United States
Environmental Protection
Agency

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

EPA/600/R-97/151
March 1998

Environmental Technology Verification Report

Field Portable X-ray Fluorescence Analyzer

Metorex X-MET 920-MP

SITE
SUPERFUND INNOVATIVE
TECHNOLOGY EVALUATION

ETV
Environmental Technology
Verification Program

070CMB98

EPA/600/R-97/151
March 1998

Environmental Technology Verification Report

Field Portable X-ray Fluorescence Analyzer

Metorex X-MET 920-MP

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-MP**
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-MP.

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-MP." The EPA document number for this report is EPA/600/R-97/151.

The EPA SW-846 Method 6200 was tested and validated using the data derived from this demonstration. This method may be used to support the general application of FPXRF for environmental analysis.

TECHNOLOGY DESCRIPTION

This analyzer operates on the principle of energy dispersive X-ray fluorescence spectroscopy where the characteristic 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 field detection limits may be 5 to 10 times greater than conventional laboratory methods. As with all field collection programs, a portion of the samples should be sent to a laboratory for confirmatory analyses.

The X-MET 920-MP is designed to produce quantitative data on metals in soils, sludges, and other solids. The X-MET 920-MP consists of a laptop computer, an electronics unit, and a surface analysis probe system (SAPS). The electronics is housed in a rugged, weatherproof, self-contained case, weighing about 5 pounds that can be operated from battery power up to 8 hours. The SAPS is designed to house one excitation source (cadmium-109 for this demonstration) and a gas-filled proportional counter detector. The SAPS weighs about 3 pounds and is specifically designed for *in situ* analysis, but can be adapted for measurement of samples in cups. The single excitation source limits the number of metals that can be quantified. The X-MET 920-MP is operated and calibrated using the "X-MET" software to analyze samples with an empirical calibration. Training and field experience is necessary to successfully derive empirical calibration curves and to operate the "X-MET" software. The X-MET 920-MP reported the analytes arsenic, barium, copper, chromium, lead, nickel, and zinc for this demonstration using source count times between 30 and 180 seconds. At the time of the demonstration, the cost of the X-MET 920-MP with the SAPS probe and cadmium-109 source (including the laptop computer) was $36,325, or it could be leased for $3,633 per month.

VERIFICATION OF PERFORMANCE

The performance characteristics of the X-MET 920-MP 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 less than or equal to 50 milligrams per kilogram (mg/kg) for all analytes except barium (330 mg/kg) and chromium (115 mg/kg). Barium is normally analyzed using an americium-241 source; therefore, its detection limit was expected to be high. A value for nickel could not be determined because the soil concentration of this analyte was too low.
- **Throughput:** Average throughput was found to be between 8 and 14 analyses per hour, depending on count times. This rate only represents the analysis time since different personnel were used to prepare the samples.
- **Drift:** Based on an evaluation of results from periodic analysis of a site-specific control sample, with a few exceptions, drift was -15 to +15 percent. Lead and arsenic displayed the least drift at both sites.
- **Completeness:** The X-MET 920-MP produced results for 1,168 of the 1,260 samples for a completeness of 92.7 percent. This was less than the demonstration objective of 95 percent. Operator error and computer software and hardware problems reduced completeness. None of the data loss was caused by mechanical or electronic malfunctions of the analyzer.

- **Blank results:** The X-MET 920-MP reported values for arsenic and copper above the precision-based method detection limits at the ASARCO site and values for chromium, lead, and zinc above the MDL at the RV Hopkins site. Analyzer blanks were composed of a pure lithium carbonate that was processed using the sample preparation steps.
- **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 limit. The RSD values for the reported analytes were less than 8 percent. Chromium and nickel were not determined due to a lack of sufficient data in the specified concentration range.
- **Accuracy:** Intramethod accuracy was assessed using site-specific soil PE samples. The results showed that 7 of 32 (21.9 percent) of the PE sample analytes had recoveries within a quantitative acceptance range of the 80 - 120 percent.
- **Comparability:** This demonstration showed that the X-MET 920-MP 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 0.95 for arsenic, 0.88 for lead, 0.69 for copper, 0.68 for chromium, and 0.55 for zinc. Using data from the RV Hopkins clay soil produced values of 0.62 for barium and 0.32 for nickel.
- **Data quality levels:** Using the demonstration derived precision RSD results and the coefficient of determination as the primary qualifiers, the X-MET 920-MP produced definitive level data for arsenic and lead and data of qualitative screening level for copper, barium, and zinc. No recommendation regarding data quality for chromium or nickel could be made due to a lack of precision or comparability data.

