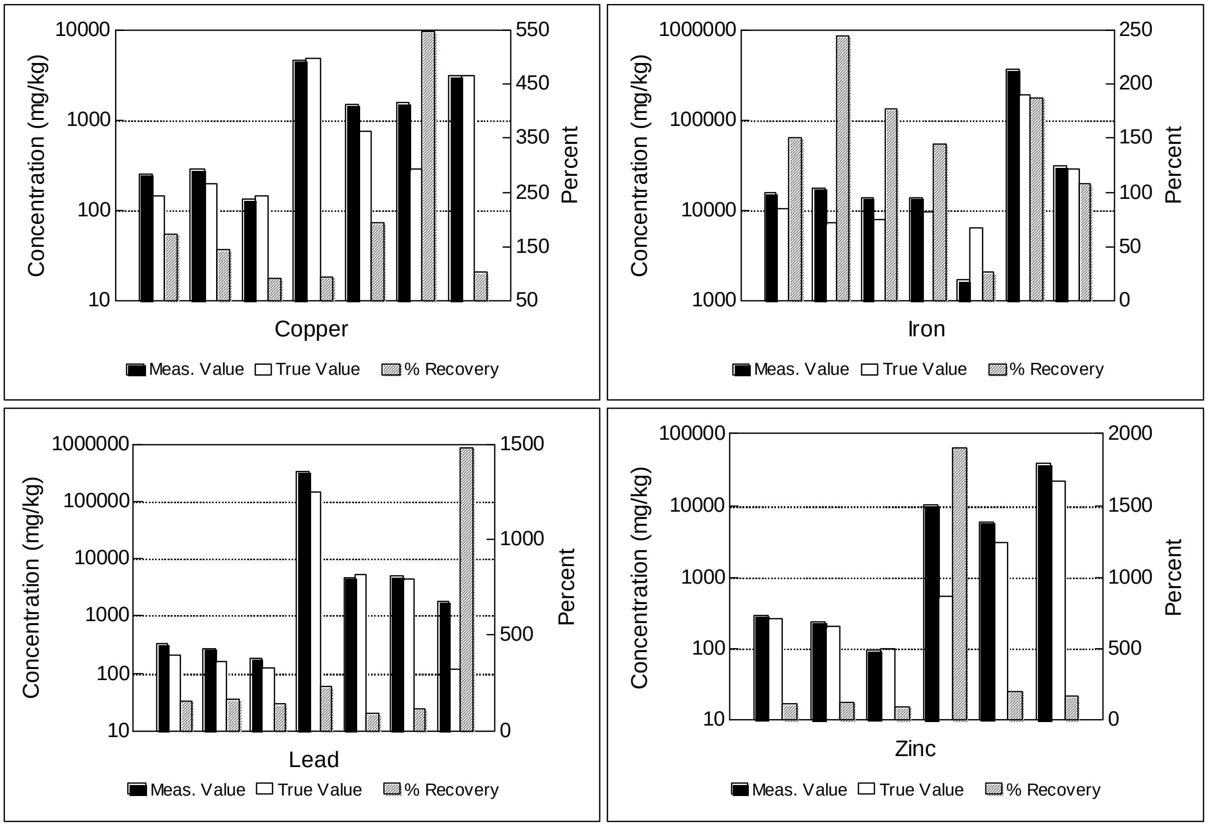


Figure 4-7. PE and CRM Results: These graphs illustrate the relationship between the analyzer's data (measured values) and the true values for the PE and CRM samples. The gray bars represent the percent recovery for the analyzer. Each set of three bars (black, white, and gray) represents a single PE or CRM sample.

All nine results outside the acceptance limits for the ERA PEs were above the upper control limit. Only two of the 14 percent recoveries were less than 100 percent. The mean percent recoveries were above 100 percent for all analytes in the ERA PEs. These results indicate that the XL Spectrum Analyzer was overestimating concentrations as compared to the certified values determined by EPA SW-846 Methods 3050A/6010A. The XL Spectrum Analyzer showed better comparability for copper and zinc than for arsenic, iron, and lead in the ERA PEs. The poor comparability in most cases was probably an artifact of the low analyte concentrations (near the detection limits) in the ERA PEs. With the exception of iron, all analyte concentrations were less than 350 mg/kg. For arsenic, copper, lead, and zinc, the analyte concentrations were all less than three times their respective precision-based MDLs and often at or below their respective field-based MDLs.

The comparability of the XL Spectrum Analyzer's results to the certified values in the CRMs did not appear to be matrix dependent. The comparability for zinc was vastly different in the CRM and ERA PEs. Comparability for zinc in the ERA PEs was good (100 percent) but was poor (0 percent) in the CRM PEs. This was not expected because the zinc concentrations were much lower in the ERA PEs than in the CRM PEs. One possible explanation for these results is that the CRM PEs contained much higher concentrations of other analytes such as copper, iron, and nickel, which may have caused interference problems for the zinc quantitation. Given these results for all the PEs, it is not advisable to use these PEs as QC checks for the XL Spectrum Analyzer.

Intermethod Assessment

The comparison of the XL Spectrum Analyzer's results to reference method's results was performed using the statistical methods detailed in Section 2. The purpose of this statistical evaluation was to determine the comparability of the data produced by the analyzer to that produced by the reference laboratory. If the $\log_{10}$ transformed FPXRF data were statistically equivalent to the $\log_{10}$ transformed reference data, and had acceptable precision (10 percent RSD or less), the data met the definitive level criteria. If the data did not meet the definitive level criteria, but could be mathematically corrected to be equivalent to the reference data, it met the quantitative screening level criteria. If the analyzer did not meet the definitive level criteria, and the statistical evaluation could not identify a predictable bias in the data, but the analyzer identified the presence or absence of contamination with at least a 90 percent accuracy rate, the data was classified as qualitative screening level.

The XL Spectrum Analyzer was configured to report arsenic, lead, chromium, copper, and zinc, all of the primary analytes for this demonstration. In a limited number of samples, it reported concentrations for two secondary analytes, iron and nickel. Other elements reported by the analyzer, but not evaluated during the demonstration, included molybdenum, zirconium, strontium, and rubidium.

During the demonstration, the developer did not provide guidance on the acceptability or use of data produced by the analyzer. In this light, the FPXRF data was originally assessed in its entirety, and no data was eliminated based on counting statistics and measurement standard deviations. Examination of this data set revealed a considerable data scatter associated with the lower concentration ranges for each analyte. Review of the raw data associated with these outliers indicated they were generally associated with high measurement SDs. Counting statistics and draft SW-846 Method 6200 identify data that is less than 3 times larger than its associated SD, as a nondetect and not useable. Based on this and through consultation with the developer, it was decided to conduct the data assessment on a revised FPXRF data set, where data less than 3 times its associated SD was considered not detected and was not used in the comparability assessment. The developer has changed its SOPs for the analyzer to include data usability criteria based on measurement SDs.

The analyzer's data for arsenic were strongly biased toward the ASARCO site, 644 data points, as compared to 36 data points from the RV Hopkins site (Table 4-7). At the ASARCO site, the arsenic concentrations were generally greater than lead concentrations, while at the RV Hopkins site samples exhibited lead concentrations 20 or more times greater than associated arsenic concentrations.

