♻EPA Environmental Technology Verification Report

Field Portable X-ray Fluorescence Analyzer

Metorex X-MET 920-P and 940

SITE
SUPERFUND INNOVATIVE
TECHNOLOGY EVALUATION

ETV✓
Environmental Technology
Verification Program

Environmental Technology Verification Report

Field Portable X-ray Fluorescence Analyzer

Metorex X-MET 920-P and 940

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:	**X-MET 920-P AND X-MET 940**
COMPANY:	**METOREX, INC.**
ADDRESS:	**PRINCETON CROSSROADS CORPORATE CENTER**
	P.O. BOX 3540
	PRINCETON, NJ 08543-3540
PHONE:	**(609) 406-9000**

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 Metorex X-MET 920-P and 940 analyzers.

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, Metorex X-MET 920-P and 940." The EPA document number for this report is EPA/600/R-97/146.

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

These analyzers operate on the principle of energy dispersive X-ray fluorescence spectroscopy where the characteristic energy 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 which enables the use of low intensity excitation sources (such as radioisotopes) and compact battery-powered, field-portable electronics. The 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.

Metorex considers the X-MET 920-P and 940 to have equivalent performance characteristics. Advances in electronics have led to the redesign of the 920-P into a smaller and lighter version, the X-MET 940. At the time of the demonstration, the 920-P was commercially available while the 940 was tested as a prototype.

These instruments are designed to produce quantitative data on the concentration of metals in soils, sludges, and other solids. Each instrument consists of a battery-operated electronics unit and a solid-state probe system (SSPS). The SSPS houses two excitation sources and a lithium-drifted (Si[Li]) detector (cooled by liquid nitrogen) for elemental excitation and detection. The SSPS allows for *in situ* analysis or the measurement of samples in cups. Either instrument can be operated and calibrated using site-specific calibration samples or through the use of fundamental parameters (FP) calibration software. During this demonstration, the FP software was used and fine-tuned with the use of one site-specific sample to improve data comparability. During this demonstration, the instruments were configured to report arsenic, barium, cadmium, copper, chromium, iron, lead, nickel, and zinc. At the time of the demonstration, each instrument cost about $55,000; either could be leased for $6,000 per month.

VERIFICATION OF PERFORMANCE

These findings do not distinguish between the two analyzers. The original study design intended to test the prototype X-MET 940; however, near the end of the data collection at the ASARCO site, a data acquisition problem occurred which prevented any additional use of this analyzer. The demonstration was resumed using an X-MET 920-P. It should also be noted that the first 920-P unit also encountered difficulties with data acquisition and was replaced by the developer. Although both instruments encountered problems, Metorex technical support was responsive to the needs of the demonstration.

The performance characteristics of the X-MET 920-P and 940 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 120 milligrams per kilogram (mg/kg) or less for arsenic, barium, cadmium, copper, lead, nickel, and zinc. The measured value for chromium was 210 mg/kg. A value for iron was not determined due to insufficient samples in the required concentration range.
- **Throughput:** Average throughput was 10 to 12 analyses per hour using a live count of 240 seconds. This rate only represents the analysis time since different personnel were used to prepare the samples.

- **Drift:** This is a measurement of the analyzer's variability in quantitating a known concentration of a standard over time. No data was produced for either analyzer to assess drift.
- **Completeness:** These instruments produced results for 1,192 of the 1,260 samples for a completeness of 94.6 percent, slightly below the demonstration objective of 95 percent. Software and mechanical problems reduced completeness.
- **Blank results:** Lithium carbonate blanks were analyzed periodically throughout the demonstration. Concentrations of copper and iron were detected in all the blanks.
- **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 values for all analytes were less than 8 percent, except chromium and nickel which had RSD values of 23 and 25 percent, respectively. Values for iron and cadmium were not reported due to insufficient data.
- **Accuracy:** Intramethod accuracy was assessed using site-specific PE soil samples and soil SRMs. The results show that 28 of 38 (73.6 percent) of the analytes in the site-specific PEs were within the quantitative acceptance range of 80 - 120 percent. The barium and cadmium concentrations were underestimated in all PE samples, and one of the three measured values for chromium was overestimated. The soil SRM data showed that 19 of 30 or 63 percent of the analytes were in the acceptable range.
- **Comparability:** This demonstration showed these instruments 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 0.94 for arsenic, 0.93 for copper, 0.94 for lead, 0.86 for zinc, 0.67 for chromium, and 0.43 for barium. Values for cadmium, nickel, and iron were not reported due to insufficient data.
- **Data quality levels:** Based on precision and comparability to the reference methods, these instruments produced definitive level data for arsenic, lead, copper, and zinc and data of qualitative screening level for chromium and barium. Values for cadmium, nickel, and iron could not be assigned without adequate precision or comparability data.

The results of this demonstration show that either the Metorex X-MET 920-P or X-MET 940 can provide useful, cost-effective data for environmental problem-solving and decision-making. Undoubtedly, these instruments 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

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 standard reference methods, (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 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 represented three distinct soil textures: sand, loam, and clay. 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 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 regarding the X-MET 920-P and X-MET 940 analyzers. Separate ETVRs have been published for the other analyzers demonstrated.

Some operational downtime was experienced by the X-MET 920-P during periods of heavy rainfall, resulting in hardware problems which limited the use of the X-MET 940 at the RV Hopkins site. Some data loss occurred at both sites. Although the cause of this data loss cannot be determined, it is likely that with additional experience and familiarity with the analyzers, the operator could have reduced this data loss.

During operations, quantitative data was provided by both analyzers on a real-time basis. The X-MET 920-P and 940 analyzers reported arsenic, chromium, copper, lead, zinc, iron, nickel, cadmium, and barium. These analyzers used total source count times of 240 live-seconds for this demonstration. These count times resulted in a sample throughput averaging between 10 and 12 samples per hour.

The X-MET 920-P and 940 analyzers provided definitive level data (equivalent to reference data) for copper, arsenic, lead, and zinc. The analyzers produced required qualitative screening level data for chromium and barium. No assignment of data quality could be made for nickel, iron, or cadmium due to a lack of sufficient data needed to calculate precision, accuracy, or the coefficient of determination.

The analyzers generally exhibited a precision similar to that of the reference methods. They exhibited precision values ranging between 3 and 25 percent relative standard deviation at 5 times the method detection limit (MDL) for all of the reported analytes. The analyzers' quantitative results were based on a calibration using the fundamental parameters method. The field-based MDLs were generally 2 to 3 times higher than the precision-based MDLs. Except for nickel, the field-based MDLs were higher than the developer's published MDLs. This difference was most likely due to count-time differences and to the differences in the developer's definition or the demonstration's definition of MDLs.

This demonstration found that the X-MET 920-P and 940 analyzers were 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. These analyzers can provide rapid, real-time analysis of the metals content of soil samples at hazardous waste sites. Either analyzer can quickly distinguish contaminated areas from noncontaminated areas, allowing investigation and remediation decisions to be made more efficiently on-site which may reduce the number of samples that need to be submitted for confirmatory analysis. The X-MET 920-P and 940 analyzers were found to be effective tools for field-based analysis of metals contamination in soil.

Table of Contents

List of Figures

List of Tables

List of Abbreviations and Acronyms

α	alpha
β	beta
ACES	automated contaminant evaluation software
Am^{241}	americium-241
CCB	continuing calibration blank
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
CRM	certified reference material
EPA	Environmental Protection Agency
ERA	Environmental Resource Associates
ETVR	environmental technology verification report
eV	electron volt
Fe^{55}	iron-55
FP	fundamental parameters
FPXRF	field portable X-ray fluorescence
ICAL	initial calibration
ICP-AES	inductively coupled plasma-atomic emission spectroscopy
ICS	interference check standard
ICV	initial calibration verification
IDW	investigation-derived waste
keV	kiloelectron volt
LCS	laboratory control samples
LED	light-emitting diode
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
NDD	Norton Disk Doctor

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
PC	personal computer
PE	performance evaluation
ppm	part per million
PRC	PRC Environmental Management, Inc.
psi	pounds per square inch
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
Si(Li)	silicon lithium
SITE	Superfund Innovative Technology Evaluation
SOP	standard operating procedure
SRM	standard reference material
SSPS	solid state probe system
TC	toxicity characteristic
USGS	United States Geological Survey
XPCS	X-MET PC System
XRF	X-ray fluorescence

Acknowledgments

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 necessary 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 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 on 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 it should reduce the number of samples that need to be submitted for confirmatory analysis. Of the seven commercially available analyzers evaluated, 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 Niton Corporation (the Niton XL Spectrum 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. The X-MET 920-P and 940 are essentially the same instruments, only their physical characteristics differ. This report presents data on the X-MET 920-P and 940 analyzers. 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 analytes 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 of the target analytes were determined by each instrument.

To demonstrate the analyzers, hazardous waste sites in Iowa (the RV Hopkins site) and in the State of Washington (the ASARCO site) were selected. These sites were selected 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 two analyzers were simple to operate in the field. However, unexpected software and hardware problems caused data loss and project downtime. The developer had provided a training course for the operator that was similar to that provided to a purchaser of the instrument. The training 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. Some of the downtime and data loss encountered could have been avoided with increased operator experience; however, Metorex provided accessible and timely field support. The analyzers were portable and capable of operating continuously over a 12-hour work day with appropriate battery changes. The almost continuous rain encountered at the ASARCO site caused an operational downtime for the X-MET 920-P Analyzer. Moisture entered the probe causing a source controlling electronics malfunction, as well as fogging of the source detector window.

The analyzers reported results for lead, arsenic, zinc, copper, nickel, iron, cadmium, barium, and chromium. Antimony was the only target analyte not reported by the analyzers. The analyzers used two radioactive sources coupled with a silicon-drifted lithium (Si[Li]), semiconductor, crystal diode detector. The type and strength of the radioactive sources allowed each instrument to produce reliable data at count times as short as 100 live-seconds. The count times used in this demonstration resulted in a sample throughput averaging 10 to 12 samples per hour.

The X-MET 920-P and 940 analyzers produced data meeting definitive level criteria (equivalent to reference data) for arsenic, copper, lead, and zinc. The analyzers produced qualitative screening level data for chromium and barium. Assignment of data quality levels for cadmium, nickel, and iron could not be made due to insufficient data.

The analyzers exhibited a precision similar to the reference methods. The chromium data generally showed the lowest precision of the primary analytes. Site and soil texture did not appear to affect data comparability. This conclusion is based on data associated with lead and zinc, the two target analytes evenly distributed, over a wide range of concentrations, at all sites and in all soil textures. The use of fundamental parameters (FP) calibration seemed to minimize any impact of inter-element interferences.

Based on performance, this demonstration found both the Metorex 920-P and 940 to be effective tools 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 X-MET 920-P and 940 analyzers. These two analyzers, as designed by Metorex, are identical with the exception of their size and weight. The design of the 940 is based on advances in microelectronics which allows the data collection, interpretation, and storage unit to be more compact and lighter. At the time of the demonstration, the X-MET 920-P was commercially available, while the X-MET 940 was a prototype unit. This ETVR presents information relative to the X-MET 920-P and 940 analyzers. The X-MET 920-P and 940 analyzers use a Si(Li) detector and any combination of three radioactive sources: cadmium-109 (Cd^{109}), americium-241 (Am^{241}), and iron-55 (Fe^{55}). The analyzers used in this demonstration were equipped with the Cd^{109} and Am^{241} sources only. The analyzers can be operated in an *in situ* or intrusive mode. The *in situ* mode is commonly called "point-and-shoot". In this mode of operation, the point of measurement on the soil surface is cleared of loose debris and organic matter, the analyzer probe is 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 placed into 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 X-MET 920-P and 940 analyzers' 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 how well FPXRF analyzers identify and quantify concentrations of 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 their performances; (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, and 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 conventional 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 this demonstration, the term "reference" was substituted for "confirmatory" since the data were 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 its position as the only commercial laboratory identified that could perform all the analyses in the required timeframe.

Method 3050A is the standard acid extraction procedure for determining metals concentrations in soil samples. It is not a total digestion method, and it potentially does 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 have 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 of data quality referred to as the PARCC parameters: precision, accuracy, representativeness, completeness, and comparability. In addition, method detection limits (MDLs) are often 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 methods was acceptable. The following paragraphs provide 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 differences (RPD) is used to provide this estimate of random errors 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 analytical methods, 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, PRC and EPA determined that the RV Hopkins and ASARCO sites met most of the site-selection criteria, and therefore, would be the sites used 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 this 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 6.68 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 for 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 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 and by the reference methods.

The X-MET 920-P and 940 analyzers can be operated in either the *in situ* or intrusive mode. The two modes of FPXRF analysis involve slightly different measurement and sampling procedures (Figure 2-1). Each procedure was designed to reflect common applications of FPXRF analyzers. 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 then took one *in situ* measurement in the middle of each sample area. This data point 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. All replicate measurements were taken from the same spot in the sampling area.

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 homogenized 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 homogenization 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. This monitoring process assumed that thorough distribution of sodium fluorescein was indicative of good sample homogenization. The effectiveness of this homogenization 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.

CHOOSE
SAMPLE LOCATION

IN SITU MEASUREMENT
(4X4 INCH GRID)

CONDUCT
PRECISION MEASUREMENT
(4% OF
LOCATIONS)

YES → COLLECT
10 REPLICATE
MEASUREMENTS
WITHOUT
MOVING PROBE

NO

COLLECT SOIL FROM
GRID TO DEPTH OF 1 INCH

COLLECT 10-20 GRAM
SUBSAMPLE FOR
WATER CONTENT
DETERMINATION

HOMOGENIZE
WITH
FLUORESCEIN DYE

SPREAD OUT SOIL IN
1-INCH DEEP PETRI DISH

CONDUCT
IN SITU
MEASUREMENT

WAS SAMPLE
PREVIOUSLY USED
FOR PRECISION
DETERMINATION?

YES → COLLECT
10 REPLICATE
MEASUREMENTS
WITHOUT
MOVING PROBE

NO

PASS SOIL THROUGH
NO. 10 SIEVE

IS SAMPLE A
FIELD DUPLICATE?

YES → SPLIT AND PACKAGE
TWO 20-GRAM ALIQUOTS
FOR REFERENCE METHOD
ANALYSIS

LABEL 30% FOR BOTH
3052/6010A ANALYSIS
AND 3050A/6010A

NO

PACKAGE 20
GRAMS FOR
CONFIRMATORY
ANALYSIS

LABEL 30% FOR
BOTH 3052/6010
AND 3050A/6010A
ANALYSIS

CONDUCT
INTRUSIVE INSTRUMENT
MEASUREMENTS
(NO PREPARATION)

WAS SAMPLE
PREVIOUSLY USED
FOR PRECISION
DETERMINATION?

YES → CONDUCT 10 REPLICATE
MEASUREMENTS

NO

PREPARE
10% SAMPLES DRY
20 GRAMS WITH
MICROWAVE OVEN

SAMPLE FOR
CONFIRMATORY
ANALYSIS

DRY SAMPLE IN
CONVECTION OVEN
1 HOUR AT 110°C

GRIND AND PASS
THROUGH NO. 40
SIEVE

CONDUCT
INTRUSIVE INSTRUMENT
MEASUREMENT OF
MATERIAL PASSING
NO. 40 SIEVE

WAS SAMPLE
PREVIOUSLY USED
FOR PRECISION
DETERMINATION?

YES → CONDUCT 10
REPLICATE
MEASUREMENTS

NO

END

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

9

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 analysis.

Following the intrusive-unprepared procedure, a portion of the 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). These samples were then analyzed in the 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).

These preparation procedures allowed the evaluation of the effects of sample preparation on FPXRF comparability to reference data.

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 learning 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. The developers trained the operators using their respective operator training manuals. 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 significant "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 can be no qualitative 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 X-MET 920-P and 940 reported all of these analytes except antimony.

Evaluations of analyzer data comparability involved examining the effects of each site, soil texture, and sample preparation technique on performance and comparability (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; then each soil texture was subdivided by the four soil preparations. These variables allowed the examination of particle size and homogenization effects on data comparability. These effects were believed to have the greatest potential 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 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 drying of samples 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 of samples. To assess this 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 the 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 it 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 the best results for specific target analytes. The developers, however, selected count times that produce 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 a user in developing a project cost analysis.

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 with the two intrusive sample preparations. Theoretically, precision and accuracy should increase as particle size decreases and becomes more 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. This effect was partially

12

examined in the comparison of analyzer performance between intrusive-unprepared and intrusive-prepared analyses. This step in sample preparation involved drying and grinding.

- 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 the 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 samples 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 samples being measured, 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 method.

The reference methods chosen for this analysis involve wet chemical analysis and partial 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, which represents the material from most surfaces, and clay and carbonate minerals. 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. This equates to a sample volume of approximately 7.5 cm^3. X-rays returning to the detector are derived from all 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 analysis in contrast to the reference methods, which represent a select or partial analysis. This difference can result in FPXRF concentrations that are higher than corresponding reference data when metals are contained within nonacid soluble complexes or constituents. 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 method but rather to the inherent differences in the nature of the analytical methods.