The results of the demonstration show that the Metorex X-MET 920-MP can provide useful, cost-effective data for environmental problem-solving and decision-making. Undoubtedly, it will be employed in a variety of applications, ranging from serving as a complement to data generated in a fixed analytical laboratory to generating data that will stand alone in the decision-making process. As with any technology selection, the user must determine what is appropriate for the application and the project data quality objectives.

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

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

Foreword

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

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

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

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

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

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

Abstract

In April 1995, the U.S. Environmental Protection Agency (EPA) conducted a demonstration of field portable X-ray fluorescence (FPXRF) analyzers. The primary objectives of this demonstration were (1) to determine how well FPXRF analyzers perform in comparison to 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: the RV Hopkins site and the ASARCO Tacoma Smelter site (ASARCO). 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 test samples analyzed during this demonstration were evenly distributed between 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 tested in this demonstration were designed to provide rapid, real-time analysis of metals concentrations in soil samples. This information will allow investigation and remediation decisions to be made on-site more efficiently and can reduce the number of samples that need to be submitted for confirmatory analysis. Of the seven commercially available analyzers tested, one is manufactured by Niton Corporation (the XL Spectrum Analyzer); two are manufactured by TN Spectrace (the TN 9000 and TN Pb Analyzer); two are manufactured by Metorex Inc. (the X-MET 920-P Analyzer and the X-MET 920-MP Analyzer); one is manufactured by HNU Systems, Inc. (the SEFA-P Analyzer); and one is manufactured by Scitec Corporation (the MAP Spectrum Analyzer). The X-MET 940, a prototype FPXRF analyzer developed by Metorex, was given special consideration and replaced the X-MET 920-P for a portion of the demonstration. This environmental technology verification report (ETVR) presents information relative to the Metorex X-MET 920-MP. Separate ETVRs have been published for the other analyzers demonstrated.

Approximately three days of operational downtime was experienced by the analyzer due to computer software and hardware problems. Most of these problems were due to operator error or inexperience. None of the downtime or data loss was associated with mechanical or electronic malfunctions of the analyzer. Quantitative data was provided by the analyzer on a real-time basis. The X-MET 920-MP Analyzer reported arsenic, chromium, copper, lead, zinc, nickel, and barium. This analyzer used count times ranging from 30 live-seconds for *in situ*-unprepared samples at the ASARCO site to 180 live-seconds for intrusive-prepared samples at the RV Hopkins site. These count times resulted in a sample throughput averaging between 8 and 14 samples per hour. The X-MET 920-MP Analyzer provided definitive data

(equivalent to reference data) for arsenic and lead; and qualitative screening level data (identifies the presence or absence of a contaminant) for copper, barium, and zinc. Insufficient precision data precluded an assignment of data quality levels for nickel or chromium.

This study showed that the analyzer produced data that exhibited a $\log_{10}$-$\log_{10}$ linear correlation to the reference data. The analyzer generally exhibited a precision similar to the reference methods. The analyzer exhibited precision of less than 10 percent relative standard deviation at 5 to 10 times the method detection limit (MDL) for all of the reported analytes except chromium and nickel. The precision evaluation was confounded by changing count times. The precision study indicated that count times probably had no effect on the precision for all target analytes except copper and lead. For copper and lead, the increasing count times caused a 2- to a 10-fold increase in precision. The analyzer’s quantitative results were based on an empirical calibration using site-specific calibration samples.

This demonstration found that the X-MET 920-MP Analyzer was generally simple to operate in the field; however, its physical configuration made it more practical for use as a benchtop unit. The auxiliary computer and cumbersome power requirements of commercial laptop computers limited its utility as an *in situ* instrument. The operator required no specialized experience or training for normal operation of the analyzer. However, ownership and operation of this analyzer may require specific licensing by a state nuclear regulatory agency. There are specific radiation safety training requirements and costs associated with this type of license.

The Metorex X-MET 920-MP Analyzer can provide rapid, real-time analysis of the metals content of soil samples at hazardous waste sites. The analyzer can quickly identify contaminated areas from noncontaminated areas allowing investigation and remediation decisions to be made more efficiently on-site and reduce the number of samples that need to be submitted for confirmatory analysis.