The reference laboratory reported no arsenic concentrations above 50 mg/kg at the RV Hopkins site; however, the analyzer reported arsenic concentrations ranging from approximately 200 - 1,000 mg/kg for the same samples. This occurrence of false positive readings at the RV Hopkins site is most likely due to the spectral proximity of the lead and arsenic X-ray emission energies. To compensate for this spectral overlap, the analyzer was programmed to quantitate arsenic based on the arsenic-K_β emission energy. However, the arsenic-K_β emission energy (11.73 keV) is close to the lead-L_β emission energy (12.61 keV). These energies have a separation of 0.88 keV. The resolution of the analyzer's silicone pin-diode detector is 0.80 keV. Apparently this resolution was not sufficient to allow arsenic quantitation at the RV Hopkins site, especially in the presence of 20 or more times greater lead concentrations.

Table 4-7. Regression Parameters[a] by Variable

Arsenic					Variable	Chromium				
n	r^2	Std. Err.	Y-Int.	Slope[b]		n	r^2	Std. Err.	Y-Int.	Slope[b]
671	0.82	0.26	0.52	0.80	All Data	327	0.50	0.22	2.60	0.30
635	0.96	0.12	-0.04	0.98	ASARCO	147	0.10	0.17	3.40	-0.21
36	0.03	0.24	2.70	-0.12	RV Hopkins	180	0.69	0.20	2.06	0.47
269	0.97	0.13	-0.12	1.00	Sand Soil	78	0.00	0.15	3.07	0.00
374	0.94	0.12	0.06	0.94	Loam Soil	69	0.22	0.18	0.35	-0.28
36	0.03	0.24	2.72	-0.12	Clay Soil	179	0.69	0.20	2.06	0.47
154	0.91	0.18	0.22	0.90	In Situ-Unprepared	90	0.49	0.25	2.49	0.35
170	0.86	0.22	0.45	0.82	In Situ-Prepared	51	0.68	0.19	2.32	0.37
166	0.96	0.12	-0.11	1.00	Intrusive-Unprepared	108	0.46	0.18	2.66	0.25
184	0.54	0.43	1.54	0.49	Intrusive-Prepared	79	0.46	0.24	2.56	0.31

Copper					Variable	Lead				
n	r^2	Std. Err.	Y-Int.	Slope[b]		n	r^2	Std. Err.	Y-Int.	Slope[b]
835	0.92	0.18	0.55	0.87	All Data	1085	0.96	0.12	0.14	0.95
774	0.96	0.13	0.26	0.95	ASARCO	710	0.96	0.12	0.14	0.95
61	0.79	0.16	1.11	0.83	RV Hopkins	375	0.96	0.12	0.17	0.94
335	0.94	0.14	0.28	0.95	Sand Soil	314	0.96	0.13	0.07	0.97
442	0.96	0.12	0.25	0.96	Loam Soil	396	0.96	0.11	0.24	0.92
61	0.79	0.16	1.12	0.83	Clay Soil	375	0.96	0.12	0.17	0.94
202	0.86	0.24	0.65	0.85	In Situ-Unprepared	284	0.88	0.21	0.38	0.88
201	0.97	0.12	0.33	0.92	In Situ-Prepared	268	0.98	0.09	0.15	0.94
202	0.94	0.15	0.46	0.89	Intrusive-Unprepared	269	0.98	0.07	0.07	0.97
231	0.93	0.19	0.68	0.84	Intrusive-Prepared	282	0.98	0.10	0.07	0.98

Zinc					Variable
n	r^2	Std. Err.	Y-Int.	Slope[b]	
784	0.89	0.17	-0.01	1.06	All Data
496	0.85	0.19	0.11	1.02	ASARCO Site
288	0.94	0.13	-0.14	1.11	RV Hopkins Site
201	0.94	0.16	-0.04	1.06	Sand Soil
302	0.54	0.25	0.67	0.83	Loam Soil
286	0.94	0.13	-0.14	1.11	Clay Soil
199	0.84	0.21	0.20	1.00	In Situ-Unprepared
181	0.93	0.13	-0.05	1.06	In Situ-Prepared
191	0.92	0.14	-0.07	1.08	Intrusive-Unprepared
212	0.89	0.18	-0.14	1.12	Intrusive-Prepared

Notes:

[a] Regression parameters based on $\log_{10}$ transformed data. These parameters were calculated for FPXRF data as the dependent variable, and thus, cannot be used to correct FPXRF data. See Section 5.

[b] Slope values determined with FPXRF data plotted on the y-axis and the reference data plotted on the x-axis.

N Number of data points.

Y-Int. Y-Intercept.

Std. Err. Standard error.

In spite of the influence of the false positive results from the RV Hopkins site, the regression analysis of the entire arsenic data set produced an r^2 meeting quantitative level data quality criteria. When the ASARCO data was examined separately, the soil variable did not appear to impact the analyzer's performance; however, sample preparation did have a major effect on instrument performance (Table 4-8). The analyzer's accuracy (as measured by a decrease in the standard error of the estimate, and in some cases, an increase in correlation) improved between the *in situ*-unprepared and *in situ*-prepared analyses. This step in the sample preparation process reflects sample homogenization. The next step in sample preparation, intrusive-prepared, resulted in an increase in accuracy; however, the strength of the correlation only increased half as much, relative to the increase exhibited after the initial sample homogenization. The act of preparing the intrusive-prepared samples involved 5 to 10 times more time, relative to the initial sample homogenization. The cost effectiveness of the additional effort in sample preparation will depend on a given project's data quality objectives. The final preparation step resulted in a slight decrease in accuracy and strength of the correlation. This may have been due to the physical removal of sample material associated with passing the sample through a No. 40-mesh sieve and to inherent instrument precision. The regression parameters for all preparation steps met the definitive level data quality criteria; however, an evaluation of the associated inferential statistics (t-test) indicated that under all four preparation steps, the analyzer's data and the reference laboratory data were significantly different. Therefore, for this demonstration, the analyzer produced quantitative screening level data for arsenic. Sample preparation was the only variable to significantly affect comparability. The greatest improvements in data comparability were exhibited after the initial sample homogenization.

Chromium exhibited a major site variable effect different from that discussed for arsenic. At the ASARCO site, chromium only occurred in its natural background concentrations, less than 40 mg/kg, well below the analyzer's precision-based MDL. However, the analyzer produced false positive chromium measurements for these samples ranging in concentration from approximately 750 to 2,500 mg/kg. This response to background chromium concentrations was used to calculate a comparability-based MDL of approximately 2,500 mg/kg for the ASARCO samples. No chromium concentrations exceeded this MDL at the ASARCO site, and therefore, no chromium results for the ASARCO site were considered valid. The RV Hopkins site was quite different. Reference laboratory chromium concentrations were well distributed in the samples, ranging from approximately 40 to greater than 5,000 mg/kg. The field-based MDL for the analyzer at this site was approximately 1,370 mg/kg. Both the ASARCO and RV Hopkins sites' field-based MDLs were almost 2 to 3 times greater than the precision-based MDL for chromium. This decrease in the MDL between the two sites may have been due to changes in the relative concentration of interfering elements such as manganese and iron. The RV Hopkins site had a much greater proportion of heavier elements which would alter the shape and intensity of the Compton peak for these samples. The iron and manganese concentrations were higher in the RV Hopkins samples relative to the ASARCO samples. It is unlikely that increased concentrations of interfering elements would reduce the analyzer's sensitivity to matrix background noise. The improved performance at the RV Hopkins site was probably due to its higher relative chromium concentrations.