The comparison of FPXRF data and the reference data employs linear regression as the primary statistical tool. Linear regression analysis intrinsically contains assumptions and conditions that must be valid for the data set. Three of the most important assumptions are: (1) the linearity of the relationship, (2)

13

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.
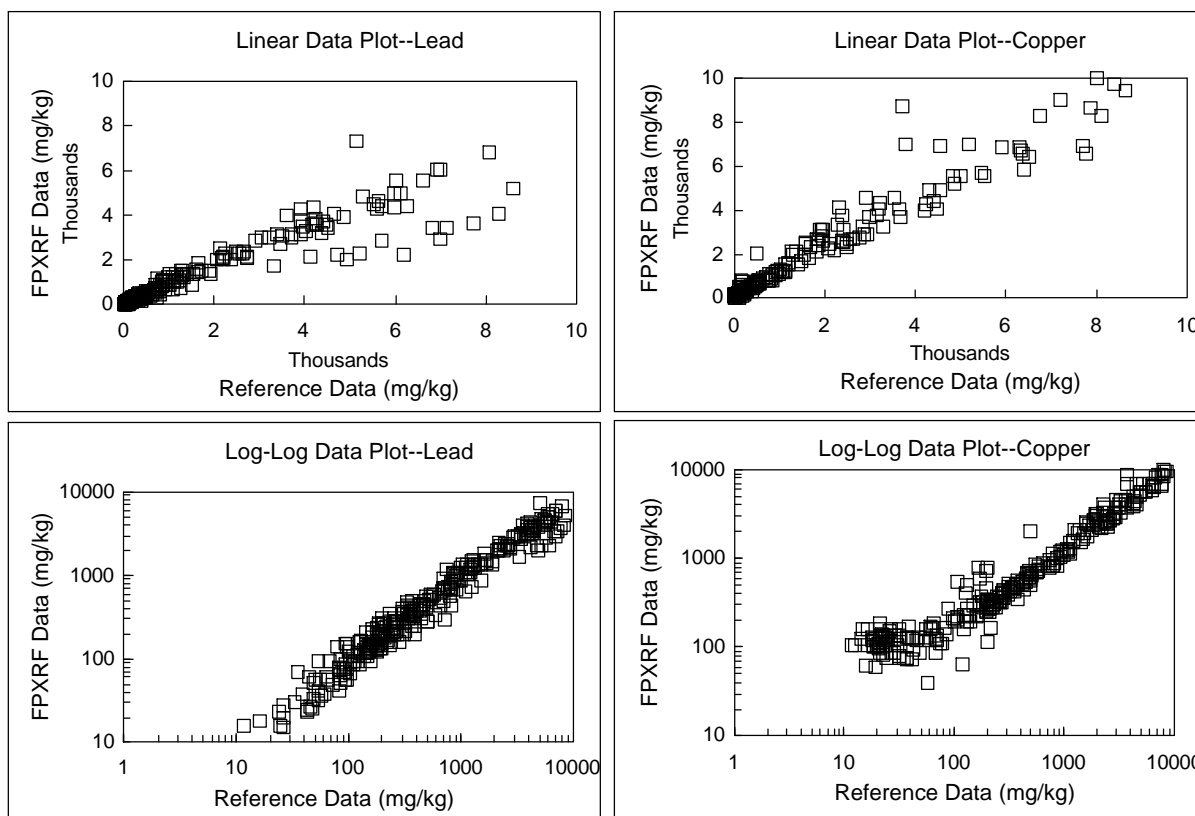
Figure 2-2. Linear and Log-log Data Plots–X-MET 920-P and 940 Analyzers: These graphs illustrate the linear relationship between the X-MET 920-P and 940 analyzer data and the reference data. The linear data plots illustrate the concentration dependence of this relationship with increased scatter at higher concentrations. The log-log plots eliminate this concentration effect. Scatter is relatively constant over the entire plot.

The 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.

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 vary 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 these findings, the reference data does allow unambiguous resolution of data quality determination.

The comparison of the reference data to the FPXRF data represents the 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).

Data 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 the screening level data.

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 inferential statistics indicate 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 in Section 4, "Intermethod Comparison."

[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.

Definitive level data are considered the highest level of quality. These data are usually generated from rigorous analytical methods such as those approved by the EPA or ASTM. 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 QA/QC requirements are satisfied.

Quantitative screening data provide unconfirmed analyte identification and quantification, although the quantification may be relatively imprecise. It is commonly recommended that at least 10 percent of the screening data must 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, 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, approved EPA methods for FPXRF did not exist. As part of this study, 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 FY-97. For purposes of this demonstration, the absence of a current 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 as defined in Table 2-2.

The comparability data set for each analyzer consisted of 1,260 matched pairs of reference method data for each target analyte. 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.

To obtain an adequate data set to evaluate the performance of the analyzers, a total of 315 soil samples was analyzed by the reference laboratory. These samples were analyzed by the Metorex analyzers for each of the four sample preparation steps. This produced an equivalent set of 1,260 data values, 630 in each mode *in situ* or intrusive. Seventy of the 315 samples submitted to the reference laboratory were split and submitted 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 variables 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. Therefore, 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.

$$Y \ = \ mX \ + \ b \hspace{6cm} \textbf{(2-1)}$$

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 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 were considered to have met the first requirement for definitive level data classification (Table 2-2). The second criteria, precision was then examined and was required to be equal or less than 10 percent RSD to retain the definitive data quality level. If both these criteria were not satisfied, certain inferential statistical parameters were then evaluated. First, 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 procedure was used to assign an overall data quality level for each analyte.

> *Slope Test for Significant Differences* $\hspace{5cm}$ **(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.*

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

18

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.

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 was less than 20 percent RSD, and the slope or intercept were not statistically equivalent to the ideal values, the analyzer was considered to produce quantitative screening level data quality (Table 2-2). 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 they had a precision greater than 20 percent RSD. 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 resultant represented 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. The value of the MDL was assigned by determining the concentration where the linear relationship disintegrates and reporting a value at two SDs above this concentration. 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 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 from the samples, the samples for moisture content determination and the intrusive-prepared samples 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 within common microwave drying times used in the field.

An analysis of the microwaved samples showed that this drying 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 altering 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 a reduction in metals concentration. 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 the homogenization was completed, 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 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 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.

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

Although the possibility for poorly homogenized samples exists under any 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.
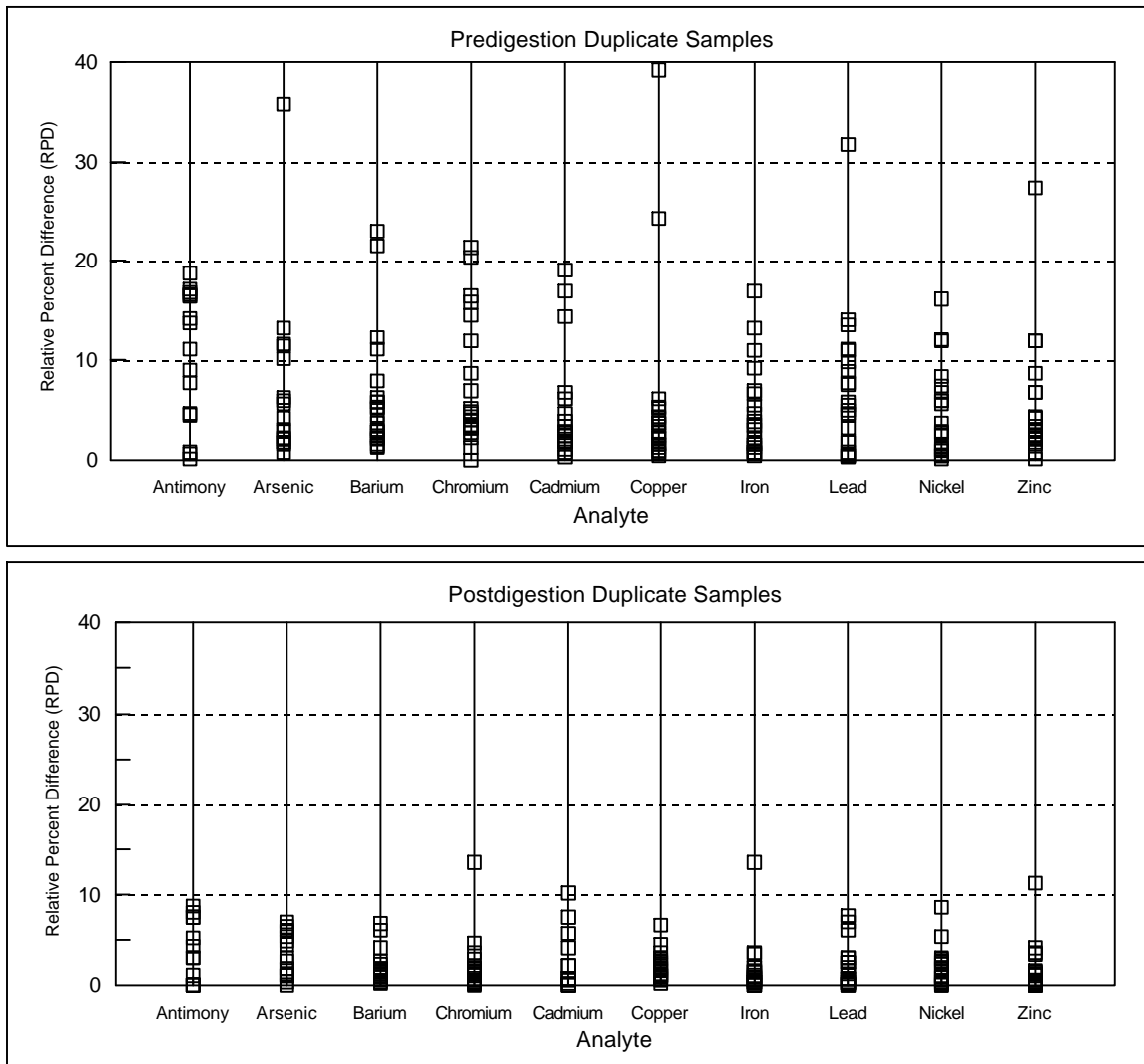
Figure 3-1. Pre- and Postdigestion Duplicate Samples: The top graph illustrates the reference laboratory's performance on analyzing <u>predigestion</u> duplicate samples. Twenty percent RPD represents the <u>predigestion</u> duplicate control limits defined in the demonstration QAPP. Two points were 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

32

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.
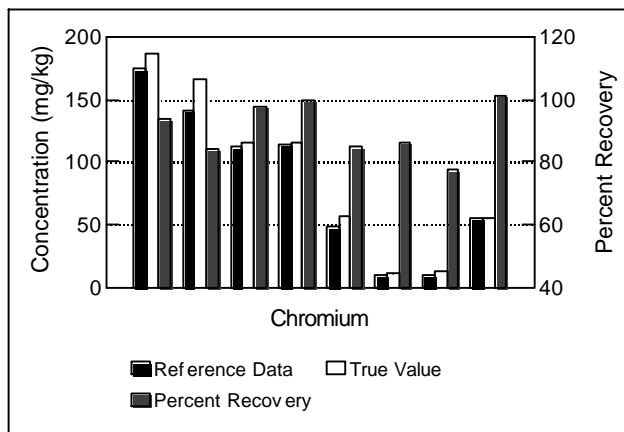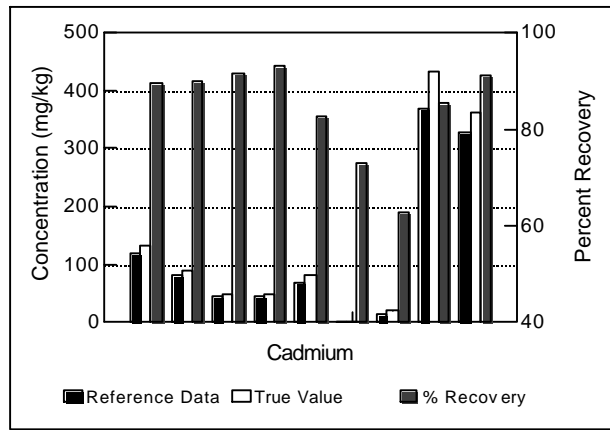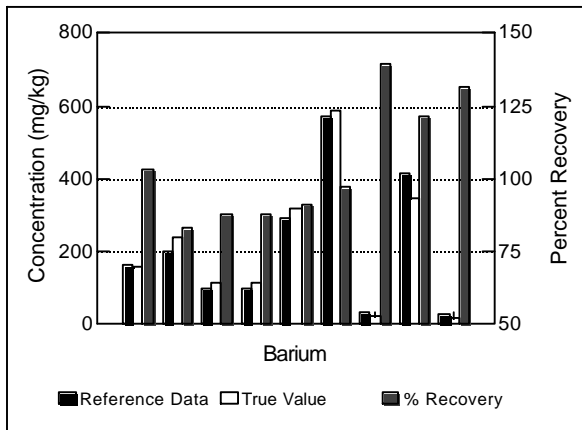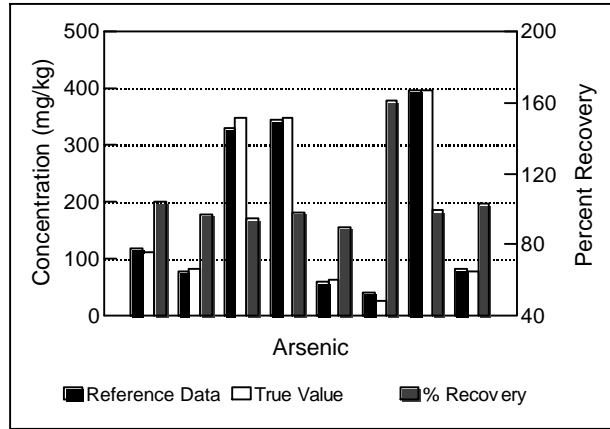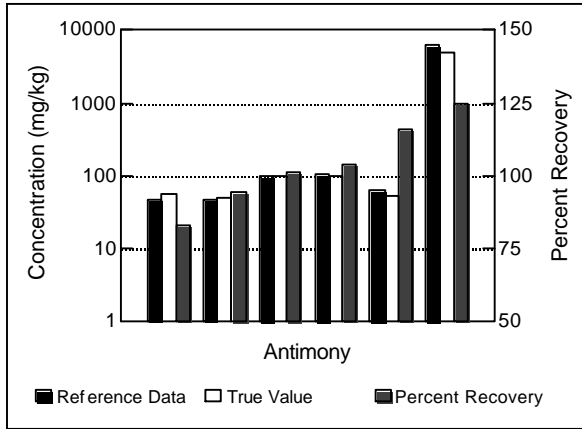
Figure 3-2. Reference Method PE and CRM Results: These graphs illustrate the relationship between the reference data and the true values for the PE or CRM samples. The gray bars represent the percent recovery for the reference data. Each set of three bars (black, white, and gray) represents a single 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.
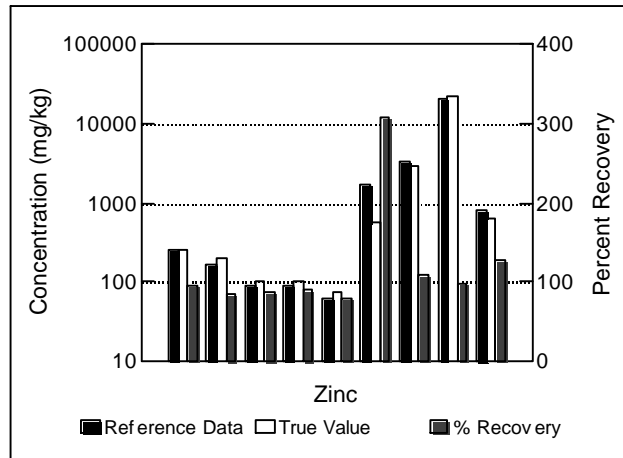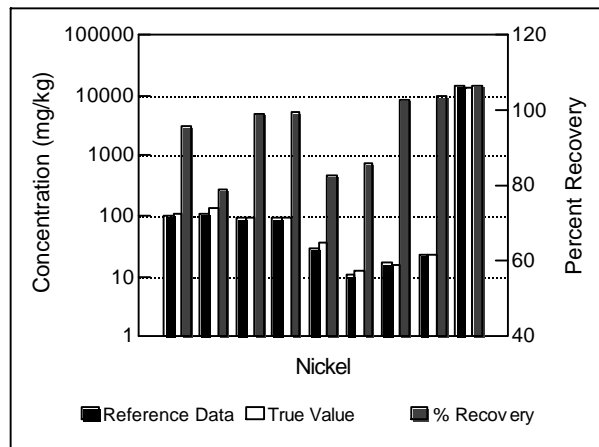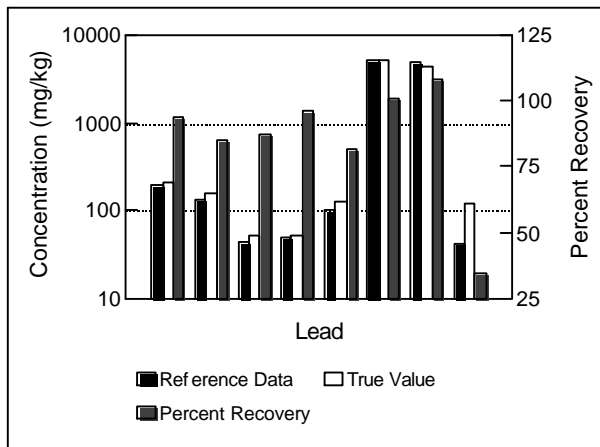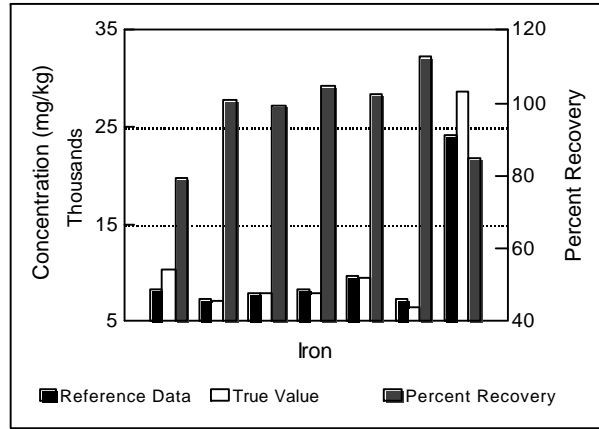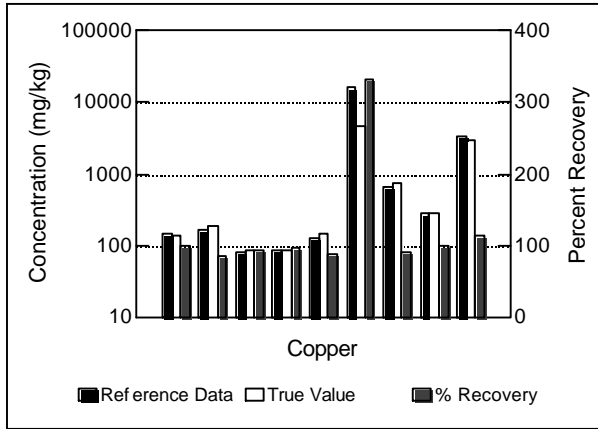
Figure 3-2 (Continued). Reference Method PE and CRM Results: These graphs illustrate the relationship between the reference data and the true values for the PE or CRM samples. The gray bars represent the percent recovery for the reference data. Each set of three bars (black, white, and gray) represents 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.
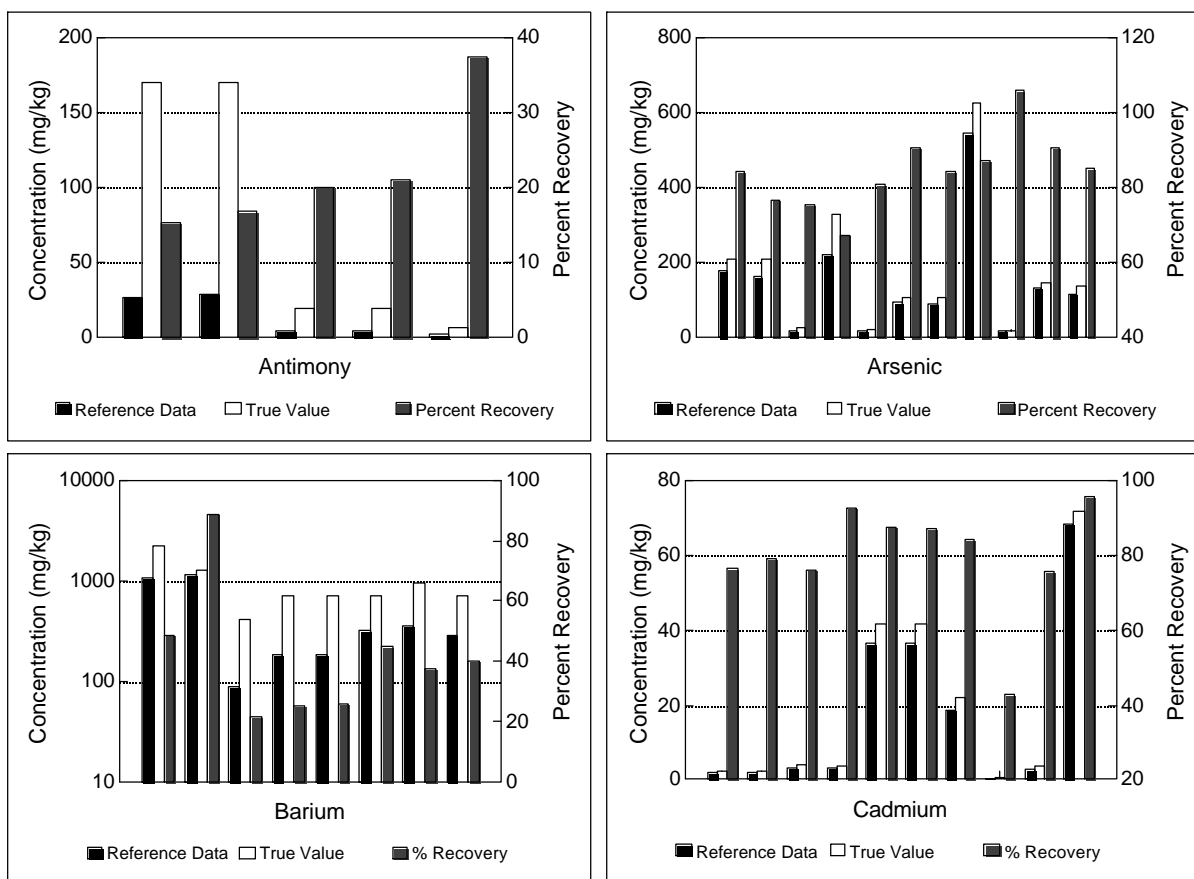
Figure 3-3. Reference Method SRM Results: These graphs illustrate the relationship between the reference data and the true values for the SRM samples. The gray bars represent the percent recovery for the reference data. Each set of three bars (black, white, and gray) represents a single SRM sample.