Table of Contents

Section **Page**

Section **Page**

List of Figures

List of Tables

List of Abbreviations and Acronyms

α	alpha
β	beta
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
Co^{57}	cobalt-57
CRM	certified reference material
DC	direct current
EPA	Environmental Protection Agency
ERA	Environmental Resource Associates
ETVR	environmental technology verification report
eV	electron volt
FPXRF	field portable X-ray fluorescence
ICAL	initial calibration
ICB	initial calibration blank
ICP-AES	inductively coupled plasma-atomic emission spectroscopy
ICS	interference check standard
ICV	initial calibration verification
IDL	instrument detection limit
IDW	investigation-derived waste
keV	kiloelectron volt
LCD	liquid crystal display
LCS	laboratory control samples
$\log_{10}$	base 10 logarithm
LRL	lower reporting limit
MCA	multichannel analyzer
mCi	millicurie
MDL	method detection limit
mg/kg	milligram per kilogram
mL	milliliter
mm	millimeter
MMTP	Monitoring and Measurement Technologies Program
mrem/hr	millirems per hour
MRI	Midwest Research Institute
NERL-ESD	National Exposure Research Laboratory—Environmental Sciences Division

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

Acknowledgments

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

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 FPXRF 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 economical resources needed to operate these technologies in the field, and (4) to test and validate an SW-846 draft method for FPXRF analysis. Secondary objectives for this demonstration were to evaluate FPXRF analyzers for their reliability, ruggedness, cost, range of usefulness, data quality, and ease of operation.

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

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

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

The target analytes for this demonstration were selected from the Resource Conservation and Recovery Act's (RCRA) Toxicity Characteristic (TC) list, analytes known to have a high aquatic toxicity and 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.

To demonstrate these analyzers, hazardous waste sites in Iowa (the RV Hopkins site) and in the State of Washington (the ASARCO site) were selected. These sites were 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 X-MET 920-MP Analyzer was generally simple to operate in the field. However, unexpected computer software and hardware problems caused data loss and project downtime. In addition, the physical configuration of the analyzer made its use in the *in situ* mode cumbersome. Its physical configuration is more adapted to a benchtop application. The developer provided a training course for the operator that was similar to that provided to a purchaser of the equipment. The training encompassed enough FPXRF theory and hands-on analyzer use to allow the operator to calibrate the analyzer, manipulate the data collection software, and adjust instrument parameters such as count times and target analytes. Metorex provided accessible and timely field support. The analyzer was portable and capable of operating continuously over a 12-hour work day with appropriate battery changes. The data collection and interpretation software is designed to operate with an auxiliary computer system. A laptop computer was used. The field portability was reduced by the size and power requirements of the laptop computer. The computer could only operate 1.5 to 2 hours on a battery. The environmental conditions encountered at the ASARCO site caused no operational downtime for the analyzer.

The analyzer used one radioactive source coupled with a gas proportional detector. The type and strength of the analyzer’s source allow it to produce reliable data at count times as short as 30 live-seconds. The count times used in this demonstration resulted in a sample throughput averaging 8 - 14 samples per hour.

The X-MET 920-MP Analyzer produced data meeting definitive level criteria (equivalent to reference data) for arsenic and lead; and produced qualitative screening level data (identifies the presence or absence of a contaminant) for barium, zinc, and copper. Data quality levels for nickel and chromium were not assigned due to lack of sufficient data.

The analyzer exhibited a similar precision relative to the reference methods. The chromium and nickel data generally showed the poorest precision of the reported analytes. No values were reported for these analytes at the 5 to 10 times MDL range. Field-based method detection limits (MDL) for this analyzer were higher than precision-based or developer-provided MDLs. However, this relationship was confounded by increased count times with increasing sample preparations. Site and soil texture did not appear to affect data comparability. The probe’s high counting efficiency and the empirical calibration seemed to minimize any impact of inter-element interferences.

Based on the performance of the analyzer, this demonstration found it to be a useful tool for characterizing the concentration of select metals in environmental soil samples. As with all 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-MP Analyzer. This analyzer was developed by Metorex, Inc. to perform elemental analyses (metals quantitation) in the field. This analyzer uses a gas-filled proportional detector and cadmium-109 (Cd^{109}) source to quantitate metals concentrations. The analyzer can be operated in either an *in situ* or intrusive mode. The *in situ* mode, commonly called a "point-and-shoot" mode, requires that the point of measurement on the soil surface is cleared of loose debris and organic matter, the analyzer's 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 then placed in a sample cup. The sample cup is placed into an analysis chamber on the probe and a measurement is taken.