Given the incomplete nature of the chromium data set, the resultant regression parameters could not be calculated. To truly assess the field performance of the analyzer for chromium, the ASARCO site's false positive data should be eliminated from the evaluation; however, given these problems, no assignment of data quality level for chromium can be made.

When the analyzer's chromium data was sorted and evaluated for each sample preparation step, the analyzer's accuracy (as measured by a decrease in the standard error of the estimate and increased correlation) improved between the *in situ*-unprepared and *in situ*-prepared sample preparation steps. This sample preparation step reflects the initial sample homogenization. The next sample preparation step,

Table 4-8. Regression Parameters[a] by the Sample Preparation Variable Sorted by Soil Texture

Arsenic						Chromium				
n	r^2	Std. Err.	Y-Int.	Slope[b]		n	r^2	Std. Err.	Y-Int.	Slope[b]
In Situ-Unprepared					Soil Texture	In Situ-Unprepared				
61	0.95	0.16	0.08	0.93	Sand Soil	25	0.03	0.15	2.74	0.20
91	0.89	0.17	0.28	0.88	Loam Soil	16	0.03	0.23	2.07	0.70
3	0.70	0.20	3.60	-1.16	Clay Soil	48	0.60	0.25	2.11	0.47
In Situ-Prepared					Soil Texture	In Situ-Prepared				
70	0.97	0.12	-0.05	0.98	Sand Soil	9	0.00	0.08	3.10	-0.08
93	0.96	0.11	-0.02	0.95	Loam Soil	3	0.00	0.23	2.87	0.06
10	0.07	0.24	2.43	0.25	Clay Soil	39	0.70	0.19	2.10	0.44
Intrusive-Unprepared					Soil Texture	Intrusive-Unprepared				
70	0.98	0.10	-0.26	1.05	Sand Soil	28	0.00	0.11	3.02	0.01
93	0.97	0.09	-0.04	0.97	Loam Soil	41	0.05	0.14	2.54	0.37
5	0.00	0.39	2.57	-0.01	Clay Soil	36	0.83	0.14	2.00	0.48
Intrusive-Prepared					Soil Texture	Intrusive-Prepared				
68	0.98	0.11	-0.23	1.04	Sand Soil	14	0.04	0.10	3.09	0.10
96	0.97	0.10	-0.05	0.98	Loam Soil	8	0.95	0.07	3.56	-0.32
18	0.10	0.22	2.76	-0.17	Clay Soil	54	0.76	0.17	2.04	0.47

Copper						Lead				
n	r^2	Std. Err.	Y-Int.	Slope[b]		n	r^2	Std. Err.	Y-Int.	Slope[b]
In Situ-Unprepared					Soil Texture	In Situ-Unprepared				
78	0.90	0.16	0.54	0.85	Sand Soil	84	0.88	0.20	0.33	0.86
110	0.89	0.21	0.40	0.93	Loam Soil	105	0.88	0.18	0.50	0.83
13	0.90	0.10	1.35	0.78	Clay Soil	93	0.88	0.20	0.56	0.85
In Situ-Prepared					Soil Texture	In Situ-Prepared				
85	0.95	0.12	0.26	0.95	Sand Soil	72	0.97	0.11	0.07	0.96
111	0.98	0.08	0.23	0.95	Loam Soil	100	0.98	0.07	0.24	0.91
11	0.78	0.15	1.19	0.78	Clay Soil	94	0.98	0.08	0.08	0.96
Intrusive-Unprepared					Soil Texture	Intrusive-Unprepared				
79	0.96	0.11	0.12	1.00	Sand Soil	78	0.98	0.10	0.03	0.99
111	0.98	0.09	0.25	0.95	Loam Soil	96	0.98	0.07	0.14	0.95
11	0.81	0.14	0.67	1.02	Clay Soil	96	0.99	0.07	0.07	0.97
Intrusive-Prepared					Soil Texture	Intrusive-Prepared				
91	0.94	0.14	0.17	1.00	Sand Soil	83	0.97	0.11	0.02	1.00
114	0.99	0.08	0.22	0.97	Loam Soil	99	0.98	0.08	0.06	1.00
25	0.74	0.19	1.20	0.77	Clay Soil	96	0.99	0.08	0.05	0.99

Zinc										
n	r^2	Std. Err.	Y-Int.	Slope[b]		n	r^2	Std. Err.	Y-Int.	Slope[b]
In Situ-Unprepared					Soil Texture	Intrusive-Unprepared				
45	0.92	0.18	0.14	0.98	Sand Soil	53	0.94	0.16	-0.05	1.07
70	0.66	0.20	0.60	0.84	Loam Soil	71	0.76	0.18	0.21	0.99
82	0.92	0.17	-0.03	1.11	Clay Soil	67	0.97	0.10	-0.27	1.15
In Situ-Prepared					Soil Texture	Intrusive-Prepared				
47	0.95	0.15	-0.13	1.08	Sand Soil	57	0.96	0.15	-0.15	1.11
70	0.74	0.16	0.37	0.92	Loam Soil	80	0.49	0.26	0.52	0.91
66	0.96	0.10	-0.16	1.10	Clay Soil	74	0.97	0.10	-0.34	1.18

Notes:
[a] Regression parameters based on $\log_{10}$ transformed data.
[b] Slope values determined with FPXRF data plotted on the y-axis and the reference data plotted on the x-axis.
n Number of usable matched pairs of data points.
Y-Int. Y-Intercept.
Std. Err. Standard Error.

intrusive-unprepared, resulted in an increase in accuracy, and the strength of the correlation also increased. The preparation of the intrusive-prepared samples involved 5 to 10 times more time than the initial sample homogenization. The cost effectiveness of this additional effort will depend on a project's data quality objectives. The final preparation step, intrusive-unprepared to intrusive-prepared, resulted in a decrease in accuracy and strength of the correlation. This decrease in comparability may have been due to the physical removal of sample material associated with passing the sample through a No. 40-mesh sieve. Sample preparation was the only variable to significantly affect comparability. The greatest increases in data comparability were exhibited after the initial sample homogenization.

Copper was detected at both sites. The range of copper concentrations was approximately 15 - 150,000 mg/kg at the ASARCO site and approximately 10 - 250 mg/kg at the RV Hopkins site. Field-based MDLs for copper were calculated to be approximately 365 mg/kg. This MDL is almost 3 times greater than the precision-based MDL. All of the RV Hopkins copper contamination was below the field-based MDL. Analyte concentrations near or below the MDLs should produce the highest measurement error. Statistical evaluation of the RV Hopkins data is not discussed in detail since the reported values, although greater than three times their measurement standard deviations, were all below the field-based MDLs and, therefore, represent concentration estimates only.