39
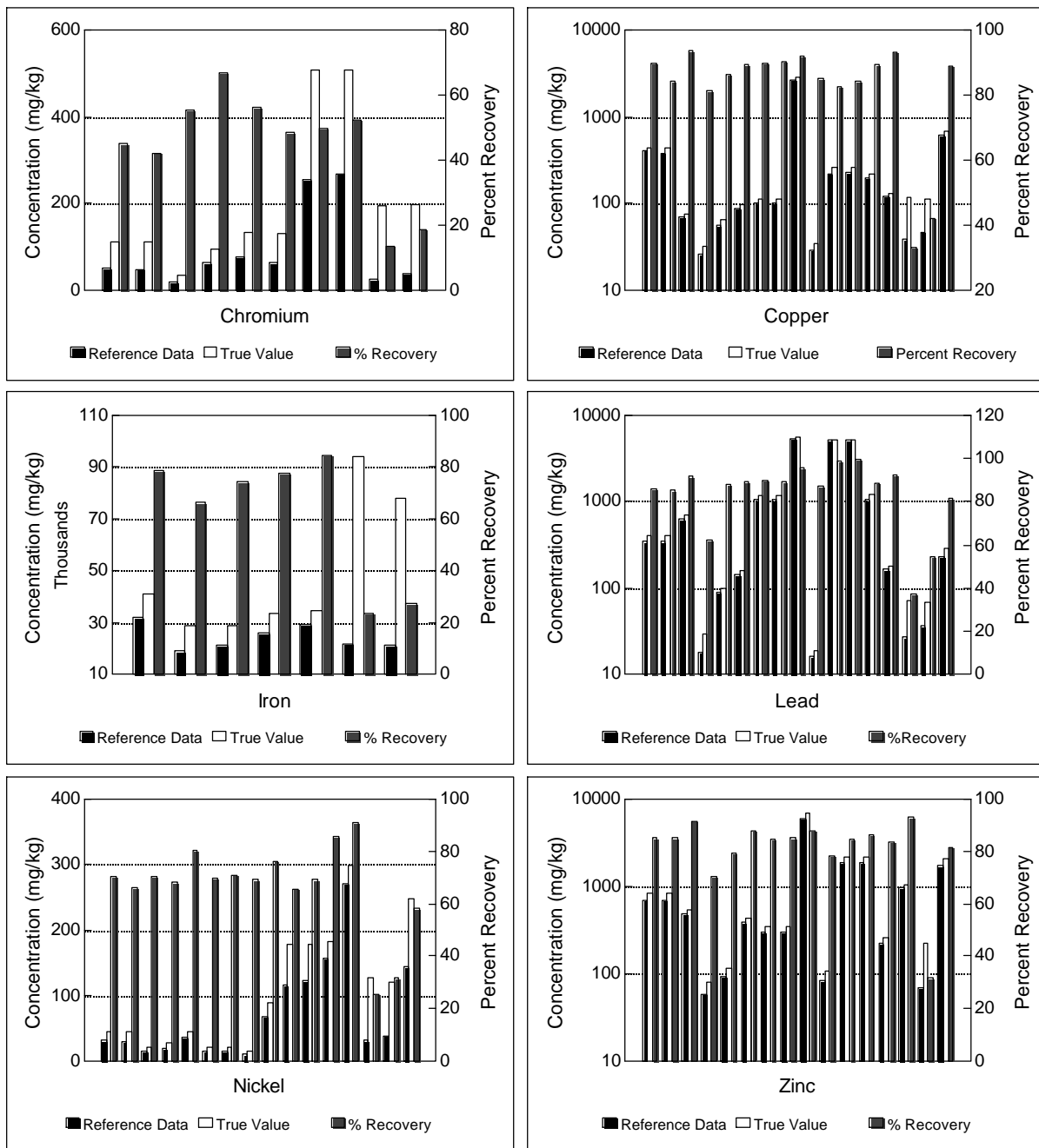
Figure 3-3 (Continued). Reference Method SRM Results: These graphs illustrate the relationship between the reference data and the true values for the SRM samples. The gray bars represent the percent recovery for the reference data. Each set of three bars (black, white, and gray) represents a single SRM sample.

Section 4
X-MET 920-P and 940 Analyzers

This section provides information on the Metorex's X-MET 920-P and 940 Analyzers including 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 outer shells. Electrons in these outer shells have higher potential energy states than inner shell electrons, and to fill the vacancies, the outer shell electrons give off energy as they cascade into the 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-rays emitted is proportional to the concentration of the source and can be used to quantitate each target 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 $K_ß$-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 $K_ß$ transition. The K_α-line is approximately 10 times more intense than the $K_ß$-line for a given element, making the K_α-line analysis the preferred choice for quantitation purposes. Unlike the K-lines, the L-lines (L_α and

41

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.

Excitation X-ray from the FPXRF Source

An excited electron is displaced, creating an electron vacancy.

K Shell Electrons

Nucleus

An outer electron shell electron cascades to the inner electron shell to fill the vacancy. As this electron cascades, it releases energy in the form of an X-ray.

L Shell Electrons

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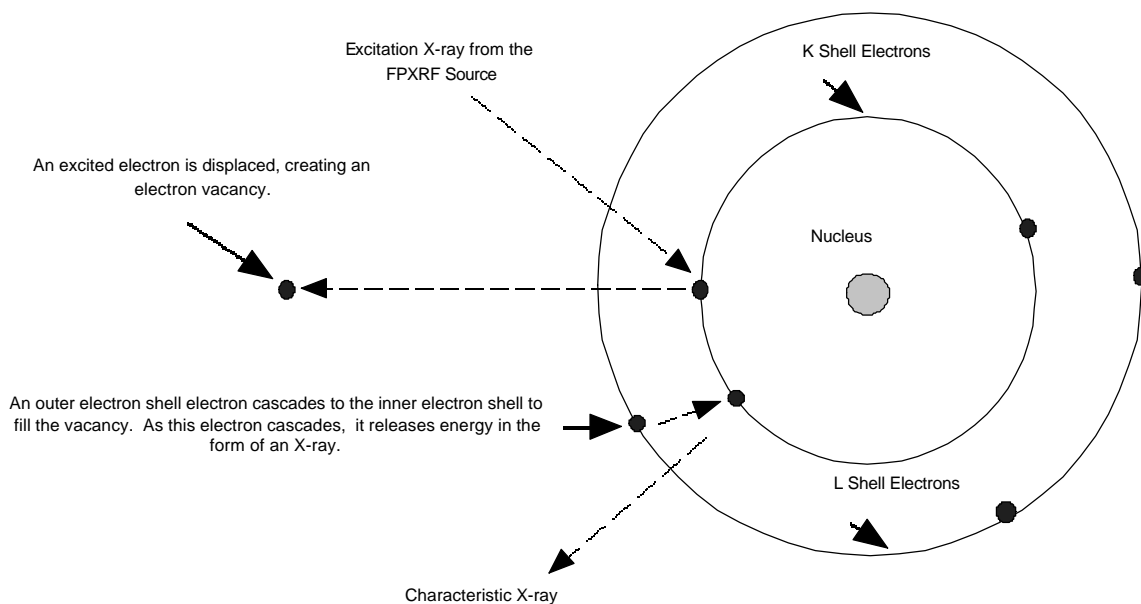
Figure 4-1. Principle of Source Excited X-ray Fluorescence: This figure illustrates the dynamics of source excited X-ray fluorescence.

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 is 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 kiloelectron 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

Metorex is an international supplier of advanced equipment for metal detection, materials testing, and chemical analysis. It offers a wide range of products from field portable and laboratory-grade metals and alloy analyzers to in-line (process) analyzers. It has more than 20 years of experience in developing X-ray analyzer technologies.

Metorex developed the X-MET 920 line as a modular system, which users can configure for their specific analytical problem. This line includes the X-MET 920-P and 940 analyzers.

The basic configuration of the analyzers includes: a disk operating system-based personal computer (PC) with Metorex electronics added in the form of a board that plugs directly into the computer expansion slot identified as the X-MET PC System (XPCS), Metorex proprietary XRF software, and the Metorex XRF analysis probe. The data acquisition, storage, and processing systems are housed in the XPCS module independent of the analysis probe which contains excitation sources as well as the X-ray detector. The X-MET 920 line uses IBM compatible computers so that the analyzers can be configured with desktop, laptop, or portable computers. The XPCS contains a 2,048 multichannel analyzer (MCA) that is used to collect the spectra. The MCA portion of the technology is contained on a single electronic board that is plugged into one of the expansion slots of the XPCS. Menu-driven software guides the user through routine analysis and calibration, and is used to generate and display the data.

With the X-MET 920 line, Metorex offers several analytical probes. The probes contain the excitation sources and either a silicon (drifted with lithium) Si(Li) detector or a gas-filled proportional counter detector.

Metorex developed the X-MET 920-P and X-MET 940 to perform elemental analysis in the petroleum and petrochemical industry, the mining and minerals industry, and the environmental field. The X-MET 920-P and 940 are field portable technologies that can be operated in the *in situ*, or intrusive mode. At the time of this demonstration, the X-MET 920-P was commercially available; the X-MET 940 was a prototype. The X-MET 920-P and 940 are Metorex's "portable" versions on the X-MET 920 line. The X-MET 940 is a smaller, lighter version of the X-MET 920-P. Other than the difference in the size and weight of the XPCS, these instruments are identical in function. Both analyzers use energy dispersive XRF spectroscopy to determine the elemental composition of soils and other solid waste materials as well as liquids and slurries. The X-MET 920-P and 940 can identify and quantify the concentrations of 70 elements, 32 of which can be identified and quantified simultaneously. Metorex offers three excitation sources, Fe^{55}, Cd^{109}, and Am^{241}, and two detectors (Si(Li) and gas-filled proportional counter) in its various probes. For this demonstration, both analyzers used the "solid state probe system" (SSPS) equipped with the Cd^{109} and Am^{241} sources and a Si(Li) detector. The SSPS is designed to house two excitation sources. The choice of sources is determined by the user depending on the target analytes.

For *in situ* analyses using either analyzer, the probe is pointed downward and placed flat against the soil surface to allow the probe window to come into contact with the soil surface. For intrusive analyses, the probe is pointed upward and a protective sample cover is attached over the probe window. The soil samples are placed in sample cups and the sample cups are placed inside the protective sample cover for analysis. In either mode, sample measurement is initiated by depressing a button located underneath the handle at the window end of the probe. This exposes the sample to primary radiation from the two excitation sources, sequentially: fluorescent and backscattered radiation reenters the probe through the source-detector window and is counted by the Si(Li) detector, which passes on the electronic pulses to the MCA for processing.

The X-MET 920-P and 940 can be operated and calibrated using either the X-MET software or Automated Contaminant Evaluation Software (ACES). The X-MET software is used to analyze samples using an empirical calibration. The ACES software is used to analyze samples with a fundamental parameters (FP) calibration. During this demonstration, the ACES software was used to run the X-MET 920-P and 940. The X-MET 920-P was used to analyze all samples from the ASARCO site and about half of the samples from the RV Hopkins site. The X-MET 940 was used to analyze the *in situ* samples at the RV Hopkins site. Equipment failure precluded its use for all the RV Hopkins samples.

Operational Characteristics

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

Equipment and Accessories

The X-MET 920-P and 940 come with all of the accessories necessary for both *in situ* and intrusive operation. Metal, foam-padded carrying cases are provided for storage of the XPCS units and the analysis probe. Specifications for the X-MET 920-P and the 940 used during this demonstration are provided in Tables 4-1 and 4-2, respectively.