This section summarizes 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-MP Analyzer, capabilities, equipment, accessories, accuracy, precision, comparability to reference methods, and other evaluation factors. Section 5 discusses potential applications of the analyzer, presents a method for data correction, and suggests a framework for a standard operating procedure (SOP). Section 6 lists the 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 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, range of usefulness, data quality, and ease of operation. The performances of the FPXRF analyzers were not compared against each other. Instead, the performance of each analyzer was independently and individually compared to that of standard conventional analytical methods commonly used in regulatory enforcement or compliance activities. In addition, the analyzer's performances were 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 project'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 discussion, the term "reference" was substituted for "confirmatory" since the data was used as a baseline for comparison. In accordance with Federal Acquisition Regulations, MRI was awarded a subcontract to analyze soil samples using the reference methods. The award was made based on MRI's costs, ability to meet the demonstration's quality assurance project plan (QAPP), requirements, and as the only commercial laboratory 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 use of the data. Data quality parameters usually include five indicators of data quality often known as the PARCC parameters: precision, accuracy, representativeness, completeness, and comparability. In addition, method detection limits (MDLs) are also used to assess data quality.

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

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

Accuracy refers to the difference between a sample result and the reference or true value. Bias, a measure of the departure from perfect accuracy, can be calculated 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 to provide results that reflect the most accurate and precise measurement it was capable of achieving. The combination of existing method requirements supplemented by the demonstration QAPP provided the guidance to assure optimum performance of the method. Representativeness was assessed by evaluating calibration standards, method blank samples, duplicate samples, and PE samples.

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

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

Site Selection

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

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

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

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

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

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

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

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

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

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

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

Predemonstration Sampling

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

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

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

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

Experimental Design

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

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

The X-MET 920-MP Analyzer can be operated in either an *in situ* or intrusive mode. These two modes of analysis involve slightly different measurement and preparation procedures. These procedures allowed for an evaluation of the effects of sample preparation or FPXRF comparability to the reference data. For *in situ* analysis, an area 4 inches by 4 inches square was cleared of vegetation, debris, and gravel larger than 2 millimeters (mm) in diameter. Each analyzer took one *in situ* measurement in each sample area. These data represented FPXRF *in situ* measurements for unprepared oils (*in situ*-unprepared). Replicate measurements were taken at 4 percent of these locations to assess analyzer precision. Figure 2-1 depicts the sample analysis flowchart for both *in situ* and intrusive analyses.

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

The homogenized sample was then spread out inside a 1-inch-deep petri dish. The 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.