When the ASARCO data set for copper was evaluated separately, the regression analysis produced an r^2 of 0.96 (Table 4-7). Sorting the ASARCO copper data by sample preparation identified a significant sample preparation effect. Between the *in situ*-unprepared and *in situ*-prepared sample preparation steps, the r^2 increased from 0.91 to 0.98 (Table 4-9). The accuracy of the analyzer (as measured by the decrease in the standard error of the estimate and increased correlation) increased between these sample preparation steps. The next step in sample preparation, intrusive-unprepared, resulted in a slight decrease in accuracy; however, this may have been an artifact of the analyzer's inherent precision. The analyzer's precision for copper measurements was calculated to be approximately 13 percent which precludes assignment at the definitive data quality level. The final sample preparation step, intrusive-prepared, resulted in a slight improvement in the regression parameters, but there was no change in the standard error of the estimate. These relatively small changes in accuracy associated with the intrusive sample preparations may not be worth the additional resources required to conduct these sample preparations. The utility of further sample preparation, past the initial homogenization, would be dependent on a project's data quality objectives. The XL Spectrum Analyzer data for copper analysis at the ASARCO site was placed in the quantitative screening level data quality criteria. This holds for the complete data set, as well as for each sample preparation step-based data set. The analyzer also produced quantitative screening level data quality for the RV Hopkins copper analyses; however, the regression analysis indicated poorer correlation for this data set. Sample preparation was the only variable to significantly affect comparability.

The lead data was more or less evenly distributed between the two sites. This allowed a more thorough assessment of potential effects of the soil and site variables, in addition to the sample preparation variable. Initial examination of the entire lead data set showed that the ASARCO site produced 712 data points for lead, ranging in concentration from less than 10 to approximately 20,000 mg/kg, and that the RV Hopkins site produced 375 lead data points, ranging from approximately 30 to approximately 16,000 mg/kg.

The regression analysis of this entire data set for lead produced an r^2 of 0.95, meeting definitive level data quality criteria. When the data set was examined by site, the resultant regression parameters and inferential statistics were almost identical. Based on this, there does not appear to be an effect associated with the site variable. A similar finding was determined when the data set was evaluated by soil texture.

Table 4-9. Regression Parameters[a] by the Sample Preparation Variable Sorted by Site Name

Arsenic						Chromium				
n	r^2	Std. Err.	Y-Int.	Slope[b]		n	r^2	Std. Err.	Y-Int.	Slope[b]
In Situ-Unprepared					Site Name	In Situ-Unprepared				
151	0.92	0.17	0.17	0.91	ASARCO	41	0.02	0.18	2.71	0.24
3	0.70	0.20	3.60	-1.16	RV Hopkins	48	0.60	0.25	2.11	0.47
In Situ-Prepared					Site Name	In Situ-Prepared				
164	0.96	0.12	0.00	0.96	ASARCO	12	0.00	0.10	2.96	0.01
10	0.07	0.24	2.43	0.25	RV Hopkins	39	0.70	0.19	2.10	0.44
Intrusive-Unprepared					Site Name	Intrusive-Unprepared				
162	0.98	0.09	-0.19	1.02	ASARCO	70	0.05	0.14	2.75	0.22
5	0.00	0.39	2.57	-0.01	RV Hopkins	36	0.83	0.14	2.00	0.48
Intrusive-Prepared					Site Name	Intrusive-Prepared				
165	0.97	0.11	-0.16	1.02	ASARCO	23	0.70	0.13	3.63	-0.34
18	0.10	0.22	2.76	-0.17	RV Hopkins	54	0.76	0.17	2.04	0.47

Copper						Lead				
n	r^2	Std. Err.	Y-Int.	Slope[b]		n	r^2	Std. Err.	Y-Int.	Slope[b]
In Situ-Unprepared					Site Name	In Situ-Unprepared				
188	0.91	0.19	0.33	0.94	ASARCO	189	0.88	0.20	0.40	0.86
13	0.90	0.10	1.35	0.78	RV Hopkins	93	0.88	0.20	0.55	0.85
In Situ-Prepared					Site Name	In Situ-Prepared				
195	0.98	0.10	0.27	0.94	ASARCO	172	0.97	0.09	0.15	0.94
11	0.78	0.15	1.20	0.78	RV Hopkins	94	0.98	0.08	0.08	0.96
Intrusive-Unprepared					Site Name	Intrusive-Unprepared				
192	0.97	0.11	0.25	0.95	ASARCO	173	0.98	0.08	0.06	0.98
11	0.81	0.14	0.67	1.02	RV Hopkins	96	0.99	0.07	0.07	0.97
Intrusive-Prepared					Site Name	Intrusive-Prepared				
204	0.98	0.11	0.22	0.97	ASARCO	183	0.97	0.10	0.05	0.99
25	0.81	0.16	1.09	0.84	RV Hopkins	96	0.99	0.08	0.05	0.99

Zinc										
n	r^2	Std. Err.	Y-Int.	Slope[b]		n	r^2	Std. Err.	Y-Int.	Slope[b]
In Situ-Unprepared					Site Name	Intrusive-Unprepared				
116	0.81	0.21	0.35	0.92	ASARCO	124	0.90	0.16	0.01	1.06
82	0.92	0.17	-0.03	1.11	RV Hopkins	67	0.97	0.10	-0.27	1.15
In Situ-Prepared					Site Name	Intrusive-Prepared				
115	0.92	0.14	0.01	1.04	ASARCO	139	0.80	0.24	0.03	1.07
66	0.96	0.10	-0.16	1.10	RV Hopkins	74	0.97	0.10	-0.34	1.18

Notes:
[a] Regression parameters based on $\log_{10}$ transformed data. These parameters were calculated for FPXRF data as the dependent variable, and thus, cannot be used to correct FPXRF data. See Section 5.
[b] Slope values determined with FPXRF data plotted on the y-axis and the reference data plotted on the x-axis.
n Number of usable matched pairs of data points.
Y-Int. Y-Intercept.
Std. Err. Standard Error.

The regression parameters for each soil texture were very similar with the y-intercept showing the greatest shift. The y-intercepts for the loam and clay soil were almost the same and were greater than the y-intercept for the sandy soil (Table 4-9). With the slopes being so close to 1.00 and the similarities of the standard errors for each regression, it is not likely that the difference in y-intercepts is either important or the result of a soil texture effect. When the lead data was sorted by sample preparation and the regression analysis run, a soil preparation effect was observed. The analyzer's accuracy (as measured by a decrease in the standard error of the estimate and increased correlation) improved between the *in situ*-unprepared and *in situ*-prepared sample preparation steps. This sample preparation step reflected sample homogenization. The next sample preparation step, intrusive-unprepared, resulted in an increase in accuracy; however, the strength of the correlation did not change. The act of preparing the intrusive-prepared samples involved 5 to 10 times more time relative to the initial sample homogenization. The cost effectiveness of this additional effort in sample preparation will depend on a project's data quality objectives. The final preparation step, intrusive-unprepared to intrusive-prepared, resulted in a slight decrease in accuracy and a similar decrease in the strength of the correlation. This decrease may have been due to the physical removal of sample material associated with passing the sample through a No. 40-mesh sieve and to the inherent instrument precision. The analyzer's precision for lead measurements was calculated to be approximately 9 percent for lead. Although the regression parameters for all sample preparation steps met the definitive level data quality criteria, the inferential statistics indicated that the data sets exhibited no significant differences only after the second sample preparation step (*in situ*-prepared). For this demonstration, this analyzer produced definitive level data for lead for all the analyses. Sample preparation was the only variable to significantly affect comparability (Figure 4-8).