Table 4-1. X-MET 920-P Instrument Specifications

Characteristic	Specification
Resolution	170 eV (Mn-K_α)
Sources	20 mCi Cd^{109} and 30 mCi Am^{241} (Fe^{55} also available)
Detector	Si(Li) - Liquid nitrogen cooled (gas-filled proportional also available)
Analyzer Size	20 cm x 10 cm x 25 cm
Probe Weight	4.1 kilograms (with liquid nitrogen)
Computer Size	38 cm x 45 cm x 9 cm
Computer Weight	5.5 kg (without external battery); 11.5 kg (with external battery)
Computer Configuration	Central processing unit 386SX 25 megahertz with a math coprocessor; 2 MB RAM; 80 MB hard disk; VGA graphics; color screen; 3.5-inch floppy disk; two serial ports and a printer port; 2,048 channel MCA
Power Source	120V (AC), internal or external batteries
Operational Checks	Pure lead check sample
Intrusive Operation	Attachment of protective sample cover
Contact: Jim Pasmore 1900 NE Division Street Bend, OR 97701 1-800-229-9209 (503) 385-6750 (FAX)	

Three main components comprise the analytical system of each instrument: the computer, electronics unit, and probe. With the X-MET 920-P, the PC and 2,048-channel MCA are packaged together in a weatherproof plastic enclosure. The X-MET 920-P features a handle which swings to the back of the unit to hold it at an angle for use on the ground or on a table. The keyboard is covered with a form-fitted soft rubber cover to prevent water from entering the unit. All ports are fitted with watertight caps. The bottom of the unit is fitted with grooves that accept an external battery pack for use in the field. The unit also contains an internal battery that is capable of powering the system for a short time. The X-MET 920-P has been given a National Electrical Manufacturers Association No. 4 enclosure rating. The external battery for the X-MET 920-P slides onto the bottom of the unit. A short power cord that is permanently attached

to the battery case plugs into the side of the X-MET 920-P. The battery is the same width and depth as the X-MET 920-P and is about 5 cm (2 inches) in height. The external battery weighs 6 kilograms (kg) (13 pounds). During the demonstration, the battery lasted about 5 to 7 hours when fully charged. To fully charge, the external battery is plugged into the combination AC adaptor and battery charger overnight. The combination AC adaptor and battery charger can be used to simultaneously power the X-MET 920-P and charge a battery.

Table 4-2. X-MET 940 Instrument Specifications

Characteristic	Specification
Resolution	170 eV (Mn-K_α)
Sources	20 mCi Cd^{109} and 30 mCi Am^{241} (Fe^{55} also available)
Detector	Si(Li) - Liquid nitrogen cooled (gas-filled proportional also available)
Analyzer Size	20 cm x 10 cm x 25 cm
Probe Weight	4.1 kilograms (with liquid nitrogen)
Computer Size	38 cm x 36 cm x 5 cm
Computer Weight	4.7 kg (without external battery); 6.4 kg (with external battery pack)
Computer Configuration	Central processing unit 386SX 25 megahertz with a math coprocessor; 4 MB RAM; 10 MB flash disk; CGA graphics; a monochrome screen; connector for an external keyboard; serial, printer, and external floppy disk drive ports; 2,048 channel MCA
Power Source	120V (AC), internal or external battery pack
Operational Checks	Pure lead check sample
Intrusive Operation	Attachment of protective sample cover
Contact: Jim Pasmore 1900 NE Division Street Bend, OR 97701 1-800-229-9209 (503) 385-6750 (FAX)	

The X-MET 940, which was used to analyze a portion of the samples at the RV Hopkins site, is a smaller, lighter version of the X-MET 920-P. It also features a handle which swings to the back of the unit to hold it at an angle for use on the ground or on a table. The keyboard on this unit was a prototype version of what will be found on production models. One feature of its layout was the addition of several keys to the right of the screen. These keys, which are separate from the rest of the keyboard, are used to operate some of the basic functions in taking measurements. All ports are fitted with watertight caps. The unit contained an internal battery that is capable of powering the unit for up to 2.5 hours. A nylon carrying bag with shoulder straps was provided for analyses in the field. Metorex provided a prototype external battery for the X-MET 940. In the production model, Metorex plans to design the battery to slide onto the bottom of the unit as with the X-MET 920-P. The prototype model, however, did not have this feature. A short power cord, which is permanently attached to the battery, plugs into the side of the X-MET 940. The battery is the same width and depth as the X-MET 940 and is about 3.8 cm (1.5 inches) in height. The prototype external battery weighed about 2 pounds. During the demonstration, each battery lasted at least

2.5 hours when fully charged. To fully charge, the prototype external battery needed to be plugged into the X-MET 940 while the unit was plugged into an AC outlet. The prototype battery required about 8 hours to fully charge.

The high resolution probe used with both analyzers is the solid state probe system (SSPS); it is a hand-held, compact unit that contains a Si(Li) detector. The detector achieves a manganese K_α X-ray resolution of 170 eV. The detector is cooled by a 0.5-liter liquid nitrogen dewar built into the probe, which allows for 8 - 12 hours of field use. A dewar is similar to a Thermos® except that it is used to store super-cooled liquids such as liquid nitrogen. This probe can be used as a surface probe to perform *in situ* analyses by pointing downward at the soil surface, or with the attachment of a protective sample cover, the probe can analyze soil samples intrusively from a sample cup. The protective sample cover is easily removed or mounted on the probe with two hex-key screws. The probe dewar can be filled with liquid nitrogen before each measurement session or can remain connected to the "mother" dewar (30-liter capacity). During this demonstration, PRC supplied the 30-liter dewar; however, a dewar can be purchased or leased from Metorex. Metorex provided a funnel assembly to fill the probe dewar from the mother dewar. Metorex provided a leather glove and safety glasses for protection from liquid nitrogen when it disconnected the funnel assembly from the probe dewar. Approximately 10 minutes is required to fill the probe dewar with liquid nitrogen and an additional 45 minutes to allow the Si(Li) detector to cool and stabilize each day prior to use.

The probe used during the demonstration contained the dual radioactive source configuration of a 20-millicurie (mCi) Cd^{109} and 30 mCi Am^{241} source. This is the most common configuration of sources for environmental work. The probe can only contain a total of two sources. These radioactive sources are in the form of an 8-mm-diameter by 5-mm-thick capsule. Both sources in the probe used during the demonstration were installed in February 1995. The sources are pneumatically driven and shifted into the measurement position using pressure produced from liquid nitrogen off-gassing in the probe's dewar. This off-gassing creates a pressure of 10 pounds per square inch (psi). The probe is equipped with an interlock mechanism to prevent operator exposure to the radioactive sources. When not in the measurement position, the radioactive sources are retracted into a tungsten shield.

One side of the probe contains a window, which displays a red color when the Cd^{109} source is in the measurement position and a green color when it is retracted into its tungsten shield. The other side of the probe contains a similar window that indicates the position of the Am^{241} source. The probe contains two buttons. One button, which is located underneath the handle at the window end of the probe, begins a measurement. The other button, a reset button located beneath the other end of the handle, causes the exposed source to be retracted into its tungsten shield. The probe's sample window is environmentally sealed by a 25-mm-diameter window of clear Kapton™ film.

The probe is connected to the analyzers with a coiled, flexible cord that is about 5 or 6 feet long. When fully extended, the cord reaches about 15 feet. The cord is permanently attached to the probe. Metorex provided a cardboard box with a fitted, protective foam insert for storing and shipping the probe. Wipe test results for the probe's sources were shipped with the analyzers. Semiannual wipe tests are required to monitor for source leakage.

Metorex brought pure element standards to the training sessions at the ASARCO site in Tacoma, Washington, and to the RV Hopkins site in Davenport, Iowa. To perform either an empirical calibration or a FP calibration, a pure element standard for each target analyte is required. These pure element standards are in the shape of a coin about the size of a $0.50 piece. They are large enough to completely cover the probe's Kapton™ window. Other equipment and supplies that are helpful when using the

X-MET 920-P or 940, which is not supplied by the developer, include protective gloves, paper towels, and a permanent marking pen.

Operation of the Analyzers

Three steps were involved in operating the analyzers: (1) performing the calibration, (2) taking measurements, and (3) managing data. The X-MET 920-P and 940 can be calibrated empirically or through the use of FPs. An FP calibration was performed using the ACES software. There were three steps in calibrating the analyzers using the ACES software.

In the first step, the operator inputs concentration data from one fully-characterized sample, which is referred to as the standard sample. Metorex recommends that the standard sample be collected at the site of interest. However, an NIST or similar standard could be used if a site-specific sample has not been analyzed. The more accurate the data for the standard sample, the better the results. At both sites, Metorex helped the operator decide which concentrations to use for these analytes. The same computer screen which contains the standard sample data will also display the abbreviated chemical names of the target analytes, the source to be used to excite each target analyte, and the peak (K_α, L_α, L_β, etc.) to be used for quantification. This step requires some knowledge of XRF theory to properly select the appropriate source and peak. This step of choosing the source and peak is no longer necessary in the latest ACES software Version 2.0.

In the second step, the operator corrects for any X-ray spectral overlaps for the target analytes. This is done by acquiring spectra from pure element standards for all of the target analytes. Once the spectra of pure elements are acquired, the software prompts the operator to create overlap tables. This requires highlighting the target analytes on a periodic table and pressing a key to begin the calculation of the overlap tables.

The final step involves analyzing the standard sample to generate response factors for each target analyte. Metorex suggested a count time of 200 seconds for each source to acquire these spectra. Once the three steps were completed, the analyzer is calibrated and ready to use. The operator noted that setting the correction factors were a trial-and-error process that would become simpler with experience.

Background of the Technology Operator

The operator chosen for analyzing soil samples using the X-MET 920-P and 940 has been a PRC employee for more than 4 years. He holds a bachelor's degree in biology and a minor in chemistry and a master's degree in environmental science. While at PRC, he has worked on projects involving site investigation and hazardous waste management. Prior to working for PRC, he spent 1 year working as an environmental scientist for the Site Investigation Section of the Indiana Department of Environmental Management, where he conducted preliminary assessments and site investigations of potential hazardous waste sites for the Superfund program.

Training

The operator received 2 days of training by the developer at the start of the demonstration. Approximately a half day of the training was dedicated to the theoretical background of XRF and the remaining 1.5 days were spent on specific operation and hands-on training for the X-MET 920-P. The hands-on training covered test measurements using the pure element samples, empirical calibrations, FP

calibrations, the analysis of various standards and soil samples, and the steps for saving spectra and quantitative results. The developer indicated that the standard training course is 3 days. The developer stated that the PRC operator received more one-on-one instruction in 2 days than users normally receive in a standard 3-day course. The course is designed for users with experience ranging from little or no scientific background to scientists with many years of experience. Metorex tailors its training course to match the level of user experience and the primary intended use of its equipment.

Two Metorex representatives stayed at the ASARCO site throughout the first day of *in situ* analysis and assisted the PRC operator that evening with calculating concentrations and downloading the spectra and results files to a floppy disk. They departed the next morning. One Metorex representative came to the RV Hopkins site for 2 days prior to the start of the field work to review the procedures with the operator.

Reliability

More than 1,800 individual measurements were collected with the X-MET 920-P and 940 analyzers using the high-resolution probe and the Si(Li) detector. This included the measurement of soil samples using the four sample preparation steps (1,260 measurements), 10 replicate measurements on 48 samples for a precision assessment (480 measurements), the measurement of QC samples such as blanks, PE samples, and SRMs, and multiple measurements of the same samples while varying count times. While collecting these measurements over a period of 20 working days, several operational problems were encountered. These problems can be divided into two categories: (1) mechanical, or (2) electronic or software. These problems are discussed below.

The first mechanical problem was encountered with the probe of the X-MET 920-P at the ASARCO site. During the third day of field work at this site, it rained much of the day and the probe got wet. Throughout the morning, one of the two pressure release valves on the probe dewar kept frosting up. Eventually, the probe stopped functioning. Thinking that the pressure in the dewar was getting low, the operator refilled the probe dewar with liquid nitrogen. After waiting for the probe to stabilize, the operator again began taking measurements. Within a few minutes, the pressure release valve began releasing nitrogen gas because the valve could not close completely due to the frost. The measurement stopped after the LED warning light came on. The operator refilled the probe dewar with liquid nitrogen, but the problem occurred once more. A call was then made to Metorex technical support for assistance. It was suggested that the operator cover the pressure release valves with a plastic bag to prevent condensation from the humid environment. The operator did this and was able to continue taking measurements. Metorex also suggested that, in the evening, the operator use a hair dryer to dry the pressure release valves.

On the morning of the fourth day of field work at the ASARCO site, a small brass nut fell out of the probe dewar when the operator opened the dewar for filling. The operator again called Metorex for assistance. The operator was instructed to remove the pressure release valves with a wrench and replace the small nut. When the operator removed the valves, he saw that a nut was missing. The operator put the nut on the end of the piston and replaced the pressure release valve. This was an easy procedure that took about 10 minutes. It is possible that this missing nut caused some of the problems encountered during the afternoon of the third day of field work.

The next problem occurred at the second sample location on the fourth day, when the data acquisition process stopped. After trying several troubleshooting measures, it was concluded that the dewar was holding pressure but there was a problem with the source logic within the software. The operator informed Metorex that there was some moisture under the Kapton™ window and on the Si(Li) detector that had

accumulated from the day before while measurements were being taken in a light rain. Metorex informed the operator that there was a photosensor beneath the window that controls source logic, which in turn affects the switching of sources during a measurement. If water droplets accumulated on the photosensor, the computer software system would be unable to tell which source was exposed and the probe would stop functioning. Metorex suggested placing the probe next to a heater to evaporate the moisture from under the Kapton™ window.

To remedy this problem, the probe was placed next to a portable heater. After several hours, moisture accumulated on the underneath side of the window but could not escape. It was decided to cut the Kapton™ window out of the probe to let the condensation escape. The following morning, the fifth day of field work at the ASARCO site, the inside of the probe was dry. The operator then placed a sheet of Mylar XRF film over the window since the Kapton™ window had been removed. This solved the problem and the probe functioned properly for the remainder of the demonstration. These problems resulted in one day of downtime and prevented the PRC operator from collecting data for 67 *in situ* samples at the ASARCO site.

While using the prototype model of the X-MET 940 for three days at the RV Hopkins site, the PRC operator encountered three problems. The first minor problem was that the number "1" key on the X-MET 940's keypad did not work. Metorex had anticipated keypad problems with its prototype model and had brought an external keypad to use with the X-MET 940. The operator connected the external keypad to the X-MET 940 and the problem was solved.

A software problem occurred during the late afternoon on the third day of using the X-MET 940. During an analysis, the measurement stopped after the probe switched from source A, the Cd^{109} source, to source B, the Am^{241} source. The following message appeared on the screen: "Unable to proceed with this operation. Disk write error. Please press enter." The operator started the measurement again, but the same message appeared. He repeated the measurement two additional times, but the same message appeared. The PRC operator then tried to calculate concentrations from spectra acquired up to that point. The same message appeared on the screen. The operator ran Norton Disk Doctor (NDD) on the X-MET 940 but was unable to correct the problem. The PRC operator called Metorex, but Metorex was unable to recommend a fix for the problem over the phone. The developer instructed the operator to return the X-MET 940 to Metorex. After Metorex received the X-MET 940 and tried to run the ACES software, it diagnosed that the X-MET 940's hard drive was full. Metorex informed the PRC operator that each time the X-MET 940 is shut off, a file is created that stores information about the X-MET 940 and the probe. After three days of using the X-MET 940, these self-creating files had used up all of the memory on the hard drive. Metorex said it would change the error message from "Disk write error" to something like "hard drive full." Metorex also said that it would inform future technology users of this self-creating file and how to remove it. No data was lost as a result of this problem because the PRC operator did not save any spectra or results files on the hard drive.

When the X-MET 940 was returned to Metorex, the operator resumed RV Hopkins sample analyses with the X-MET 920-P. When the RV Hopkins calibrations were set up on the X-MET 920-P, correction factors for zinc and nickel produced erroneous data for these analytes. All of the RV Hopkins samples analyzed by the X-MET 920-P were biased high for nickel and zinc by approximately 600 mg/kg and 800 mg/kg, respectively. This bias is shown in Figure 4-2. The cause of this error could not be established. This problem caused the loss of zinc and nickel data for samples 380 - 399 for the *in situ*-prepared analysis, and all of the intrusive zinc and nickel data at the RV Hopkins site. This data loss was not considered in the discussion of analyzer data completeness, since this did not reflect an analyzer-induced data loss.

When the operator began acquiring spectra with the X-MET 920-P at the RV Hopkins site, several error messages appeared. One of the messages read: "Sector not found reading drive C." NDD was used to diagnose the problem and encountered several bad sectors and marked them so they would not be used. Similar problems were encountered during the following 2 days. Eventually, enough bad sectors were encountered on the X-MET 920-P's hard drive that they could not be fixed with NDD.