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

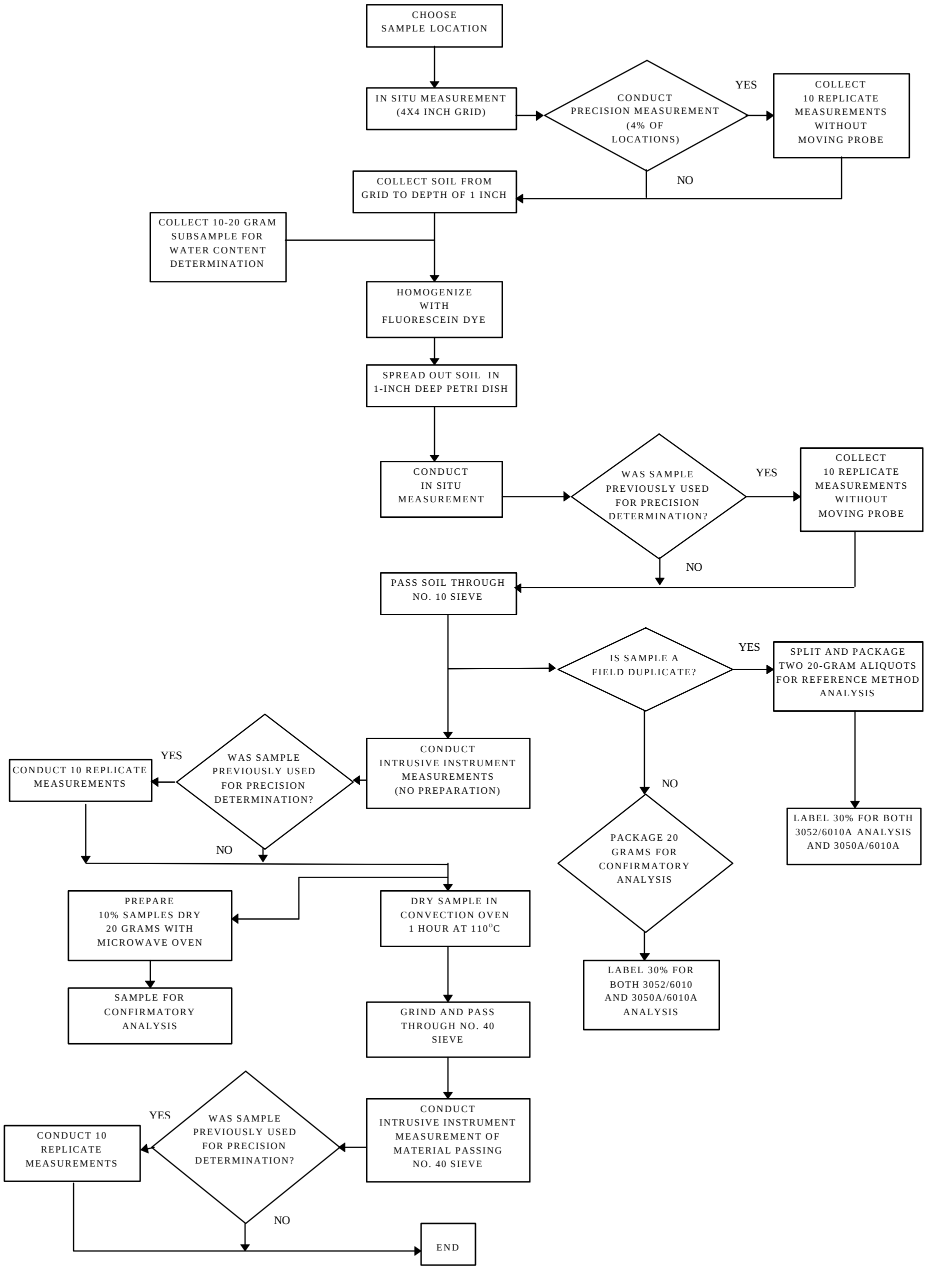

Figure 2-1. Sample Preparation and Analysis: This flowchart depicts the handling procedures for each sample taken during the demonstration.

Following the intrusive-unprepared analysis, 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). The sample was 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).

Qualitative Factors

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

Many analytical methods exhibit "operator effects," in which individual differences in sample preparation or operator technique result in a significant effect on the numerical results. To reduce the possible influence of operator effects, a single operator was used to operate each FPXRF analyzer. While this reduced some 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 for all the FPXRFs demonstrated except the X-MET 920-MP. Due to unforeseen circumstances, a second operator was needed to analyze samples for the RV Hopkins site. Sample preparation variation effects were minimized by using the same personnel to prepare samples. To eliminate the influence of operator effects on the reference method analysis, only one reference laboratory was used to analyze the samples. Based on this design, there could be no quantitative estimate of "operator" effect.

Quantitative Factors

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

The data developed by the FPXRF analyzers were compared to reference data for the following primary analytes: arsenic, barium, chromium, copper, lead, and zinc; and for the following secondary analytes: nickel, iron, cadmium, and antimony. The specific target analytes determined by the X-MET 920-MP in this demonstration were arsenic, barium, copper, chromium, lead, nickel, and zinc.

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

ASARCO data was divided into sand and loam soil textures; each soil texture was subdivided by the soil preparation steps. This design allowed for the examination of particle size and sample homogenization. These effects were believed to have the greatest impact on data comparability.

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

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.

In addition to the quantitative factors discussed above, the common FPXRF sample preparation technique of microwave drying 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 temperature range achievable during microwave drying. To assess this effect, 10 percent of the homogenized, crushed, oven-dried, and sieved (No. 40) 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 was only evaluated for the reference laboratory in an attempt to identify potential effects on data comparability.

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

An important health and safety issue during the demonstration was the effectiveness of radioactivity shielding of each FPXRF analyzer. Qualitative radiation readings were 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 lease, expendable supplies, such as liquid nitrogen and sample cups, and nonexpendable costs, such as labor, licensing agreements for the radioactive sources, operator training costs, and disposal of investigation-derived waste (IDW). This information is provided to assist the user in 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 using the two intrusive sample preparation procedures. 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.

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

Evaluation of Analyzer Performance

Metals concentrations measured by each analyzer were compared to the corresponding reference laboratory data and to other QA/QC sample results. These comparisons were conducted independently for each target analyte. These measurements were used to determine an analyzer's accuracy, data quality level, method precision, and comparability to reference methods. PE 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 from which the data are obtained is not identical. While there is significant overlap in the nature of the analysis, there are also major differences. These differences, or "perspectives," allow the user to characterize the same sample in slightly different ways. Both methods have a role in site characterization and remediation. It is important to consider these differences and the measurement error intrinsic to each method when comparing the FPXRF method against a reference analytical method.

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

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

The first assumption requires that the independent variable (reference data) and the dependent variable (FPXRF data) are linearly related and are not described 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 the FPXRF data and the reference data are, in fact, related linearly and that this assumption is correct.