The zinc concentrations were more or less evenly distributed between the two sites, similarly to the lead concentrations. This allowed an assessment of potential effects of the soil and site variables, in addition to the sample preparation variable. Initial examination of the entire zinc data set showed that the ASARCO site produced 496 data points for zinc, ranging in concentration from approximately 20 to 5,000 mg/kg, and that the RV Hopkins site produced 286 zinc data points ranging from approximately 30 to 12,000 mg/kg. The outliers removed from the regression analysis were primarily associated with the ASARCO site data and with samples exhibiting copper concentrations from 11,000 to 50,000 mg/kg. Only 2 of the 27 outliers for the entire data set were from the RV Hopkins site. Twenty-one out of the 25 outliers identified for the ASARCO site were collected from the same area. This area represented the highest copper contamination sampled during the demonstration. A total of seven points was sampled in this area. The samples from these points were analyzed at each of the four sample preparation steps, producing 28 potential measurements from this area. Two of these measurements were valid analyzer data points and were not identified as outliers. The remaining 26 data points were identified as either outliers or they were not considered valid because of their high measurement standard deviations. The soils in this area were light blue in color due to the high copper concentrations. The reference data for these samples indicated copper concentrations ranging from 11,210 to 154,460 mg/kg. High concentrations of copper can cause interference for corresponding zinc measurements. Copper has a K_β emission energy of 8.6 keV and zinc has a K_α emission energy of 9.5 keV. The XL Spectrum Analyzer quantitates zinc from the K_α peak. The proximity of these emission energies and the resolution of the analyzer's detector (0.80 keV) may have produced spectral overlap and as a result the error associated with the zinc measurements. This data suggests that copper concentrations in the 11,000 - 50,000 mg/kg range will produce significant interference for zinc measurements. This interference causes the analyzer to artificially elevate the zinc concentrations, resulting in false positive readings.

When the zinc data set was examined by the site variable, the resultant regression parameters were different. The correlation or r^2 between the FPXRF data and the reference data was poorer for the ASARCO site data relative to the RV Hopkins data. In addition, the standard error of the estimate was greater for the ASARCO site data. This apparent site effect is most likely an artifact of the elevated copper concentrations associated with some of the ASARCO site samples. Further examination of the data when sorted by soil texture indicated that this apparent site effect was simply an artifact of the soil texture. The loam soil exhibited a much poorer correlation and greater standard error. This could be due to the fact that the loam soils occurred in the area of greatest copper contamination sampled at the ASARCO site. The interference effect associated with the high copper concentrations is most likely the cause of this observed soil texture effect and not an effect associated with soil texture or particle size.

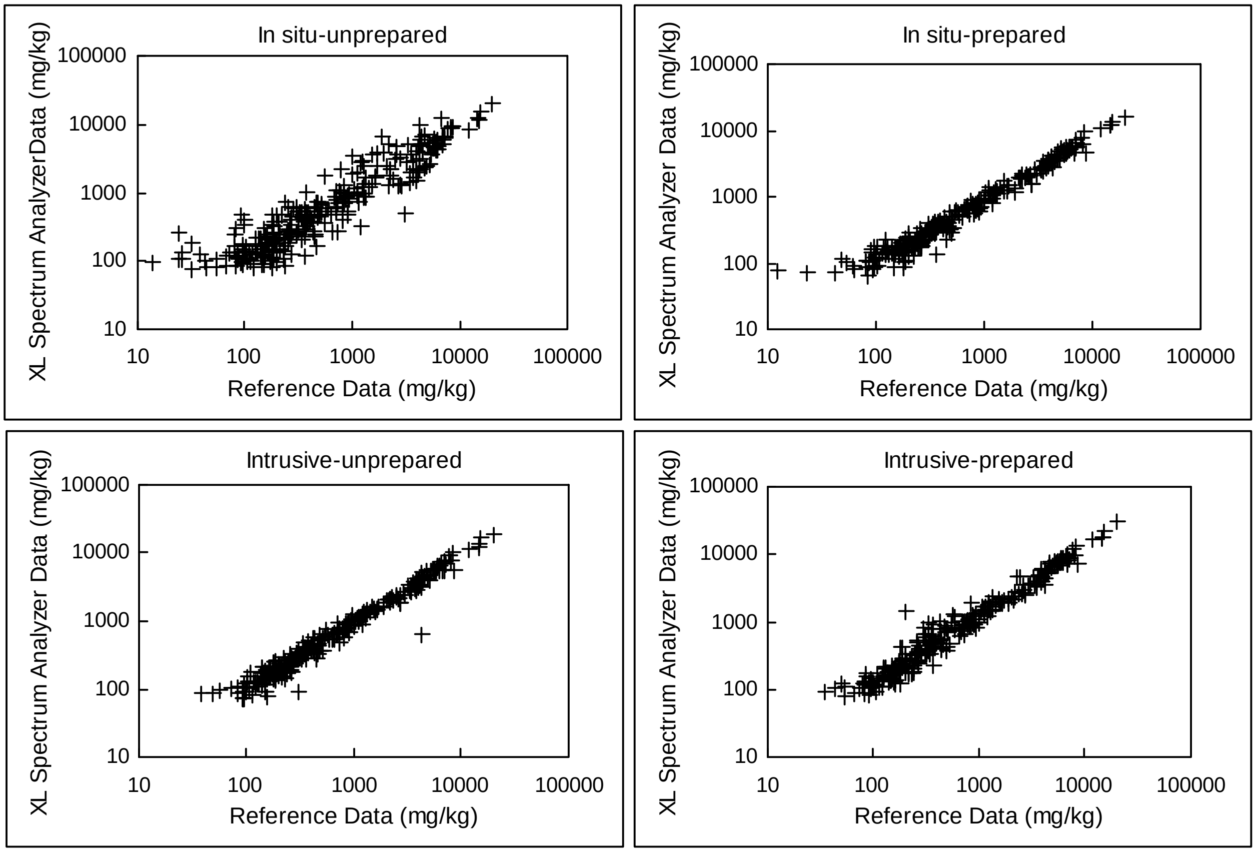


Figure 4-8. Sample Preparation Effect on Lead Results: These graphs illustrate the change in comparability with changes in sample preparation.

When the zinc data was sorted by sample preparation step and the regression analysis run, a soil preparation effect measured by a decrease in the standard error of the estimate and increased correlation improved between the *in situ*-unprepared and *in situ*-prepared sample preparation steps (Table 4-7). The cost effectiveness of this additional effort in sample preparation will depend on a project's specific data quality objectives. The final preparation step resulted in a slight decrease in accuracy and strength of the correlation. The limited changes in analyzer performance after the initial sample homogenization may have been due the inherent instrument precision, calculated to being approximately 11 percent for zinc. Additional sample preparation, past the initial homogenization step, did not affect the comparability of the data. Although the regression parameters for all preparation steps met the definitive level data quality criteria, the inferential statistic indicated that the data sets were significantly different. Therefore, for this demonstration, this analyzer produced quantitative screening level data quality for zinc.

Within the sample preparation steps, the effects of contaminant concentration were also examined. The data sets for the primary analytes were sorted into the following concentration ranges: 0 - 100 mg/kg, 100 - 1,000 mg/kg, and greater than 1,000 mg/kg (Table 4-10). The regression analysis for each primary analyte and for each preparation step was rerun on the concentration-based data sets. For target analytes that exhibited susceptibility to interferences, the data affected by the interference was removed, as was done in the intermethod evaluation. Of the primary analytes, only lead and zinc exhibited field-based MDLs spanning the upper two tiers of the three concentration ranges. The remaining primary analytes generally had field-based MDLs in the mid to upper range of the middle concentration tier. This left only the upper tier of concentration ranges for the evaluation of concentration effects. Having complete data for only one of the tiers of concentration ranges made it impossible to meaningfully evaluate the effect of contaminant concentration ranges for these other target analytes. However, lead and zinc had complete data for both the middle and upper concentration range; this allowed at least a qualitative assessment of concentration effect. Based on this data, both lead and zinc exhibited stronger comparability and less error at the upper concentration range (Table 4-10). This relationship is expressed through improved r^2 values and standard error of the estimates for the upper concentration range relative to the middle concentration range; however, for both concentration ranges the analyzer produced quantitative screening level data quality. This relationship of increasing accuracy and reduced relative error as concentrations increase is generally exhibited by most analytical methods.