Metorex sent a new X-MET 920-P XPCS unit, the PRC operator removed the MCA card from the nonfunctional X-MET 920-P, and placed it in the new X-MET 920-P. This was a relatively easy process that took about 30 minutes. The operator then copied the software and all files containing the ASARCO calibration onto the new X-MET 920-P. This was an easy process that took about 30 minutes. The operator then finished analyzing all the samples without any problems.
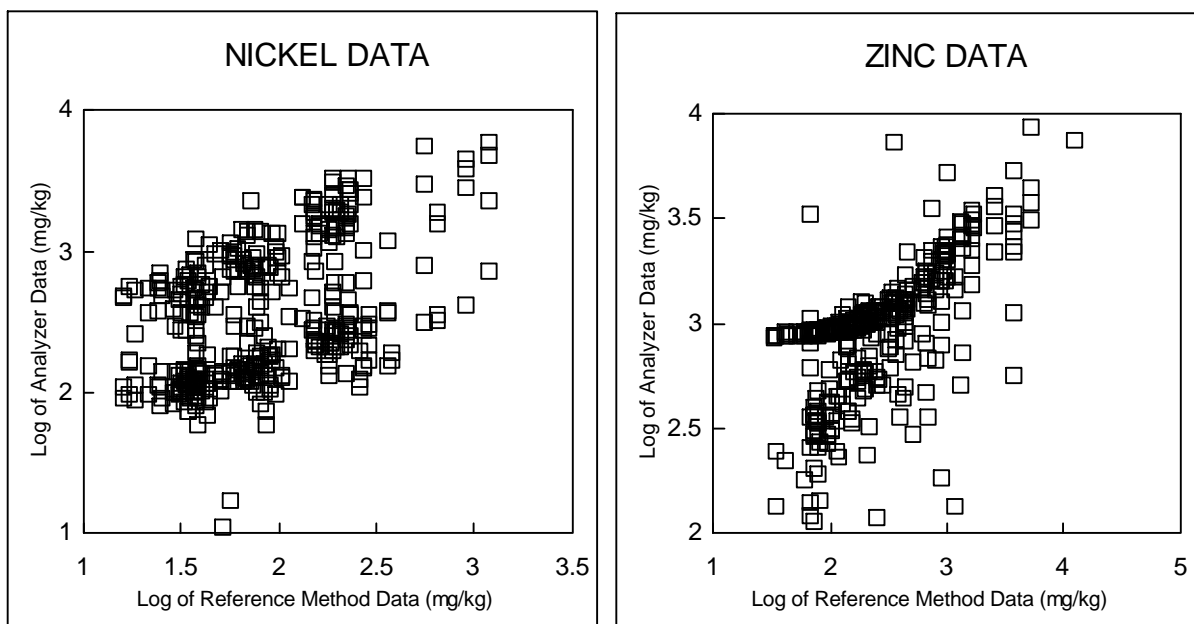
Figure 4-2. Analyzer Comparison Data: These graphs illustrate the shift in analyzer data that occurred when the X-MET 940 was replaced with the X-MET 920-P at sample number 380 during the in situ-prepared analysis. The nickel data appears to have generally increased by approximately 600 mg/kg and the analyzer's zinc data clusters around 1,000 mg/kg over a wide range of reference data concentrations.

Health and Safety

The potential for exposure to radiation from the excitation source was the largest health and safety consideration while using the analyzers. The X-MET 920-P and the 940 are sold under a general license, meaning that the analyzers are designed and constructed in such a way that while in use, per the instruction manual, an operator would not accumulate a radiation dose higher than that from naturally occurring radiation. One objective of the demonstration was to evaluate radiation exposure to operators from the analyzers. Radiation was monitored with a gamma-ray detector radiation survey meter. Background radiation at the two sites was between 0.006 and 0.012 millirems per hour (mrem/hr). Radiation was monitored in the *in situ* and intrusive modes while the probe's source was exposed (during a measurement), obtaining a worst-case scenario. The radiation was measured within 5 cm of the probe face while analyzing a sample. Radiation exposure also was monitored at a point on the probe where the operator's

hand was located during analysis to provide a realistic value of operator exposure. Although the X-MET 920-P and the 940 are sold under a general license, meaning that they have met all safety requirements according to the Nuclear Regulatory Commission, many states still recommend that radiation from survey instruments be below a certain level. For example, in the State of Kansas, the permissible occupational exposure is 5,000 millirems per year, which equates to approximately 2 to 3 mrem/hr assuming constant exposure for an entire work year.

While taking *in situ* measurements (probe pointing down), maximum radiation values of 0.15 - 0.20 mrem/hr and 1.2 to 1.5 mrem/hr at the probe face were obtained with the Cd^{109} source and Am^{241} source exposed, respectively. The radiation values dropped to 0.020 - 0.030 mrem/hr at the probe handle when either of the two sources was exposed. While collecting intrusive measurements (probe pointing up), radiation values of 0.15 mrem/hr at the side of the protective sample cover and 0.015 - 0.020 mrem/hr above the protective sample cover were obtained with the Cd^{109} source exposed. With the Am^{241} exposed, radiation values of 2.3 to 2.5 mrem/hr at the side of the protective sample cover and 0.015 - 0.020 mrem/hr above the protective sample cover were obtained. Background radiation levels were obtained at the trigger under the handle with either source exposed. All radiation values at the probe handle were below the occupational level of 2.0 mrem/hr.

Transferring liquid nitrogen from an external dewar to the internal dewar of the SSPS used for both the X-MET 920-P and 940 was another health and safety consideration. Due to the extremely low temperature of liquid nitrogen, the operator must take care to avoid contact during the filling operation. Safety goggles and gloves must be worn during this process. It is also recommended that a laboratory coat be worn when filling the SSPS with liquid nitrogen.

Cost

The primary cost benefit of field analysis is the quick access to analytical data. This allows the process depending on the test results to move efficiently to the next stage. Costs associated with field analysis are very dependent on the scope of the project. Since most of the mobilization costs are fixed, analyzing a large number of samples usually 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 finding may be resolved completely without the need to return for another sample collection event.

A representative list of costs associated with the Metorex 920-P is presented in Table 4-3. 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 in this table.

Since it was tested as a prototype, no-cost data was available for the X-MET 940. At the time of the demonstration, the X-MET 920-P and SSPS with two radioisotope sources cost $52,470. This includes the X-MET and ACES software, eight pure element standards, a liquid nitrogen dewar, and 3 days of training for two people at Metorex. Travel and accommodation costs for the training are not included. Spare batteries are available for $425 and spare battery chargers are available for $340. Periodic maintenance includes replacement and disposal of the Cd^{109} source every 2 years at a cost of $4,500 with an additional $500 disposal fee.

The X-MET 920-P can be rented from Metorex. There is a 1-month minimum rental. The cost is 10 percent of the purchase price per month, and all shipping costs. Users have a choice of training options.

The first is a 3-day class offered at Metorex's facility at $685 per person plus travel and lodging expenses. On-site training classes are also available. Metorex must be contacted for details regarding the on-site training classes.

Table 4-3. Instrument and Field Operation Costs

Item	Amount	
X-MET 920-P	$ 52,740	Purchase Price
	5,500	Per Month Lease
Replacement Source	5,000	For Cd^{109}
Operator Training (Vendor Provided)	685	—
Radiation Safety License	500	(State of Kansas)
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	10 - 12	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 of site-specific soil samples with metals concentrations 2 to 5 times the expected MDLs. These data were obtained during the measurement of instrument precision. Based on this precision data, a standard deviation was calculated and the MDLs were defined as 3 times the standard deviation 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 X-MET 920-P and 940 are shown in Table 4-4.

The precision-based MDLs were obtained using a 140-second count time for the Cd^{109} source and a 100-second count time for the Am^{144} source. Table 4-4 also lists MDLs reported by the developer. The developer's MDLs were acquired using a 200-second count time for each source. The counting statistics for FPXRF analysis indicate that it would take a fourfold increase in count time to increase the precision and therefore reduce MDLs by 50 percent.

Another method of determining MDLs involves the direct comparison of the log$_{10}$ transformed FPXRF data and the log$_{10}$ transformed reference data. When these sets of data were plotted against each other, the resultant plots were linear. As the plotted line approached zero for either method, there was a point at which the FPXRF data appeared to respond to the same reading for decreasing concentrations of the reference data. Figure 4-3 illustrates this effect for copper. This point was determined by observation and

was somewhat subjective; however, a sensitivity analysis showed that even a 25 percent error in identifying this point resulted in up to 10 percent changes in MDL calculation. By determining the mean values of this FPXRF data and subsequently two standard deviations (SD) around this mean, it was possible to determine a field or performance-based MDL for the analyzer. For the X-MET 920-P and 940 analyzers, these field-based MDLs are also shown in Table 4-4.

Table 4-4. Method Detection Limits—X-MET 920-P and 940 Analyzers

Analyte	Developer-based MDL (mg/kg)	Precision-based MDL (mg/kg)	Field-based MDL (mg/kg)
Arsenic	75	55	90
Barium	20	30	1,320
Chromium	200	210	470
Cadmium	25	25	55
Copper	80	75	210
Iron	85	Not determined	Not determined
Lead	60	45	45
Nickel	80	120	200
Zinc	75	70	120

Note: mg/kg Milligrams per kilogram
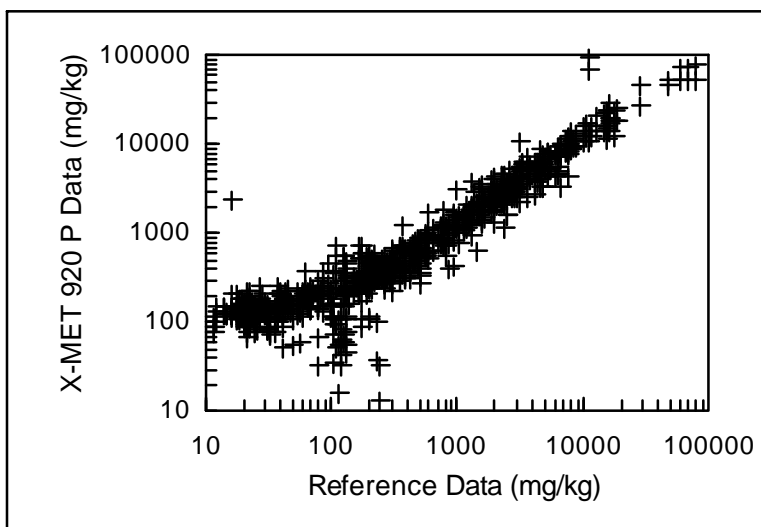
Figure 4-3. Critical Zone for the Determination of a Field-based Method Detection Limit for Copper: Between 100 and 200 mg/kg for the reference data, the linear relationship between the two data sets changes. This point of change identified the point at which field-based MDLs for the analyzer were determined.

Because iron was generally present at tens of thousands of milligrams per kilogram, reasonable field-based detection limits could not be calculated. For the other elements, the precision-based MDLs were very

similar to the developer MDLs. The field-based MDLs for most of the analytes were 1.5 to 3 times higher than the developer's or precision-based MDLs. The field-based and precision-based MDL for lead were the same.

Antimony, a secondary analyte for this demonstration, was not reported by this analyzer, although it is in the analyzer's range and can be determined. At the time of the demonstration, a standard with a certified value for antimony was unavailable. Antimony can be analyzed by the X-MET 920-P and the 940 when they are equipped with an Am^{241} source.

Throughput

Both analyzers used a total source live-second count time of 240 seconds. With the additional "dead" time of the detector and the time required to label each sample and store data between sample measurements, the time required to analyze one soil sample was 5 to 6 minutes. This resulted in a throughput of approximately 10 to 12 samples per hour in the intrusive mode. The throughput for the *in situ* samples was 8 to 10 measurements per hour which includes the time required to walk to the next sample location.

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 an average of approximately 5 minutes per sample (*in situ*-prepared), 20 minutes per sample was required for No. 10 sieving (intrusive-unprepared), and 10 minutes per sample was required for grinding and sieving (intrusive-prepared). The operator noted that it took about 10 - 15 minutes to fill the probe dewar with liquid nitrogen and 45 minutes for the probe to cool and stabilize. This time was used to calculate concentrations from spectra collected the previous day and to perform data management tasks such as printing a hard copy of the data. On average, it took about 1 hour after arriving at the site before the operator could start taking measurements.

Drift

Drift is a measurement of an analyzer's variability in quantitating a known amount of a standard over time. Normally drift is evaluated by reviewing results from the periodic analysis of an SRM or other check samples. No data was produced by either analyzer to assess drift.

Intramethod Assessment

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

Blanks

Analyzer blanks consisted of pure lithium carbonate. The blanks were placed directly in a sample cup after all four sample preparation steps like the dried and ground soil samples. The blanks were used to monitor for contamination of the probe by material such as residual soil left on the probe face. Four blanks were analyzed at the ASARCO site using the ASARCO site-specific FP calibration model. Four blanks also were analyzed at the RV Hopkins site using the RV Hopkins site-specific FP calibration model. Iron was detected in all blanks at levels ranging from 237 to 304 mg/kg. These concentrations were above the

MDL listed by the developer; however, they were not of concern as most of the soil samples contained iron concentrations exceeding 20,000 mg/kg. Copper also was detected in all blanks at levels ranging from 85 to 103 mg/kg. These levels were slightly above the developer and precision-based MDLs, but were below the field-based MDLs.

Completeness

The analyzers produced data for 1,192 out of the 1,260 samples for a completeness of 94.6 percent, slightly below the demonstration objective of 95 percent. Data was not obtained for 68 samples at the ASARCO site. The mechanical problems with the probe while using the X-MET 920-P caused the loss of data for 67 *in situ*-unprepared samples, most of which were in the loam textured soil. The remaining missing data point was for one *in situ*-prepared ASARCO soil sample. This omission was caused by an oversight of the operator.

Precision

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

These analyzers 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 the four sample preparation steps. Therefore, there was a total of 48 precision points for the analyzers. The replicate measurements were taken using the same source count times used for regular sample analysis. For each detectable analyte in each precision sample, a mean concentration, SD, and RSD were calculated. The percent RSD is defined as the SD divided by the mean concentration times 100.

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 results, reflected by the precision data in the 5 to 10 times MDL range, were all below the 10 percent RSD required for definitive level data quality classification with the exception of chromium and nickel. Cadmium and iron did not have sufficient data to allow data quality conclusions based on precision. Table 4-5 shows that chromium and nickel precision was slightly greater than 20 percent, placing the chromium and nickel results in the qualitative level data quality classification based solely on precision. The lower precision for chromium was expected because chromium is a problematic analyte in FPXRF analysis. The lower precision for nickel may be an artifact of the low concentrations of nickel in the soil samples and the influence of iron interference in the samples.

There was no observable effect of sample preparation on precision. This was expected because the method used to assess precision during this demonstration assessed analyzer precision and not total method precision. These samples were purposely chosen to span a large concentration range to test the effect of analyte concentration on precision. As the concentration of the target analyte increased, the precision improved. Figure 4-4 shows an asymptotic relationship between concentration and precision. In this figure, precision shows little improvement at concentrations greater than 400 mg/kg; however, at concentrations below 400 mg/kg, precision is highly concentration dependent. Although lead is shown in Figure 4-4, a similar trend was exhibited by the other target analytes. Table 4-5 shows that the RSD

values were less than 10 percent for all analytes except chromium and nickel at concentrations greater than 500 mg/kg.

Table 4-5. Precision Summary—X-MET 920-P and 940 Analyzers

Analyte	Mean % RSD Values by Concentration Range			
	5 - 10 Times MDL[a] (mg/kg)	50 - 500 (mg/kg)	500 - 1,000 (mg/kg)	>1,000 (mg/kg)
Arsenic	3.36 (8)	10.48 (12)	3.04 (4)	1.28 (4)
Barium	3.38 (24)	3.16 (32)	2.34 (12)	2.46 (4)
Chromium	22.72 (4)	35.21 (4)	29.56 (8)	22.72 (4)
Copper	7.80 (8)	11.35 (28)	3.95 (8)	2.70 (12)
Iron	ND	ND	ND	1.43 (48)
Lead	4.80 (12)	11.19 (24)	3.62 (8)	1.87 (16)
Cadmium	ND	24.95 (48)[b]	ND	ND
Nickel	24.92 (16)	27.35 (8)	22.53 (8)	ND
Zinc	4.26 (16)	9.18 (24)	4.12 (20)	2.65 (4)

Notes: [a] The MDLs referred to in this column are the precision-based MDLs shown in Table 4-4.

[b] This value may be biased high because the cadmium concentration in the soil samples was near the detection limit.

mg/kg Milligrams per kilogram.

ND No data.

( ) Number of samples, including the four sample preparation steps, each consisting of 10 replicate analyses. Numbers do not always add up to 48 because some samples had analyte concentrations below the analyzer's MDL.
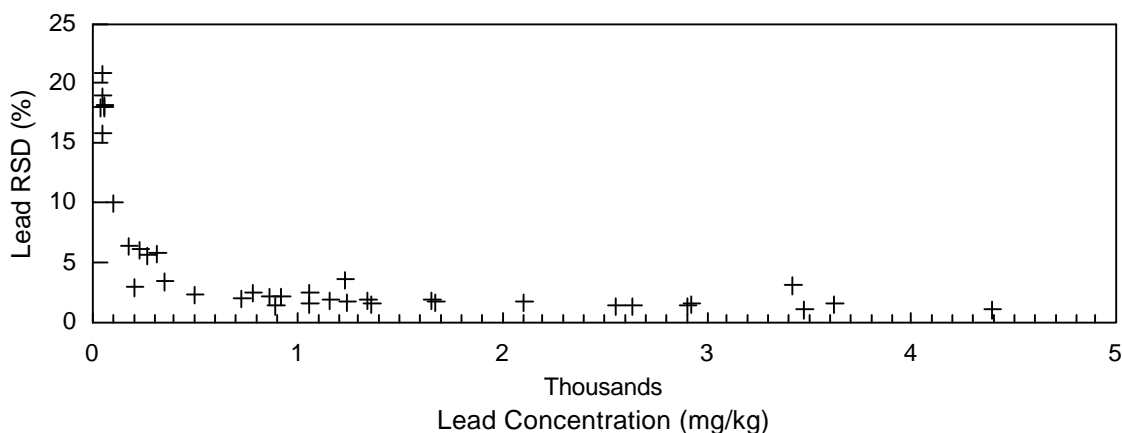
Figure 4-4. Precision vs. Concentration—X-MET 920-P and 940 Analyzers: This graph illustrates the analyzers' precision as a function of analyte concentration.

Accuracy

Accuracy refers to the degree to which a measured value for a sample agrees with a reference or true value for the same sample.