The second assumption requires that the error be normally distributed, the sum to equal zero, be independent, and exhibit a constant error variance for the data set. Figure 2-2 illustrates that for raw data, this assumption is not correct (at higher concentrations the scatter around the regression line increases), but that for the logarithmic transformation (shown as a log-log plot) of the data, this assumption is valid (the scatter around the regression line is 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.

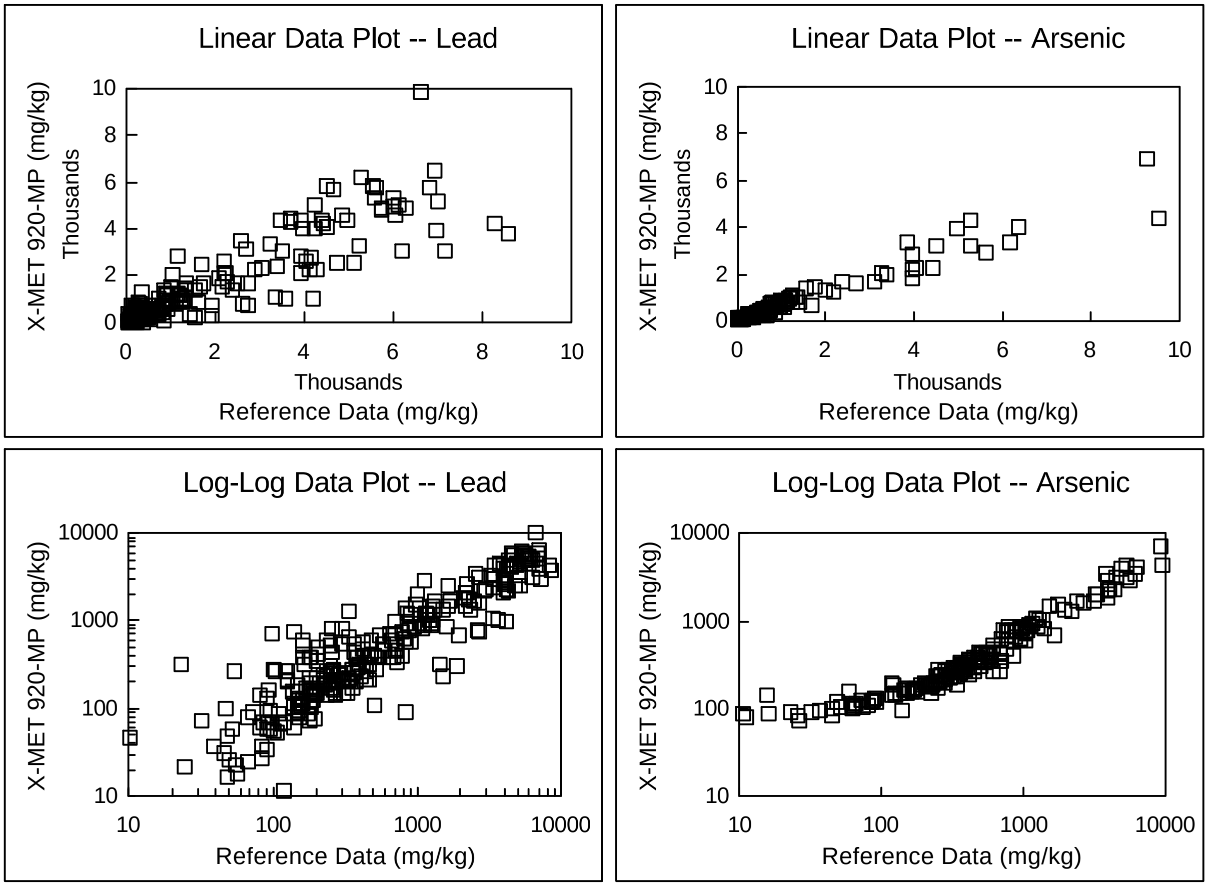

Figure 2-2. Linear Log-log Data Plots: These graphs illustrate the linear relationship between the FPXRF data and the reference data. The linear data plots illustrate the concentration dependence of this relationship with the increased scatter at higher concentrations. The log-log plots eliminate this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

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

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

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

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

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

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

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

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

Table 2-2. Criteria for Characterizing Data Quality

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

Notes:

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

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

r^2 Coefficient of determination.

RSD Relative standard deviation.

At the time of this demonstration, approved EPA methods for FPXRF did not exist. As part of this demonstration, PRC prepared draft Method 6200 "Field Portable X-Ray Fluorescence Spectrometry for the Determination of Elemental Concentrations in Soil and Sediment." The draft method has been submitted for inclusion in Update 4 of SW-846 scheduled for approval in FY-97. For purposes of this demonstration, the absence of an EPA-approved final method did not preclude the analyzers' data from being considered definitive. The main criterion for data quality level determination was based on the comparability of each analyzer's data to that produced by the reference methods, as well as analyzer-specific criteria such as precision.