Table 4-10. Concentration Effect Data for Lead and Zinc

	In Situ-Unprepared			In Situ-Prepared			Intrusive-Unprepared			Intrusive-Prepared		
Concentration Range (ppm)	No. of Samples	r^2	Std. Err.	No. of Samples	r^2	Std. Err.	No. of Samples	r^2	Std. Err.	No. of Samples	r^2	Std. Err.
Lead												
0 - 100	9	0.12	0.20	18	0.21	0.11	14	0.00	0.07	28	0.15	0.08
100 - 1,000	159	0.62	0.20	153	0.88	0.09	159	0.91	0.08	163	0.84	0.12
>1,000	95	0.67	0.19	97	0.96	0.06	96	0.97	0.06	96	0.95	0.07
Zinc												
0 - 100	17	0.01	0.08	4	0.00	0.09	10	0.01	0.06	8	0.60	0.08
100 - 1,000	150	0.59	0.21	145	0.78	0.14	150	0.78	0.15	166	0.78	0.16
>1,000	32	0.48	0.21	32	0.94	0.07	31	0.92	0.08	32	0.93	0.08

To examine the effect of count times on the analyzer's comparability, a subset of 26 samples, intrusive-prepared, from the RV Hopkins site was reanalyzed using twice the original count times. This did not significantly affect the comparability, as measured by the r^2, slope, intercept, and standard error of the regression, except for chromium. For chromium, only the slope and y-intercept showed a significant change with increased count times. The slope and y-intercept shifted from 0.49 and 2.0, respectively, when the 60 live-second count times were used, to 0.87 and 0.89, respectively, when the count times were doubled to 120 live-seconds.

In summary, the XL Spectrum Analyzer produced quantitative screening level data quality for arsenic, copper, and zinc. The analyzer produced definitive level data for lead. The precision-based MDLs were generally 2 to 3 times lower than the field-based MDLs. The precision-based MDLs represent the optimal estimate of the MDL, where the field-based MDL may be more representative of actual instrument performance relative to contaminated soil samples. These limits can range from the thousands of milligrams per kilogram for chromium to the hundreds of milligrams per kilogram for lead and zinc. Of the three variables examined, site, soil, and preparation, sample preparation was the only variable to exhibit significant impact on the data comparability. In fact, the most significant increases in data comparability were exhibited after the initial sample homogenization. Subsequent sample preparation generally improved the comparability of the data at a fraction of the level exhibited by the initial homogenization. The need for the improved comparability associated with the more involved levels of sample preparation should be determined on a project specific basis and driven by data quality objectives.

A final decision regarding the assignment of data quality leads involves an assessment of both r^2 and precision RSD results. Using the criteria presented in Table 2-2, a summary of the Niton XL Spectrum Analyzer's data quality performance in this demonstration is presented in Table 4-11.

Table 4-11. Summary of Data Quality Level Parameters

Target Analytes	XL Spectrum Analytes	Precision Mean % RSD 5 - 10 X MDL	Method Detection Limits (mg/kg) (Precision-based)	Coefficient of Determination (r^2 All Data)	Data Quality Level
Arsenic	Arsenic	9.2	120	0.82	Quantitative
Barium	—	—	—	—	—
Chromium	Chromium	ND	900	0.50	Insufficient Data
Copper	Copper	13.2	130	0.92	Quantitative
Lead	Lead	6.5	75	0.96	Definitive
Zinc	Zinc	11.2	115	0.89	Quantitative
Nickel	—	—	—	—	—
Iron	—	—	—	—	—
Cadmium	—	—	—	—	—
Antimony	—	—	—	—	—

Section 5
Applications Assessment and Considerations

The Niton XL Spectrum Analyzer was originally designed to analyze for lead in paint. This demonstration found that it could also provide analytical data on metals contamination in soil for lead at the definitive level and for arsenic, copper, and zinc at the quantitative screening level. The Niton "SOILAIR" software used for calibration and quantitation maximized instrument performance and accounted for most common soil-related matrix interferences. Although this instrument was not designed for field use, with a few minor weatherizing measures, it was operated under a variety of environmental conditions. The analyzer never experienced failure resulting in down time throughout the 1-month field demonstration. During this time, almost 1,260 samples were analyzed. The limited training provided by the developer was sufficient to allow basic field operation. The developer provided on-line technical support that was sufficient to allow uninterrupted operation throughout the demonstration. A summary of key operational features is listed in Table 5-1.

Table 5-1. Summary of Test Results and Operational Features

Weighs less than 3 pounds, battery lifetime of 8 hours
Sample throughput of 20 to 25 samples per hour
Can conduct in situ and intrusive measurements
Generally low purchase price
Easy operation; brief training period
Uses Compton Ratio Method of Calibration
Calibration drift RSD values less than 8 percent for all analytes monitored
Precision percent RSD values less than 15 percent at 5 to 10 times the MDL for all analytes; a value for Cr was not determined at this concentration
Can be used on soils exhibiting 30 percent water saturation by weight
Produced data of definitive quality for lead. Data of quantitative screening level for arsenic, copper, and zinc
Single excitation source (Cd^{109})
Requires a specific radiation license to operate in most states

Comparison of $\log_{10}$ transformed Niton XL and $\log_{10}$ transformed reference laboratory data indicated that for most metals, the analyzer provided quantitative screening level quality data. This data quality level is applicable to most field applications. The data produced by this analyzer was $\log_{10}$-$\log_{10}$ linearly correlated to the reference data. This linear correlation held more than 5 orders of magnitude. The relationship between the analyzer data and the reference data indicates that if 10 - 20 percent of the samples analyzed were submitted for reference method analysis, the XL Spectrum Analyzer raw data could be corrected to more closely match the reference data. In the case of lead, after the initial sample homogenization, the XL Spectrum Analyzer data was statistically equivalent to the reference data. This instrument exhibited precision slightly lower than the reference method, generally between 10 and 20 percent, indicating a high degree of reproducibility.

The XL Spectrum Analyzer is generally operated with relatively short count times (60 - 120 live-seconds) and uses only one radioactive source. This single radioactive source limits the number of analytes that can be detected. The XL Spectrum Analyzer's "SOILAIR" software can report concentrations for molybdenum, zirconium, strontium, rubidium, nickel, arsenic, chromium, iron, copper, lead, and zinc in soil samples. The relatively short count times and the single radioactive source combine to increase the sample throughput and detection limits but decrease the analyzer accuracy and precision.

This demonstration identified sample preparation as the most important variable with regard to analyzer performance and data comparability. The Niton analyzer can be used in an *in situ* or intrusive mode. The data from this demonstration indicated that when operated in the *in situ* mode, the results did not show a strong correlation between FPXRF and reference data. This may not be due to instrument error but rather to the inherent spatial variability of contamination, even within an area as small as the 4-inch by 4-inch grid sampled during this demonstration. The greatest increase in correlation between the FPXRF data and reference data was achieved after the initial sample homogenization. However, further sample preparation, such as sieving or drying and sieving, did not greatly improve the comparability.