Intramethod accuracy was assessed for the analyzers using site-specific PE samples and SRMs. Accuracy was evaluated through a comparison of percent recoveries for each primary and secondary analyte reported by the analyzers. The analyzers measured 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 six site-specific PE samples consisted of three from each of the two demonstration sites that 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 seven stream or river sediments, two ash, and one 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.

These 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 4-6 summarizes the accuracy data for the target analytes in these samples. Figures 4-5 and 4-6 show the true value, the measured value, and percent recovery for the individual site-specific PEs and SRMs, respectively. Although nickel was reported by the X-MET 920-P, no nickel data is presented or discussed in this accuracy assessment. There was an error in the nickel calibration that caused this data to be unusable. No figure is presented for the cadmium data for the site-specific PEs or for the chromium data for the SRMs because there were two or less detects in both cases.

Based on the 80 - 120 percent recovery acceptance range, the analyzers' accuracy varied from 0 percent for barium and cadmium to 100 percent for arsenic, copper, iron, and lead in the site-specific PEs. Overall, the analyzers produced 28 out of 38 results or 73.7 percent within the 80 - 120 percent recovery acceptance range for all analytes in the six PEs. Nine out of the 10 results falling outside of the acceptance range were below the lower limit of 80 percent recovery. Only a 140 percent recovery for chromium in one PE sample was above the upper limit of 120 percent recovery. Table 4-6 shows that the mean percent recoveries for seven of the eight analytes in the PEs were less than 100 percent. The mean percent recoveries and range of percent recoveries indicate that, in general, the analyzers were producing results that were biased slightly low for iron, lead, and zinc, and substantially low for barium. The analyzers were underestimating barium concentrations by a factor of 2 to 3. This underestimation of barium is somewhat difficult to explain because the site-specific PE samples were used to assist in the calibration of both analyzers and, therefore, the barium results should have been similar to the true value. The cadmium data for the site-specific PEs were inconclusive with only two data points and one of those samples having a true cadmium value of 27.9 mg/kg, which was very close to the precision-based MDL of 25 mg/kg. The one chromium result above the acceptance range was for a sample with a chromium concentration near the precision-based MDL and below the field-based MDL. The one zinc result that fell below the 80 percent acceptance range was only slightly below with a 78 percent recovery.

Table 4-6. Accuracy Summary for Site-Specific PE and SRM Results—X-MET 920-P and 940 Analyzers

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	4	100	95	89 - 108	8.7	92 - 22,444
Barium	6	0	40	21 - 57	13	792 - 7,240
Cadmium	2	0	64	57 - 72	NA	27.9 - 353
Chromium	3	67	117	92 - 140	24	247 - 3,800
Copper	5	100	99	91 - 109	7.9	300 - 7,132
Iron	6	100	86	81 - 91	3.3	27,320 - 70,500
Lead	6	100	85	80 - 92	4.6	292 - 14,663
Zinc	6	83	88	78 - 94	6.9	164 - 4,205
Soil Standard Reference Materials						
Arsenic	3	33	50	0 - 91	58	105 - 626
Barium	5	20	70	47 - 117	29	707 - 2,240
Cadmium	3	100	100	81 - 115	18	21.8 - 71.8
Copper	4	25	162	100 - 210	53	76 - 2,950
Iron	3	100	97	93 - 103	5.8	28,900 - 35,000
Lead	5	80	80	44 - 94	21	101 - 5,532
Zinc	7	86	111	96 - 161	23	81 - 6,952
Sediment Standard Reference Materials						
Arsenic	1	0	71	71	NA	211
Barium	3	67	84	63 - 103	20	335 - 414
Chromium	1	0	22	22	NA	509
Copper	4	25	83	10 - 175	73	99 - 452
Iron	1	100	94	94	NA	41,100
Lead	4	25	72	61 - 81	9.2	161 - 5,200
Zinc	4	75	81	46 - 102	25	264 - 2,200
Ash and Sludge Standard Reference Materials						
Arsenic	2	0	128	126 - 131	NA	136 - 145
Barium	2	0	59	55 - 63	NA	709 - 1,500
Copper	3	67	120	116 - 125	4.2	113 - 696
Iron	2	100	84	84	NA	77,800 - 94,000
Lead	3	33	46	0 - 84	53	68 - 286
Zinc	3	100	106	99 - 112	6.7	210 - 2,122

Notes: n   Number of samples with detectable analyte concentrations.

          SD   Standard deviation.

     mg/kg   Milligrams per kilogram.

          NA   Not applicable, standard deviation not calculated for less than three results.

Figure 4-5. Site-specific PE Sample Results—X-MET 920-P and 940 Analyzers: These graphs illustrate the relationship between the analyzers' data (measured values) and the true values for the site-specific PE samples. The gray bars represent the percent recovery for the analyzers. Each set of three bars (black, white, and gray) represents a single site-specific PE sample.
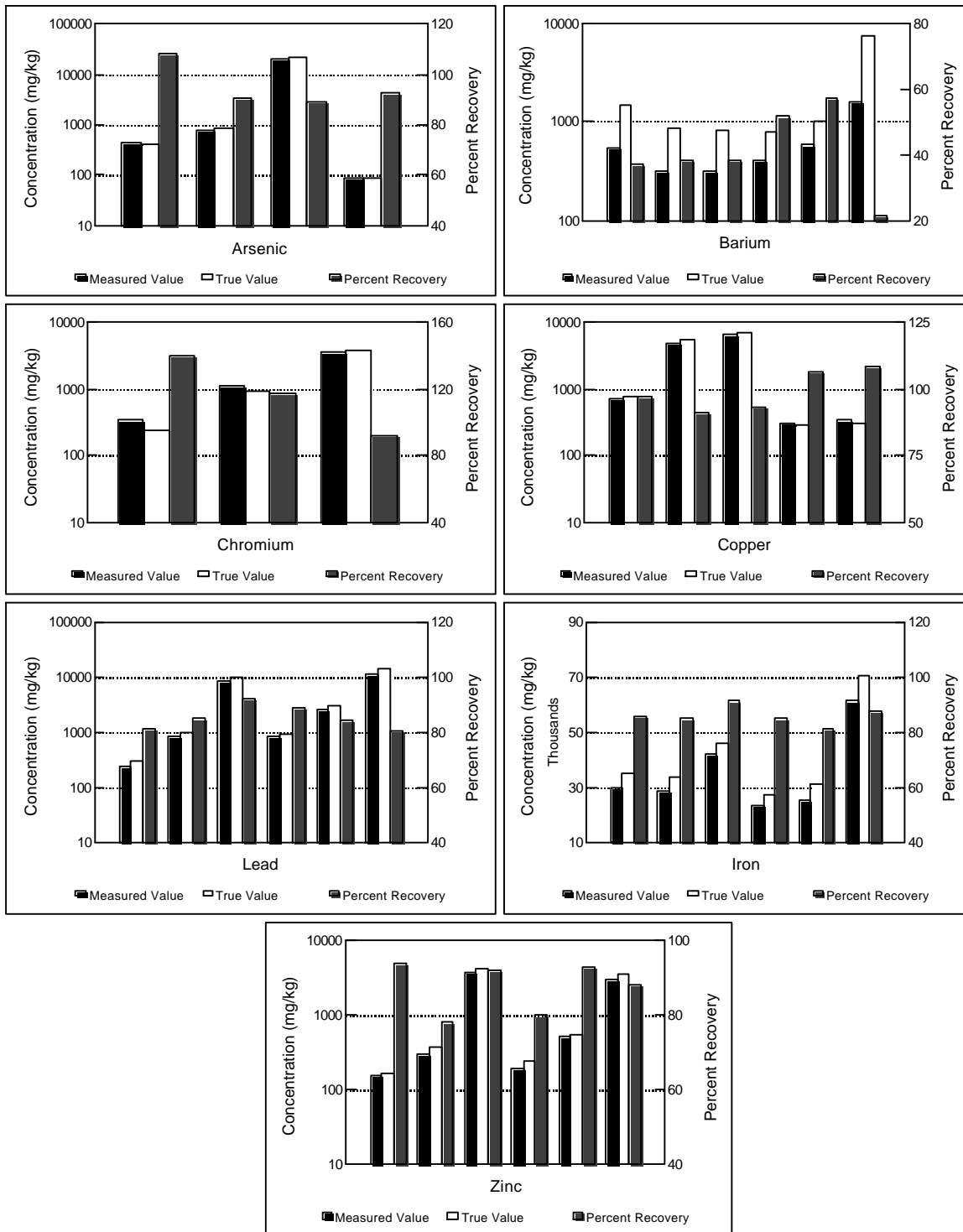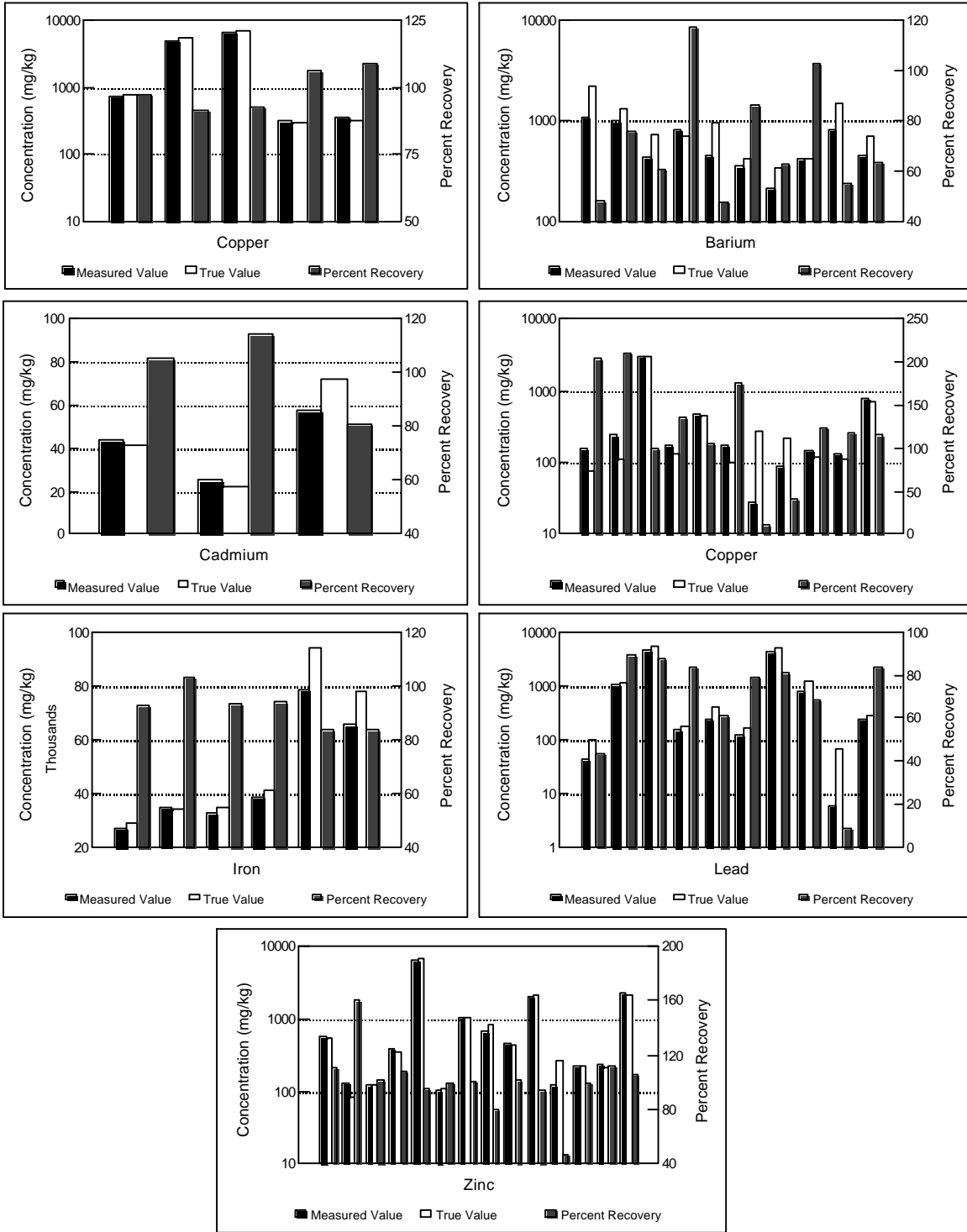
Figure 4-6. SRM Results—X-MET 920-P and 940 Analyzers: These graphs illustrate the relationship between the analyzers' data (measured values) and the true values for the SRMs. The gray bars represent the percent recovery for the analyzers. Each set of three bars (black, white, and gray) represents a single SRM sample.

Table 4-6 provides a summary of the accuracy data for the SRMs. A more detailed analysis of the SRM data is presented in Figure 4-6. The analyzers' accuracy for the SRMs varied from 0 percent for arsenic and chromium in the sediment SRMs and arsenic and barium in the ash and sludge SRMs to 100 percent for iron in all SRMs, cadmium in the soil SRMs, and zinc in the ash and sludge SRMs. The iron concentrations were in the tens of thousands of milligrams per kilogram, which is in a concentration range that the analyzers should perform well. Some analytes such as copper, lead, and zinc had concentrations spanning 1 or more orders of magnitude in all SRMs. Overall, for all analytes in all SRMs, the analyzers produced 35 out of 63 results within the 80 - 120 percent recovery acceptance range for an accuracy of 55.5 percent.

A more detailed analysis of the SRM data showed that there was a matrix effect on the analyzers accuracy. The analyzers produced 19 out of 30 results or 63.3 percent within the acceptance range for all target analytes in the seven soil SRMs; eight out of 18 or 44.4 percent within the acceptance range for all target analytes in the four sediment SRMs; and eight out of 15 results or 53.3 percent within the acceptance range for all target analytes in the ash and sludge SRMs. This demonstrated that the analyzers were more accurate when measuring SRMs of a soil matrix, than sediment, sludge, or ash. This was expected because the soil SRMs more closely matched the matrix of the site-specific soil samples used to calibrate the analyzers.

Although, in general, the analyzers displayed greater accuracy for the soil SRMs, the trend was not consistent among analytes. The accuracy for arsenic and lead was greatest for the soil SRMs. The accuracy for copper and zinc was greatest for the ash and sludge SRMs. Barium accuracy was greatest for the sediment SRMs. Iron accuracy was 100 percent for all SRMs.

The X-MET 920-P and 940 displayed similar or greater accuracy for all analytes except barium and cadmium for the site-specific PEs as compared to the soil SRMs. These results were expected for three reasons. First, site-specific PEs were used to calibrate the X-MET 920-P. Second, 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. Third, the analyte concentrations were often higher in the site-specific PEs versus the soil SRMs. This data supported the developer's contention that a site-specific PE sample is a better assessment of accuracy than an SRM for these analyzers.

Comparability

Intramethod comparability for the analyzers was assessed through the analysis of four ERA PEs and four CRM PEs. This was done to present users 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 Methods 3050A/6010A. Therefore, since these methods do not necessarily determine total metals concentrations in soil, it was expected that the analyzer would overestimate analyte concentrations relative to PALs. The ability of the analyzers to produce results within the PALs and the percent recovery for each of the analytes was used to evaluate intramethod comparability. As with the site-specific PEs and SRMs, the arsenic, cadmium, copper, iron, lead, and zinc data generated from the ASARCO calibration application and the chromium data generated from the RV Hopkins calibration application were used for the comparability assessment. A combination of both applications' data was used in an assessment of the barium data. No nickel data is presented or discussed because of the problem with the nickel calibration.

The analyzers' performance data for all target analytes for the eight CRMs and PEs are summarized in Table 4-7. The measured values, true values, and percent recoveries for all detectable analytes are shown in Figure 4-7. No figure is shown for chromium because there was only one detect. For the ERA PEs, the analyzers produced 15 out of 28 results or 53.6 percent within the acceptance range. For the CRM PEs, the analyzers produced 13 out of 21 results or 61.9 percent within the acceptance range. With the ERA and CRM PEs combined, the analyzers produced 28 out of 47 results or 59.6 percent within the acceptance range. Based on the data presented in Table 4-7, the analyzers' results were slightly more comparable to the CRM PEs than the ERA PEs. The better comparability to the CRM PEs was expected because the ERA PEs had lower analyte concentrations than the CRM PEs. With the exception of iron, the analyte concentrations in the ERA PEs were all less than 350 mg/kg, which was less than 5 times the MDL for most of the analytes. The arsenic and zinc data were more comparable for the ERA than the CRM PEs, while the barium, copper, and iron data were more comparable for the CRM than the ERA PEs. Lead and cadmium comparability were the same for the ERA and CRM PEs. The chromium data were inconclusive.

The analyzers overestimated barium, copper, iron, lead, and zinc concentrations in the ERA PEs by a factor of 1.3 to 2.0, shown by the mean percent recoveries in Table 4-7. The analyzers produced only six out of 28 calculated percent recoveries for the ERA PEs less than 100 percent and four of these were for cadmium. This is consistent with the fact that FPXRF method analysis is a total metal technique whereas EPA SW-846 Methods 3050A/6010A used to certify the results in the ERA PEs are not. The poor comparability for some analytes in the ERA PEs may be an artifact of analyte concentration. All copper concentrations were less than 3 times the precision-based MDL and less than the field-based MDL. All iron concentrations were 3 to 4 times less than the iron concentrations in the site-specific PEs used to calibrate both analyzers. The one lead result outside of the acceptance range was for a sample with 52.4 mg/kg lead, which was slightly above the precision and field-based MDL of 45 mg/kg. The one zinc result outside of the acceptance range was for a sample with 101 mg/kg zinc, which was above the precision-based MDL of 70 mg/kg but less than the field-based MDL of 120 mg/kg.