The comparability data set for the X-MET 920-MP Analyzer consisted of matched pairs of FPXRF and reference data for each target analyte. Each 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 for an independent assessment of the influence of each variable on comparability.

For the purposes of this demonstration, a total of 315 soil samples was collected for analysis. These samples were analyzed by the reference methods and by the FPXRF analyzer up to four times, using each of the four sample preparation steps. This would produce 1,260 data values for these analyzers which used all four of the sample preparation steps. 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 an examination of the lead and zinc data. These were the only primary analytes that exhibited a wide distribution of concentrations across all sites and soil textures. The effects of sample preparation were evaluated for all target analytes. If the evaluation of the influence of a given variable did not result in a better correlation, as exhibited by a higher coefficient of determination (r^2) and smaller standard error of the estimate (using $\log_{10}$ transformed data), then the influence was considered to be insignificant. However, if the correlation worsened, the cause was examined and explained. If the correlation improved, resulting in a higher r^2 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. A linear regression analysis was then conducted on these data sets. If this did not result in improved r^2 values and reduced standard errors of the estimate, then the relationship between the analyzer's $\log_{10}$ transformed data and the $\log_{10}$ transformed reference data was considered linear over the entire range of concentrations encountered during the demonstration. This would mean that there was no concentration effect.

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

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

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

where

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

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

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

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

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

where

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

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

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

where

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

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

transformed data meeting these criteria were considered statistically equivalent to the $\log_{10}$ transformed reference data.

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 ideal values, the analyzer was considered to produce quantitative screening level data quality. However, the linear regression was deemed sufficiently significant that bias could be identified and corrected. Results in this case could be mathematically corrected if 10 - 20 percent of the samples are sent to a reference laboratory. Reference laboratory analysis results for a percentage of these samples would provide a basis for determining a correction factor.

Data placed in the qualitative screening level category exhibit r^2 values less than 0.70. These data either were not statistically similar to the reference data based on inferential statistics or had a precision of 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 of samples exhibiting contamination 5 to 10 times the estimated detection levels of the analyzers were multiplied by 3. The result represents the MDL for the analyzer.

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 at a point below the MDL. By determining the point where the linear relationship disintegrates, the MDL is assigned at two SDs above this concentration. This procedure represents a field- or performance-based MDL.

Deviations from the Demonstration Plan

Seven deviations were made from the demonstration plan during on-site activities. The first dealt with the determination of the moisture content of the 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 immediately adjacent to the original sample location was used for determining moisture content. This was done to conserve sample volume 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, the samples for moisture content determination were dried in a convection oven at 110 °C for 1 hour.

The third deviation involved an assessment of analyzer drift due to changes in temperature. The demonstration plan required 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. This interval is within common microwave drying times used in the field.

An analysis of the microwaved samples showed that this process had a significant impact on the analytical results. The mean RPD for the microwaved and nonmicrowaved raw data were significantly different at a 95 percent confidence level. This suggests that the microwave drying process somehow increases error and sample concentration variability. This difference may be due to the extreme heat that altered 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 up to 3 days 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 homogenous 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 complete sample homogenization. Approximately one-quarter teaspoon of dry sodium fluorescein salt was added to each sample prior to homogenization. After mixing, the sample was examined under an ultraviolet light to assess the distribution of sodium fluorescein throughout the sample. If the fluorescent dye was evenly dispersed in the sample, homogenization was considered complete. If the dye was not evenly distributed, the mixing was continued and 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 all of 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 homogenization routine, at the scale of analysis used by this demonstration, the samples were considered to be completely homogenized.

Section 3
Reference Laboratory Results

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

Reference Laboratory Methods

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

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

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

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

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

Reference Laboratory Quality Control

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

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

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

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

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

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

Notes:
[a] Quality control parameters were evaluated on the raw reference data.
[b] RPD control limits only pertain to original and laboratory duplicate sample results that were greater than 10 times the instrument detection limit (IDL).
[c] RPD control limits only pertain to original and laboratory duplicate sample results that were greater than or equal to 10 times the LRL.