The XL Spectrum Analyzer was strongly affected by arsenic and lead interference at the RV Hopkins site. Samples from this site exhibited lead to arsenic concentration ratios greater than 10:1. In these samples, the instrument produced false positive readings for arsenic in all cases. These false positive results ranged from 4 to 20 times the actual arsenic concentration in the samples. A similar interference was experienced at the ASARCO site. Chromium analysis proved to be problematic for the analyzer, with MDLs ranging from 900 to 1,500 mg/kg. Iron or manganese interference may have caused this poor response. The remaining target analytes did not appear to be significantly affected by interfering elements unless they represent in abnormally high concentrations, e.g., copper interfering with zinc.

The steps needed to correct the field measurements to more closely match the reference data are as follows:

1. Conduct sampling and FPXRF analysis.
2. Select 10 - 20 percent of the sampling locations for resampling. These resampling locations can be evenly distributed over the range of concentrations measured, or they can focus on an action level concentration range.
3. Resample the selected locations. Thoroughly homogenize the samples and have each sample analyzed by FPXRF and a reference method.
4. Tabulate the resulting data with reference data in the y-axis column (dependent variable) and the FPXRF data in the x-axis column (independent variable). Transform this data to the equivalent $\log_{10}$ value for each concentration.
5. Conduct a linear regression analysis and determine the r^2, y-intercept and slope of the relationship. The r^2 should be greater than 0.70 to proceed.

6. Place the regression parameters into Equation 5-1:

$$Y(\log_{10}\ \mathit{corrected\ FPXRF\ data}) = \mathit{slope} * (\log_{10}\ \mathit{FPXRF\ data}) + Y\text{-}\mathit{intercept} \tag{5-1}$$

7. Use the above equation with the $\log_{10}$ transformed FPXRF results from Step 4 above and calculate the equivalent $\log_{10}$ corrected FPXRF data.

8. Take the anti-$\log_{10}$ ($10^{[\log_{10}\text{ transformed corrected FPXRF data}]}$) of the equivalent $\log_{10}$ corrected FPXRF data calculated in Step 7. These resulting values (in milligrams per kilogram) represent the corrected FPXRF data.

To show the effect of correcting the FPXRF data, the change in average relative bias and accuracy can be examined. The average relative bias between the FPXRF data and the reference data is a measure of the degree to which the FPXRF over- or underestimates concentrations relative to the reference methods. The relative bias is an average number for the entire data set and may not be representative of an individual measurement. An example of this can be seen in an analyzer's data where measurements are underestimated at low concentrations but overestimated at high concentrations. On average, the relative bias for this analyzer is zero; however, this bias is not representative for high or low concentration measurements. To avoid this dilemma, three approaches can be taken: (1) the evaluation of average relative bias can be focused on a narrow concentration range, (2) the analyzer's data can be corrected using the regression approach described above, or (3) average relative accuracy can be calculated. Average relative accuracy represents the percentage that an individual measurement is over- or underestimated relative to a reference measurement. Table 5-2 shows the average relative bias and accuracy exhibited by the Niton XL, before and after data correction using the eight-step approach previously discussed.

The average relative bias and accuracy for lead (Pb), which fell into the definitive level data quality category, were generally smaller than for the other elements. The analytes falling into the quantitative and qualitative screening level data quality categories had generally larger average relative bias and accuracy.

Once the data is corrected using the regression approach presented earlier, in most cases, the average relative bias and accuracy were reduced. The average relative bias numbers are no longer strongly influenced by a concentration effect since the regression approach used to correct the data used $\log_{10}$ transformed data. The average relative bias and accuracy for the corrected data are similar to the acceptable average relative bias between the reference data and PE samples (true values), as shown by the last column in Table 5-2.

Based on this demonstration, the XL Spectrum Analyzer is well suited for the rapid real-time assessment of selected metals contamination in soil samples. Although in several cases the analyzer produced data statistically equivalent to the reference data, generally confirmatory analysis will be required or requested for FPXRF analysis. If 10 - 20 percent of the samples analyzed by the analyzer are submitted for reference method analysis, instrument bias, relative to standard methods such as 3050A/6010A, can be corrected. Bias correction allows FPXRF data to be enhanced so that it approximates the reference data. The demonstration showed that this analyzer exhibited a strong linear relationship with the reference data more than a 5 orders of magnitude concentration range. For optimum correlation, samples with high, medium, and low concentration ranges from a project must be submitted for reference analysis.

The Niton XL Spectrum Analyzer can provide rapid assessment of the distribution of metals contamination in soils at a hazardous waste site. This data can be used to characterize general site contamination, guide critical conventional sampling and analysis, and monitor removal actions. This demonstration suggested that for some elements, the field data is equivalent to the reference data. SW-846 Method 6200 for FPXRF analysis will help in the acceptance of this data for quantitative screening level applications and possibly definitive level applications. FPXRF data can be produced and interpreted in the field on a daily or per sample basis. This real-time analysis allows the use of contingency-based sampling for any application and greatly increases the potential for meeting project objectives on a single mobilization. This analyzer is an effective tool for site characterization and remediation. It provides a faster and less expensive means of analyzing metals contamination in soil.

Table 5-2. Effects of Data Correction on FPXRF Data Comparability to Reference Data for All In Situ-Prepared Samples

Target Analyte	Average Relative Bias on Raw Data[a]	Average Relative Bias on Corrected Data[b]	Average Relative Accuracy on Raw Data[c]	Average Relative Accuracy on Corrected Data[d]	Acceptable Accuracy for PE Samples[e]
Arsenic	0.88	1.04	1.33	1.33	1.76
Chromium	12.1	1.52	17.67	2.44	1.55
Copper	1.31	1.04	1.10	1.74	1.18
Iron	1.42	1.01	1.48	1.61	1.54
Lead	0.97	1.02	1.52	1.51	1.63
Zinc	1.35	1.03	1.52	1.27	1.64

Notes:

[a] A measurement of average relative bias, measured as a factor by which the FPXRF, on average, over- or underestimates results relative to the reference methods. This measurement of bias is based on raw (not $\log_{10}$ transformed) data. This average relative bias does not account for any concentration effect on analyzer performance.

[b] A measurement of average relative bias on the FPXRF data after it has been corrected using the eight-step regression approach.

[c] A measurement of average relative accuracy at the 95 percent confidence interval, measured as a factor by which the raw FPXRF, on average, over- or underestimates individual results relative to the reference methods. This measurement of accuracy is based on raw (not $\log_{10}$ transformed) data. This average relative accuracy is independent of concentration effects.

[d] A measurement of average relative accuracy at the 95 percent confidence interval, of the corrected FPXRF data obtained using the eight-step regression approach.

[e] A measurement of accuracy represents a factor and 95 percent confidence interval that define the acceptable range of differences allowed between the reference method reported concentrations and the true value concentrations in the PE samples. This bias is included only as a general reference for assessing the improvement on comparability of FPXRF data and reference data after FPXRF data correction.