As with the ERA PEs, the barium results in the CRM PEs were severely overestimated. The comparability of the analyzers' results to the certified values in the CRM PEs was sample dependent. Seven of the eight results outside of the acceptance limits were found for the two soil CRM PEs. This was not expected since the matrix of these CRM PEs more closely matched the matrix of the site-specific PEs than did the sludge or ash CRM PEs. It is possible that interferences were causing the poor comparability in these two CRMs. One soil CRM PE contained nearly 20 percent iron which was greater than the normal 2 - 10 percent iron found in most soil samples. The other soil CRM PE contained nearly 15 percent lead. The recoveries of arsenic, copper, lead, and zinc in these CRM PEs may have been affected by the high concentrations of iron or lead. In both CRM PEs, it is possible that the FPs may not have been able to compensate for the high concentrations of iron or lead.

Intermethod Assessment

The comparison of the X-MET 920-P and 940 results to those of the reference method's was performed using the statistical methods detailed in Section 2. The purpose of this statistical evaluation was to determine the comparability between data produced by the analyzers and 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), 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, they 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 regression analysis of the entire $\log_{10}$ transformed data set for the primary analytes showed that arsenic, copper, and lead had r^2 values at or above 0.93. Barium, chromium, and zinc had r^2 values ranging from 0.86 to 0.43. Based on a comparison of the raw data, the analyzers tended to overestimate concentrations of all the target analytes except lead. Section 5 discusses the average relative bias of the analyzers and how confirmatory analysis and data correction reduces FPXRF bias.

The next step in the data evaluation involved the assessment of the potential impact of the variables: site, soil type, and sample preparation on the regression (Table 4-8). The effect of the site variable was assessed for lead and zinc only. These were the only target analytes exhibiting a wide and similar concentration distribution at each site. The evaluation of the soil variable is presented in Table 4-9.

Table 4-7. PE and CRM Results—X-MET 920-P and 940 Analyzers

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	4	100	102	92 - 114	9.16	65 - 349
Barium	4	25	196	107 - 308	88	111 - 319
Cadmium	4	100	79	64 - 86	10.3	47 - 131
Chromium	0	NA	NA	NA	NA	NA
Copper	4	0	202	178 - 247	32	88 - 196
Iron	4	0	177	153 - 232	37	7,130 - 10,400
Lead	4	75	134	92 - 212	54	52 - 208
Zinc	4	75	138	127 - 160	15.3	76 - 259
Certified Reference Materials						
Arsenic	1	0	19	19	NA	397
Barium	2	50	514	180 - 848	NA	342 - 586
Cadmium	2	100	94	88 - 100	NA	362 - 432
Chromium	1	100	103	103	NA	161,500
Copper	4	50	80	20 - 129	49	279 - 4,792
Iron	3	100	81	52 - 100	25	6,481 - 191,650
Lead	4	75	100	23 - 185	67	120 - 144,740
Zinc	4	25	96	6.4 - 181	76	546 - 22,217

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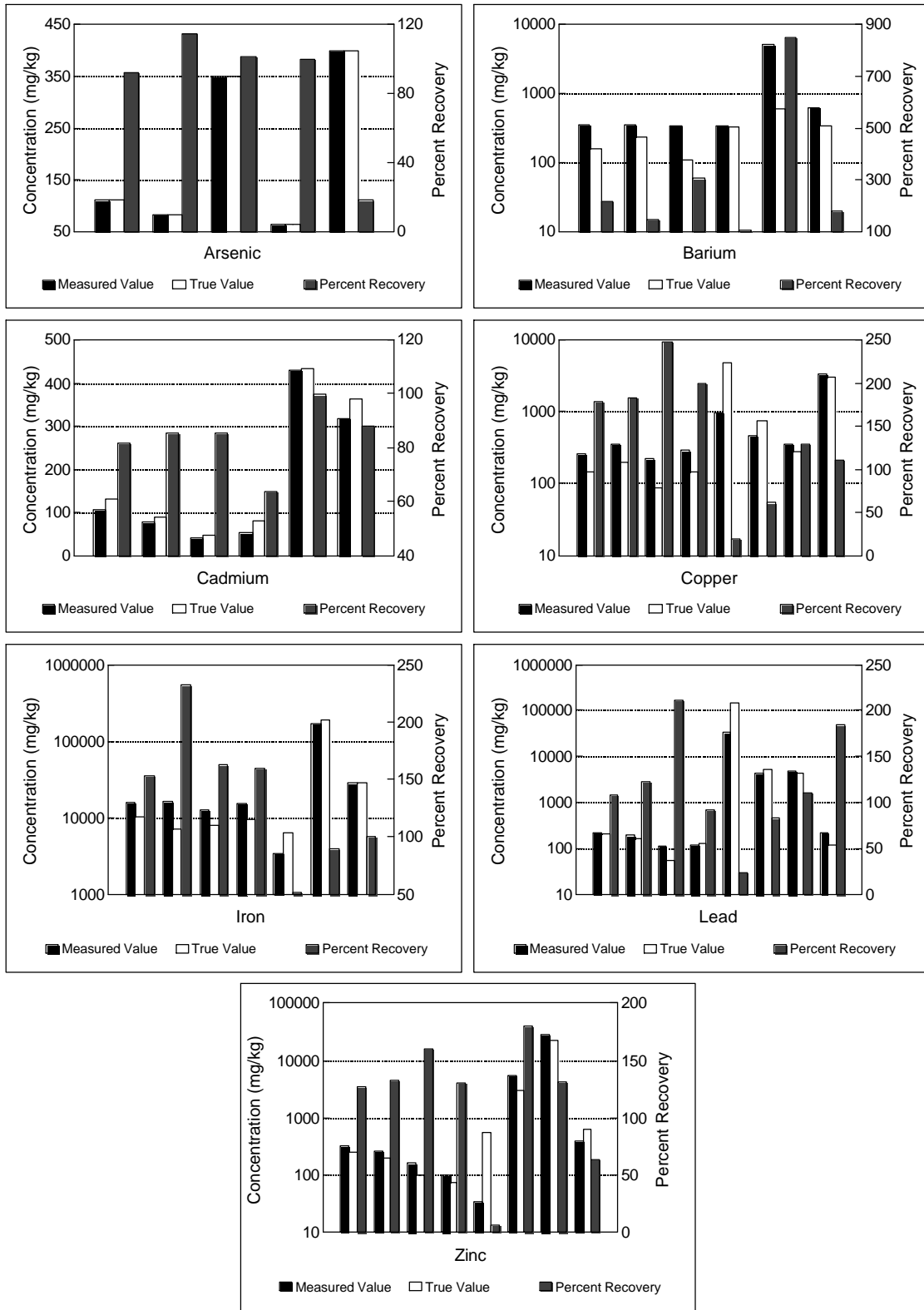   NA   Not applicable, analyte not present above LRL.

Figure 4-7. PE and CRM Results—X-MET 920-P and 940 Analyzers: These graphs illustrate the relationship between the analyzers' data (measured values) and the true values for the PE and CRM samples. The gray bars represent the percent recovery for the analyzers. Each set of three bars (black, white, and gray) represents a single PE or CRM sample.

64

Table 4-8. Regression Parameters[a] by Primary Variable–X-MET 920-P and 940 Analyzers

		Arsenic			Variable			Barium		
n	r²	Std. Err.	Y-Int.	Slope		n	r²	Std. Err.	Y-Int.	Slope
993	0.941	0.24	0.11	0.96	All Data	891	0.43	0.23	1.48	0.52
773	0.964	0.13	0.27	0.91	ASARCO Site	746	0.04	0.22	2.07	0.18
248	0.066	0.54	0.06	0.85	RV Hopkins Site	172	0.77	0.10	2.15	0.34
384	0.964	0.14	0.31	0.89	Sand Soil	368	0.00	0.14	2.38	0.03
391	0.961	0.12	0.21	0.94	Loam Soil	380	0.14	0.28	1.55	0.41
248	0.066	0.54	0.06	0.85	Clay Soil	398	0.66	0.15	2.04	0.39
215	0.913	0.33	0.04	0.95	In Situ-Unprepared	228	0.73	0.17	1.49	0.56
266	0.957	0.21	-0.11	1.03	In Situ-Prepared	253	0.35	0.34	1.31	0.58
254	0.964	0.17	0.16	0.94	Intrusive-Unprepared	309	0.61	0.21	1.36	0.61
259	0.954	0.20	0.27	0.95	Intrusive-Prepared	206	0.62	0.19	1.52	0.56

		Chromium			Variable			Copper		
n	r²	Std. Err.	Y-Int.	Slope		n	r²	Std. Err.	Y-Int.	Slope
1154	0.673	0.18	1.72	0.49	All Data	1147	0.93	0.19	0.92	0.75
771	0.000	0.13	2.42	0.02	ASARCO Site	778	0.97	0.12	0.44	0.90
387	0.665	0.24	1.62	0.53	RV Hopkins Site	371	0.07	0.21	1.87	0.17
378	0.030	0.12	2.76	-0.19	Sand Soil	381	0.95	0.12	0.53	0.86
392	0.053	0.13	2.10	0.20	Loam Soil	395	0.97	0.11	0.37	0.92
387	0.665	0.24	1.62	0.53	Clay Soil	371	0.07	0.21	1.87	0.17
237	0.555	0.21	1.83	0.42	In Situ-Unprepared	234	0.86	0.21	1.18	0.63
305	0.594	0.19	1.77	0.44	In Situ-Prepared	298	0.95	0.17	0.85	0.77
308	0.753	0.15	1.69	0.52	Intrusive-Unprepared	303	0.95	0.16	0.86	0.76
304	0.774	0.15	1.66	0.55	Intrusive-Prepared	307	0.94	0.18	0.88	0.79

		Lead			Variable			Zinc[b]		
n	r²	Std. Err.	Y-Int.	Slope		n	r²	Std. Err.	Y-Int.	Slope
1147	0.945	0.16	0.06	0.96	All Data	860	0.860	0.18	0.45	0.86
749	0.959	0.14	-0.06	1.01	ASARCO Site	747	0.905	0.14	0.38	0.87
396	0.906	0.18	0.39	0.85	RV Hopkins Site	141	0.710	0.23	1.18	0.67
373	0.956	0.14	0.06	0.95	Sand Soil	390	0.941	0.12	0.41	0.87
378	0.959	0.14	-0.19	1.06	Loam Soil	358	0.844	0.16	0.29	0.89
396	0.906	0.18	0.39	0.85	Clay Soil	141	0.710	0.23	1.18	0.67
233	0.866	0.24	0.35	0.85	In Situ-Unprepared	214	0.770	0.23	0.63	0.79
306	0.965	0.12	0.08	0.95	In Situ-Prepared	253	0.810	0.20	0.51	0.83
303	0.971	0.12	-0.10	1.00	Intrusive-Unprepared	186	0.960	0.10	0.23	0.93
307	0.968	0.13	0.02	1.01	Intrusive-Prepared	195	0.950	0.11	0.35	0.91

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]　Zinc results for intrusive-unprepared and intrusive-prepared do not include samples from the RV Hopkins site.

n　Number of data points.

Y-Int.　Y-Intercept.

Std. Err.　Standard error.

65

Table 4-9. Regression Parameters[a] for the Sample Preparation Variable and Soil Texture— X-MET 920-P and 940 Analyzers

Arsenic						Barium				
n	r²	Std. Err.	Y-Int.	Slope		n	r²	Std. Err.	Y-Int.	Slope
In Situ-Unprepared					Soil Texture	**In Situ-Unprepared**				
95	0.933	0.18	0.43	0.81	Sand Soil	85	0.037	0.08	2.28	0.10
50	0.877	0.21	0.38	0.85	Loam Soil	48	0.678	0.12	1.36	0.53
71	0.074	0.53	0.05	0.88	Clay Soil	97	0.774	0.11	2.08	0.37
In Situ-Prepared					Soil Texture	**In Situ-Prepared**				
95	0.989	0.08	0.16	0.95	Sand Soil	97	0.004	0.33	1.99	0.15
114	0.989	0.06	0.15	0.95	Loam Soil	112	0.053	0.36	1.51	0.30
66	0.089	0.53	-0.17	0.95	Clay Soil	75	0.798	0.08	2.25	0.31
Intrusive-Unprepared					Soil Texture	**Intrusive-Unprepared**				
96	0.984	0.10	0.16	0.95	Sand Soil	91	0.031	0.06	2.34	0.07
113	0.986	0.07	0.11	0.96	Loam Soil	111	0.568	0.14	1.30	0.54
53	0.033	0.56	0.39	0.58	Clay Soil	100	0.636	0.19	1.82	0.47
Intrusive-Prepared					Soil Texture	**Intrusive-Prepared**				
96	0.980	0.11	0.39	0.90	Sand Soil	98	0.104	0.06	2.23	0.14
114	0.985	0.08	0.29	0.95	Loam Soil	110	0.423	0.11	1.88	0.33
59	0.034	0.59	0.33	0.64	Clay Soil	97	0.784	0.09	2.24	0.35

Chromium						Copper				
n	r²	Std. Err.	Y-Int.	Slope		n	r²	Std. Err.	Y-Int.	Slope
In Situ-Unprepared					Soil Texture	**In Situ-Unprepared**				
94	0.049	0.08	2.77	-0.15	Sand Soil	94	0.909	0.14	0.78	0.74
49	0.005	0.14	2.28	0.11	Loam Soil	50	0.840	0.18	0.49	0.87
92	0.715	0.25	1.27	0.63	Clay Soil	87	0.215	0.19	1.72	0.32
In Situ-Prepared					Soil Texture	**In Situ-Prepared**				
95	0.035	0.14	2.82	-0.25	Sand Soil	96	0.963	0.10	0.50	0.87
111	0.028	0.10	2.23	0.11	Loam Soil	114	0.985	0.07	0.40	0.90
97	0.618	0.26	1.57	0.52	Clay Soil	87	0.053	0.22	1.82	0.16
Intrusive-Unprepared					Soil Texture	**Intrusive-Unprepared**				
94	0.046	0.13	2.83	-0.24	Sand Soil	96	0.972	0.09	0.34	0.93
14	0.082	0.10	2.11	0.19	Loam Soil	114	0.990	0.06	0.32	0.92
100	0.752	0.17	1.81	0.48	Clay Soil	95	0.034	0.17	1.96	0.09
Intrusive-Prepared					Soil Texture	**Intrusive-Prepared**				
93	0.016	0.10	2.65	-0.11	Sand Soil	97	0.961	0.10	0.48	0.91
111	0.074	0.12	2.09	0.21	Loam Soil	14	0.990	0.06	0.43	0.93
99	0.790	0.17	1.74	0.53	Clay Soil	98	0.030	0.22	2.00	0.11

Table 4-9. *Continued*

Lead						Zinc[b]				
n	r^2	Std. Err.	Y-Int.	Slope		n	r^2	Std. Err.	Y-Int.	Slope
In Situ-Unprepared					Soil Texture	**In Situ-Unprepared**				
86	0.833	0.20	0.44	0.75	Sand Soil	94	0.937	0.11	0.56	0.78
49	0.849	0.23	0.38	0.84	Loam Soil	48	0.762	0.19	0.48	0.81
99	0.833	0.22	0.73	0.75	Clay Soil	83	0.717	0.23	1.18	0.69
In Situ-Prepared					Soil Texture	**In Situ-Prepared**				
95	0.985	0.08	0.02	0.97	Sand Soil	95	0.965	0.10	0.35	0.89
111	0.986	0.08	-0.19	1.05	Loam Soil	103	0.850	0.15	0.24	0.90
100	0.931	0.15	0.50	0.81	Clay Soil	58	0.728	0.22	1.91	0.63
Intrusive-Unprepared					Soil Texture	**Intrusive-Unprepared**				
97	0.980	0.10	-0.03	0.98	Sand Soil	95	0.977	0.08	0.22	0.94
113	0.973	0.12	-0.40	1.11	Loam Soil	103	0.913	0.11	0.28	0.90
99	0.951	0.15	-0.03	0.96	Clay Soil	ND	ND	ND	ND	ND
Intrusive-Prepared					Soil Texture	**Intrusive-Prepared**				
96	0.970	0.12	0.01	1.00	Sand Soil	97	0.973	0.09	0.39	0.91
111	0.975	0.11	-0.19	1.10	Loam Soil	101	0.862	0.15	0.30	0.92
98	0.968	0.11	0.28	0.92	Clay Soil	ND	ND	ND	ND	ND

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 the FPXRF data. See Section 5.

    [b] Zinc results for intrusive-unprepared and intrusive-prepared do not include samples from the RV Hopkins site.

    Y-Int. Y-Intercept.

    Std. Err. Standard Error.

    n Number of data points.

    ND Analytes not present in significant quantities to provide meaningful regression.

Copper, arsenic, and cadmium did not exhibit a wide concentration distribution in the clay soil. Therefore, the effect of soil type on the quantitation of these analytes was restricted to sand and loam soil. Barium, chromium, and nickel did not exhibit a wide concentration distribution in either the sand or loam soil. Therefore, the effect of soil type on the quantitation of these analytes was restricted to loam soil only. This single soil type evaluation could not be used to assess the potential effect of soil type on analyzer performance. Based on this evaluation, there was no apparent impact of either the site or soil variables on the regression. The sample preparation variable exhibited the greatest impact on the regression (Table 4-10). Generally, the largest shift in the r^2 was exhibited between the *in situ*-unprepared and *in situ*-prepared analyses (Figure 4-8); this step represents sample homogenization. Generally, r^2 values increased with increasing sample preparation. Except for zinc, sample homogenization accounted for between 80 and 100 percent of the total increase in the r^2 experienced across all sample preparation steps. Zinc exhibited its greatest increase in comparability after the intrusive-unprepared analysis. This sample preparation effect makes sense due to the fact that the homogenization step assured that the analyzers and the reference methods were analyzing essentially the same sample. The initial sample homogenization (*in situ*-prepared)

improved the comparability for arsenic and zinc between the two $\log_{10}$ transformed data sets to the point that the analyzers met the definitive level criteria. Lead data produced by the analyzers met definitive level criteria after the final sample preparation step. The analyzer data for copper met the definitive level quality criteria at the initial sample preparation step. Increasing sample preparation increased comparability for barium; however, the data for this analyte never met a higher level criteria. Chromium failed to meet quantitative screening level criteria for the regression analysis and its precision data, slightly greater than 20 percent, placed it in the qualitative screening level category.

Table 4-10. Regression Parameters[a] for the Sample Preparation Variable and Site Name– X-MET 920-P and 940 Analyzers

		Arsenic						Barium		
n	r^2	Std. Err.	Y-Int.	Slope		n	r^2	Std. Err.	Y-Int.	Slope
In Situ-Unprepared					**Site Name**	**In Situ-Unprepared**				
145	0.925	0.19	0.39	0.83	ASARCO	132	0.151	0.10	2.09	0.19
71	0.074	0.53	0.05	0.88	RV Hopkins	97	0.774	0.11	2.08	0.37
In Situ-Prepared					**Site Name**	**In Situ-Prepared**				
208	0.990	0.07	0.14	0.95	ASARCO	213	0.005	0.38	2.00	0.10
66	0.089	0.53	-0.17	0.95	RV Hopkins	75	0.798	0.08	2.25	0.31
Intrusive-Unprepared					**Site Name**	**Intrusive-Unprepared**				
209	0.985	0.08	0.13	0.95	ASARCO	206	0.219	0.14	1.85	0.30
53	0.033	0.56	0.39	0.58	RV Hopkins	100	0.686	0.19	1.82	0.47
Intrusive-Prepared					**Site Name**	**Intrusive-Prepared**				
210	0.981	0.09	0.35	0.92	ASARCO	207	0.306	0.08	2.09	0.22
59	0.034	0.59	0.33	0.64	RV Hopkins	97	0.784	0.09	2.24	0.35

		Chromium						Copper		
n	r^2	Std. Err.	Y-Int.	Slope		n	r^2	Std. Err.	Y-Int.	Slope
In Situ-Unprepared					**Site Name**	**In Situ-Unprepared**				
146	0.038	0.12	2.84	-0.23	ASARCO	147	0.888	0.17	0.70	0.79
92	0.715	0.25	1.27	0.63	RV Hopkins	87	0.215	0.19	1.72	0.32
In Situ-Prepared					**Site Name**	**In Situ-Prepared**				
207	0.001	0.13	2.47	-0.03	ASARCO	209	0.984	0.08	0.44	0.89
97	0.618	0.26	1.57	0.52	RV Hopkins	87	0.053	0.22	1.82	0.16
Intrusive-Unprepared					**Site Name**	**Intrusive-Unprepared**				
207	0.003	0.12	2.37	0.05	ASARCO	210	0.986	0.08	0.35	0.92
100	0.752	0.17	1.81	0.48	RV Hopkins	95	0.034	0.17	1.96	0.09
Intrusive-Prepared					**Site Name**	**Intrusive-Prepared**				
204	0.011	0.12	2.31	0.09	ASARCO	211	0.984	0.08	0.44	0.92
99	0.790	0.17	1.74	0.53	RV Hopkins	98	0.030	0.22	2.00	0.11

Table 4-10. *Continued*

	Lead					Zinc[b]				
n	**r^2**	**Std. Err.**	**Y-Int.**	**Slope**	**Site Name**	**n**	**r^2**	**Std. Err.**	**Y-Int.**	**Slope**
		In Situ-Unprepared			**Site Name**			**In Situ-Unprepared**		
132	0.892	0.20	0.30	0.84	ASARCO	144	0.890	0.15	0.55	0.78
99	0.833	0.22	0.73	0.75	RV Hopkins	85	0.686	0.25	1.12	0.71
		In Situ-Prepared			**Site Name**			**In Situ-Prepared**		
207	0.983	0.08	-0.07	1.01	ASARCO	200	0.864	0.17	0.29	0.89
100	0.931	0.15	0.50	0.81	RV Hopkins	83	0.717	0.23	1.18	0.69
		Intrusive-Unprepared			**Site Name**			**Intrusive-Unprepared**		
206	0.977	0.10	-0.17	1.03	ASARCO	196	0.957	0.10	0.23	0.93
99	0.951	0.15	-0.03	0.96	RV Hopkins	ND	ND	ND	ND	ND
		Intrusive-Prepared			**Site Name**			**Intrusive-Prepared**		
208	0.969	0.13	-0.11	1.05	ASARCO	199	0.927	0.13	0.38	0.90
98	0.968	0.11	0.28	0.92	RV Hopkins	ND	ND	ND	ND	ND

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 the FPXRF data. See Section 5.

        [b] Zinc results for intrusive-unprepared and intrusive-prepared do not include samples from the RV Hopkins site.

Y-Int. Y-Intercept.

Std. Err. Standard Error.

n Number of data points.

ND Analytes not present in significant quantities to provide meaningful regression.

Within the sample preparation steps, the effect of contaminant concentration was also examined. The $\log_{10}$ transformed data sets for the primary analytes were sorted into the following concentrations ranges: 0 - 100 mg/kg, 100 - 1,000 mg/kg, and greater than 1,000 mg/kg. The regression analysis for each target analyte and for each sample preparation step was rerun on these concentration-sorted data sets. A review of these results showed general improvement in the r^2 and standard error with increasing concentration. The 0 - 100 mg/kg concentration range showed the poorest comparability. This is most likely due to this range generally occurring just above or below the analyzers field-based MDLs. The analyzers' precision and accuracy are lowest in this concentration range. This proximity to MDLs was also exhibited for most of the target analytes in the 100 - 1,000 mg/kg range. Generally, the r^2 values improved between 5 and 30 percent between 100 and 1,000 mg/kg and greater than 1,000 mg/kg ranges. This effect was minimized by examining the $\log_{10}$ transformed data and the slight effect appears to be linked to the general proximity of a measurement to its associated MDL. The further away from the MDL, the less effect concentration will have on quantitation and comparability.

A final decision regarding the assignment of data quality levels derived from this demonstration involves an assessment of both r^2 and the precision RSD. Using the criteria presented in Table 2-2, a summary of the Metorex X-MET 920-P and 940 data quality performance in this demonstration is provided in Table 4-11.
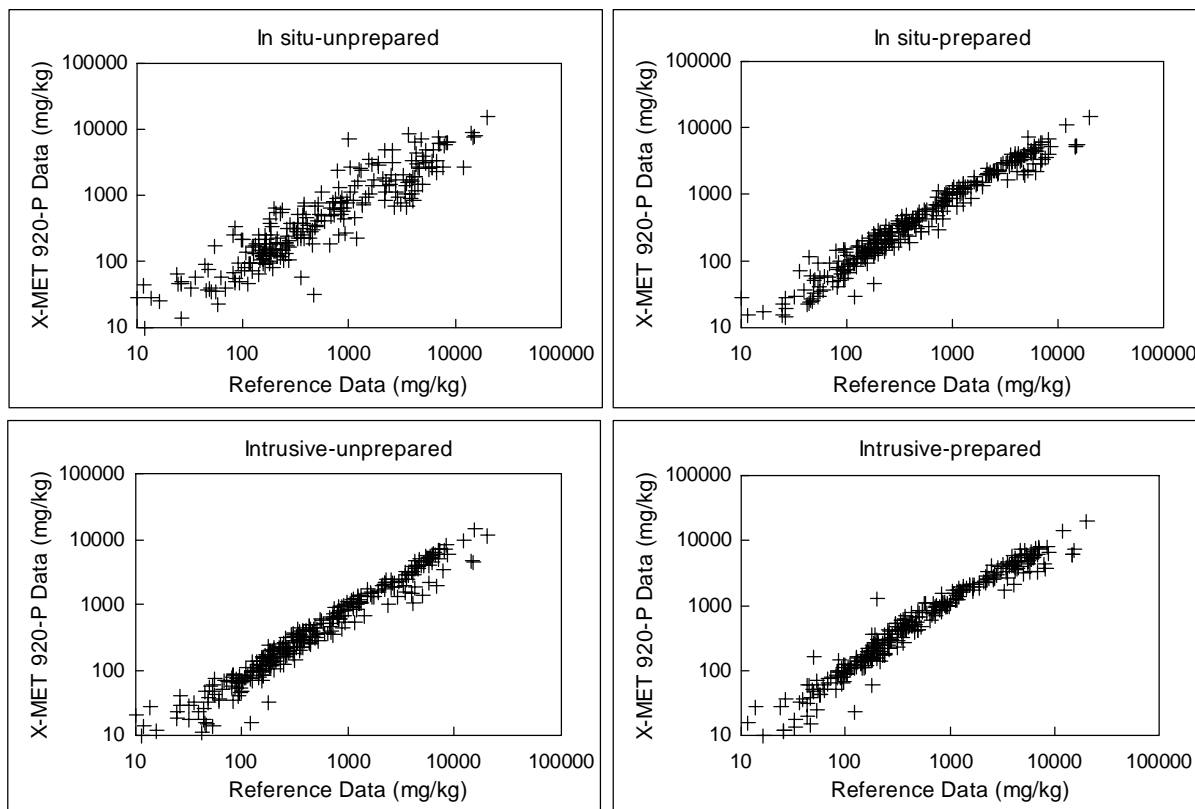
Figure 4-8. Sample Preparation Effect on Lead Results: These log-log plots illustrate the change in comparability with changes in sample preparation step.

Table 4-11. Summary of Data Quality Level Parameters

Target Analytes	920-P/940 Analytes	Precision (mg/kg) 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	3.36	55	0.94	Definitive
Barium	Barium	3.38	30	0.43	Qualitative
Chromium	Chromium	22.72	210	0.67	Qualitative
Copper	Copper	7.80	75	0.93	Definitive
Lead	Lead	4.80	45	0.94	Definitive
Zinc	Zinc	4.26	70	0.86	Definitive
Nickel	Nickel	24.92	120	Not Determined	Insufficient Data
Iron	Iron	Not Determined	Not Determined	Not Determined	Insufficient Data
Cadmium	Cadmium	Not Determined	25	Not Determined	Insufficient Data
Antimony	Not Reported	—	—	—	—

Section 5
Applications Assessment and Considerations

The Metorex X-MET 920-P and 940 analyzers are designed to analyze for metals in soils. Developer provided ACES software was used for calibration and quantitation to maximize instrument performance and account for common soil-related matrix interferences. In addition, the calibrations can be fine tuned using site-specific samples to further improve data comparability. These analyzers were designed for field use and exhibited ruggedness through a variety of environmental operating conditions. The analyzers experienced two failures resulting in downtime and lost data during the field demonstration. During the demonstration, more than 1,260 samples were analyzed by these instruments. The training provided by the developer was sufficient to allow basic field operation of the analyzers; however, developer assistance was required to address software and equipment problems. The developer provided highly accessible and timely field support.

The X-MET 940 was lighter and more compact than the X-MET 920-P. The data storage problems associated with the X-MET 940's hard drive need to be resolved, allowing a greater number of samples to be analyzed and reducing the potential for data loss or equipment downtime. A summary of the operational features of both instruments is shown in Table 5-1.

Comparison of the $\log_{10}$ transformed FPXRF data to the $\log_{10}$ transformed reference data indicated that the FPXRF and reference data are $\log_{10}$-$\log_{10}$ linearly related. Based on this relationship, both analyzers can produce definitive level data for lead, arsenic, copper, and zinc. This indicated that the $\log_{10}$ transformed FPXRF data were statistically equivalent to the $\log_{10}$ transformed reference data for these analytes. For the target analytes barium and chromium, the analyzers produced qualitative screening level data. These analyzers, for the above elements, exhibited instrument precision similar to the reference methods, indicating high measurement reproducibility.

Both analyzers can use up to two radioactive sources allowing analysis of a large number of metals in soils. The analyzers generally use count times 100 - 240 live-seconds. Longer count times and multiple sources generally increase accuracy, the number of analytes and lower the detection limits but decrease sample throughput.

There were no apparent effects of site or soil type on instrument performance. Both analyzers can be applied in an *in situ* or intrusive mode. This demonstration identified sample preparation as the most important variable with regard to comparability of the data from analyzers to the reference methods. The data from this demonstration indicated that when operated in the *in situ* mode, the user most probably would not be able to show a strong correlation between FPXRF and reference data. This may not be due to instrument error, but rather to inherent spacial 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 and reference data for the analyzers was achieved after the initial sample preparation step (sample homogenization). Further sample preparation, such as sieving or drying and sieving, in most cases did improve the comparability; however, relative to the demonstration-defined data quality criteria, no increases in data quality level occurred. Based on the comparability study, the analyzers produced field-based MDLs 2 to 3 times greater than the developer-supplied or precision-based MDLs. This difference may have been due to differences in the definition of MDLs used by the developer and the demonstration.

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

• 920-P: Computer weighs 25 pounds, probe weighs 9 pounds, battery life of 5 to 7 hours
• 940: Computer weighs 14 pounds, probe weighs 9 pounds, battery life of 2 to 3 hours
Sample throughput of 10 to 12 samples per hour
Three radioactive sources available
Up to two sources can be fitted on an analyzer's probe increasing the number of potential analytes
Si(Li) or gas-filled proportional detectors available
Arsenic, lead, copper, and zinc–definitive level data Barium, chromium–qualitative screening level data
Fundamental parameters calibration can be fine tuned with site-specific samples (optional empirical calibration available)
Precision—percent RSD values less than 10 percent at 5 to 10 times the MDL for all analytes except nickel and Cr for those reported (Cd and Fe NR)
Can be used on soils exhibiting up to 30 percent water saturation by weight
Can conduct in situ or intrusive measurements

Based on this demonstration, both analyzers are well suited for the rapid real-time assessment of metals contamination in soil samples. This demonstration indicated that the use of these analyzers in a rainy environment requires special handling to keep the probe dry. In addition, more extensive training or operator experience would be needed to reduce the potential for data loss and downtime associated with software problems. The addition of calibration check samples would also reduce the potential for data loss. Although in several cases the analyzers produced data statistically equivalent to the reference data, generally confirmatory analysis would be required or requested for FPXRF analysis. Ten to 20 percent of the samples measured by the analyzers should be submitted for reference method analysis; then instrument bias relative to standard methods such as 3050A/6010A could be determined. This would only hold true if the analyzers and the reference laboratory measure similar samples. This was accomplished in this demonstration by thorough sample homogenization. Bias correction allows FPXRF data to be corrected so that it approximates the reference data. The demonstration showed that both analyzers exhibit strong $\log_{10}$-$\log_{10}$ linear relationships with the reference data over a concentration range of 5 orders of magnitude. For optimum correlation, samples in the high, medium, and low concentration ranges should be submitted for reference method analysis.

The steps to correct the FPXRF measurements to more closely match 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 x-axis column (dependent variable) and the FPXRF data in the y-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:

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

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 the analyzers' 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 analyzers' 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 FPXRF, before and after data correction using the eight-step approach previously discussed.

The average relative bias and accuracy for the analytes which fell into the definitive level data quality category were generally small. The analytes falling into the quantitative and qualitative screening level data quality categories had generally larger average relative bias and accuracy.

Once the FPXRF data is corrected using the regression approach presented earlier, both the average relative bias and accuracy were reduced. These numbers are no longer strongly influenced by 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 th2e last column in Table 5-2.

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 Relative Accuracy Based on PE Samples[e]
Antimony	4.60	1.19	10.40	2.65	2.94 ± 0.56
Arsenic	1.07	1.05	2.47	2.42	1.76 ± 0.28
Barium	3.21	1.31	9.21	2.76	1.36 ± 0.06
Chromium	8.15	1.27	18.19	3.04	1.55 ± 0.15
Copper	2.18	1.13	6.33	2.66	1.18 ± 0.47
Iron	1.28	1.02	1.69	1.40	1.54 ± 0.14
Lead	1.92	1.03	1.54	1.62	1.63 ± 0.23
Nickel	3.91	1.30	14.18	3.48	1.56 ± 0.14
Zinc	1.67	1.20	5.05	2.41	1.64 ± 0.12

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 = $((\sum_i[\text{FPXRF}_i/\text{Reference}_i])\diagup$ 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 $(\sum_i([\text{FPXRF}_i/\text{Reference}_i]-1)^2\diagup$ 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.

Both analyzers can provide rapid assessment regarding the distribution of metals contamination in soil 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 in some applications and for some analytes, the FPXRF data may be statistically similar to the reference data. The approval of Method 6200 will help in the acceptance of FPXRF data. 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. Both analyzers are powerful tools for site characterization and remediation. They provide a faster and less expensive means of analyzing metals contamination in soil, relative to conventional approaches.

General Operational Guidance

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

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 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. Developers state that their FPXRF instruments will perform adequately for soils with moisture contents of 5 - 20 percent, but 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 indicated 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 in the sample, 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 the method blanks are below the MDLs of the procedure.

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.