The average relative bias is calculated as follows:

Average relative bias = $((\sum_i[\mathrm{FPXRF}_i/\mathrm{Reference}_i])/\text{number of paired samples})-1$

This value represents the percentage that the FPXRF over- or underestimates the reference data, on average, for the entire data set. To convert this calculated value to a factor, 1.0 is added to the calculated average relative bias. The above table presents the average relative bias as a factor.

The average relative accuracy is calculated as follows:

Average relative accuracy = SQRT $(\sum_i([\mathrm{FPXRF}_i/\mathrm{Reference}_i]-1)^2/\text{number of paired sample})$

This value represents the percentage that an individual FPXRF measurement over- or underestimates the reference data. The relative accuracy numbers in the table are calculated at the 95 percent confidence interval. This is accomplished by adding two standard deviations to the above formula before the square root is taken. To convert this calculated value to a factor, 1.0 is added to the calculated average relative accuracy. The above table presents the average relative bias as a factor.

General Operational Guidance

The following paragraphs describe general operating considerations for FPXRF analysis. This information is derived from SW-846 Method 6200 for FPXRF analysis.

General operation of FPXRF instruments will vary according to specific developer protocols. For all environmental applications, confirmatory or reference sampling should be conducted so that FPXRF data can be corrected. Before operating any FPXRF instrument, the developer's manual should be consulted. Most developers recommend that their instruments be allowed to warm up for 15 - 30 minutes before analysis of any samples. This will help alleviate drift or energy calibration problems.

Each FPXRF instrument should be operated according to the developer's recommendations. There are two modes in which FPXRF instruments can be operated: *in situ* and intrusive. The *in situ* mode involves analysis of an undisturbed soil or sediment sample. Intrusive analysis involves collecting and preparing a soil or sediment sample before analysis. Some FPXRF instruments can operate in both modes of analysis, while others are designed to operate in only one mode. The two modes of analysis are discussed below.

For *in situ* analysis, one requirement is that any large or nonrepresentative debris be removed from the soil surface before analysis. This debris includes rocks, pebbles, leaves, vegetation, roots, and concrete. Another requirement is that the soil surface be as smooth as possible so that the probe window will have good contact with the surface. This may require some leveling of the surface with a stainless-steel trowel. Most developers recommend that the soil be tamped down to increase soil density and compactness. This step reduces the influence of soil density variability on the results. During the demonstration, this modest amount of soil preparation was found to take less than 5 minutes per sample location. The last requirement is that the soil or sediment not be saturated with water. Most FPXRF instruments will perform adequately for soils with moisture contents of 5 - 20 percent, but it will not perform well for saturated soils, especially if ponded water exists on the surface. Data from this demonstration did not see an effect on data quality from soil moisture content. Source count times for *in situ* analysis usually range from 30 to 120 seconds, but source count times will vary between instruments and depending on required detection limits.

For intrusive analysis of surface soil or sediment, it is recommended that a sample be collected from a 4- by 4-inch square that is 1 inch deep. This will produce a soil sample of approximately 375 grams or 250 cm^3, which is enough soil to fill an 8-ounce jar. The sample should be homogenized, dried, and ground before analysis. The data from this demonstration indicates that sample preparation, beyond homogenization, does not greatly improve data quality. Sample homogenization can be conducted by kneading a soil sample in a plastic bag. One way to monitor homogenization when the sample is kneaded in a plastic bag is to add sodium fluorescein dye to the sample. After the moist sample has been homogenized, it is examined under an ultraviolet light to assess the distribution of sodium fluorescein throughout the sample. If the fluorescent dye is evenly distributed, homogenization is considered complete; if the dye is not evenly distributed, mixing should continue until the sample has been thoroughly homogenized. During the demonstration, the homogenization procedure using the fluorescein dye required 3 to 5 minutes per sample.

Once the soil or sediment sample has been homogenized, it can be dried. This can be accomplished with a toaster oven or convection oven. A small portion of the sample (20 - 50 grams) is placed in a suitable container for drying. The sample should be dried for 2 to 4 hours in the convection or toaster oven at a temperature not greater than 150 °C. Microwave drying is not recommended. Field studies have shown that microwave drying can increase variability between the FPXRF data and reference data. High levels of metals in a sample can cause arcing in the microwave oven, and sometimes slag will form in the sample.

The homogenized, dried sample material can also be ground with a mortar and pestle and passed through a 60-mesh sieve to achieve a uniform particle size. Sample grinding should continue until at least 90 percent of the original sample passes through the sieve. The grinding step normally averages 10 minutes per sample.

After a sample is prepared, a portion should be placed in a 31-mm polyethylene sample cup (or equivalent) for analysis. The sample cup should be completely filled and covered with a 2.5-micrometer Mylar™ (or equivalent) film for analysis. The rest of the soil sample should be placed in a jar, labeled, and archived. All equipment, including the mortar, pestle, and sieves, must be thoroughly cleaned so that the method blanks are below the MDLs of the procedure.

Technology Update

The text below was taken verbatim from information submitted by Niton Corporation.

In the fall of 1994, we first agreed to participate in the demonstration. We had worked out a simple method for correcting composition-related soil matrix effects that we felt made our XL Spectrum Analyzer a practical tool for the field analysis of lead in soil. By the time of the demonstration in April 1995, we had developed a more refined calculation method that corrected for spectral overlaps and provided multiple element capability. At the time of the demonstration, we had not yet completed the instrument-resident soil application software, so we used a palm-top personal computer linked to the instrument to perform the needed calculations and store data. The system used in the demonstration was a prototype of what became the XL-LISA.

In the summer of 1995, we commercially introduced the XL-LISA (XL Lead In Soil Application). Intended as a software upgrade to the standard XL Spectrum Analyzer (our lead paint analyzer), XL-LISA includes instrument-resident software to analyze for lead, arsenic, zinc, and copper in soil, and store up to 500 complete readings with spectra on-board the instrument. XL-LISA also includes a kit of equipment and supplies for preparing soil samples for accurate XRF analysis.

Since the demonstration and introduction of XL-LISA, we have made a number of technical improvements in the XL hardware design, electronics, software, and calibration procedures. These improvements have led to more consistent, stable operation, somewhat better accuracy and lower detection limits.

We are presently launching a companion product, the NITON 700, which combines the XL's small size and weight with exceptionally high sensitivity in thin sample application (such as air sampling filters). The NITON 700 offers a much expanded range of multiple element capability for soil and thick sample analysis. We will continue to push for improvements in sensitivity and accuracy as we further develop this product.

Section 6
References

Havlick, Larry L., and Ronald D. Crain. 1988. *Practical Statistics for the Physical Sciences.* American Chemical Society. Washington, D.C.

Kane, J. S., S. A. Wilson, J. Lipinski, and L. Butler. 1993. "Leaching Procedures: A Brief Review of Their Varied Uses and Their Application to Selected Standard Reference Materials." *American Environmental Laboratory.* June. Pages 14-15.

Kleinbaum, D. G., and L. L. Kupper. 1978. *Applied Regression Analysis and Other Multivariable Methods*. Wadsworth Publishing Company, Inc., Belmont, California.

Morgan, Lewis, & Bockius. 1993. RODScan$_{®}$.

PRC Environmental Management, Inc. 1995. "Final Demonstration Plan for Field Portable X-ray Fluorescence Analyzers."

U.S. Environmental Protection Agency. 1993. "Data Quality Objectives Process for Superfund-Interim Final Guidance." Office of Solid Waste and Emergency Response. Washington, D.C. EPA/540/R-93/071.