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

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

Quality Control Review of Reference Laboratory Data

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

Reference Laboratory Sample Receipt, Handling, and Storage Procedures

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

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

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

Sample Holding Times

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

Initial and Continuing Calibrations

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

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

Detection Limits

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

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

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

Analyte	LRL (mg/kg)	Analyte	LRL (mg/kg)
Antimony	6.4	Copper*	1.2
Arsenic*	10.6	Iron	600[a]
Barium*	5.0	Lead*	8.4
Cadmium	0.8	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.

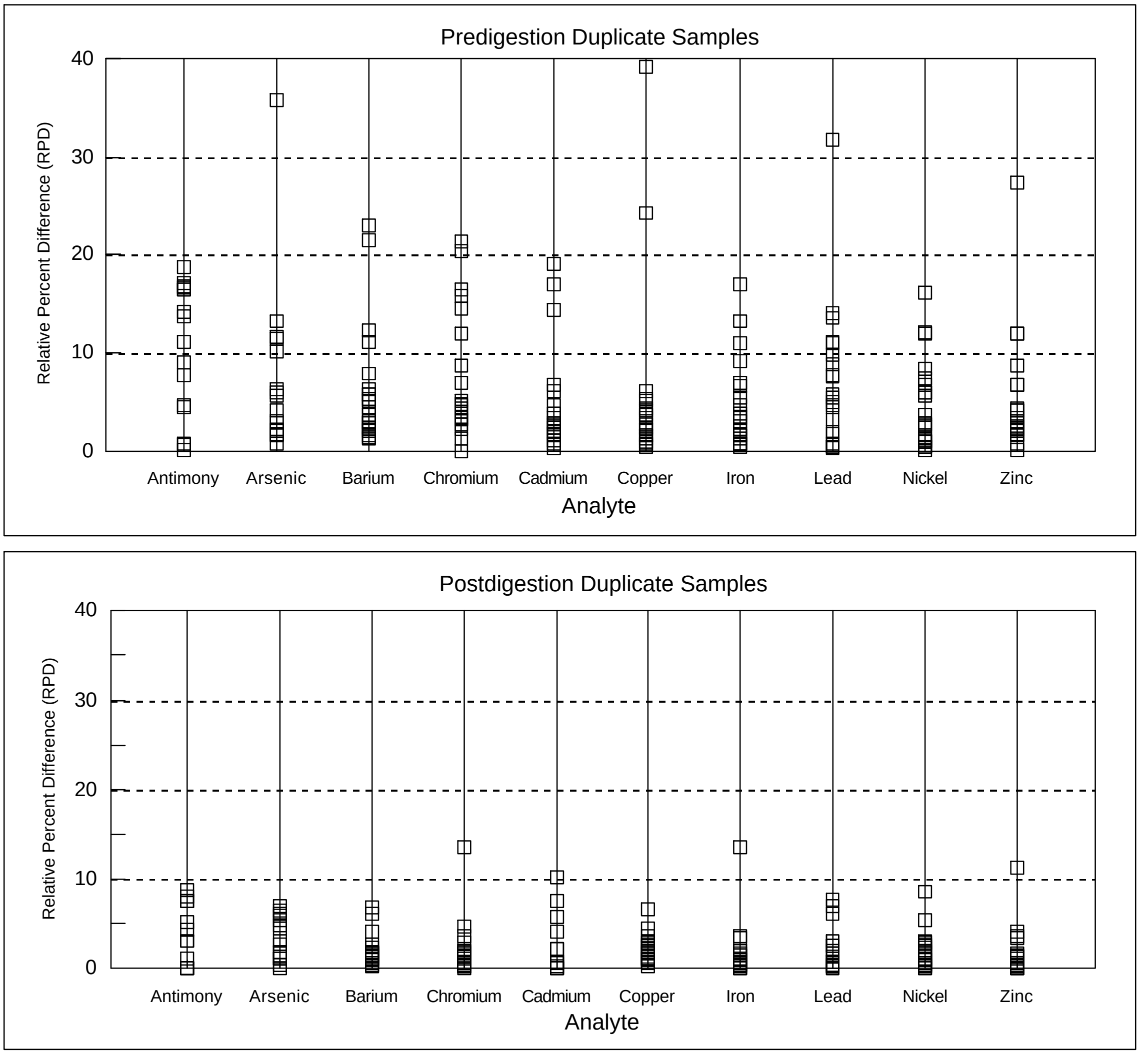

Figure 3-1. Pre- and Postdigestion Duplicate Samples: The top graph illustrates the reference laboratory's performance on analyzing predigestion duplicate samples. Twenty percent RPD represents the predigestion duplicate control limits defined in the demonstration QAPP. Two points were deleted from this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

Accuracy

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

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

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

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

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

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

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

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

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

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

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

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

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

Representativeness

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

Completeness

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

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

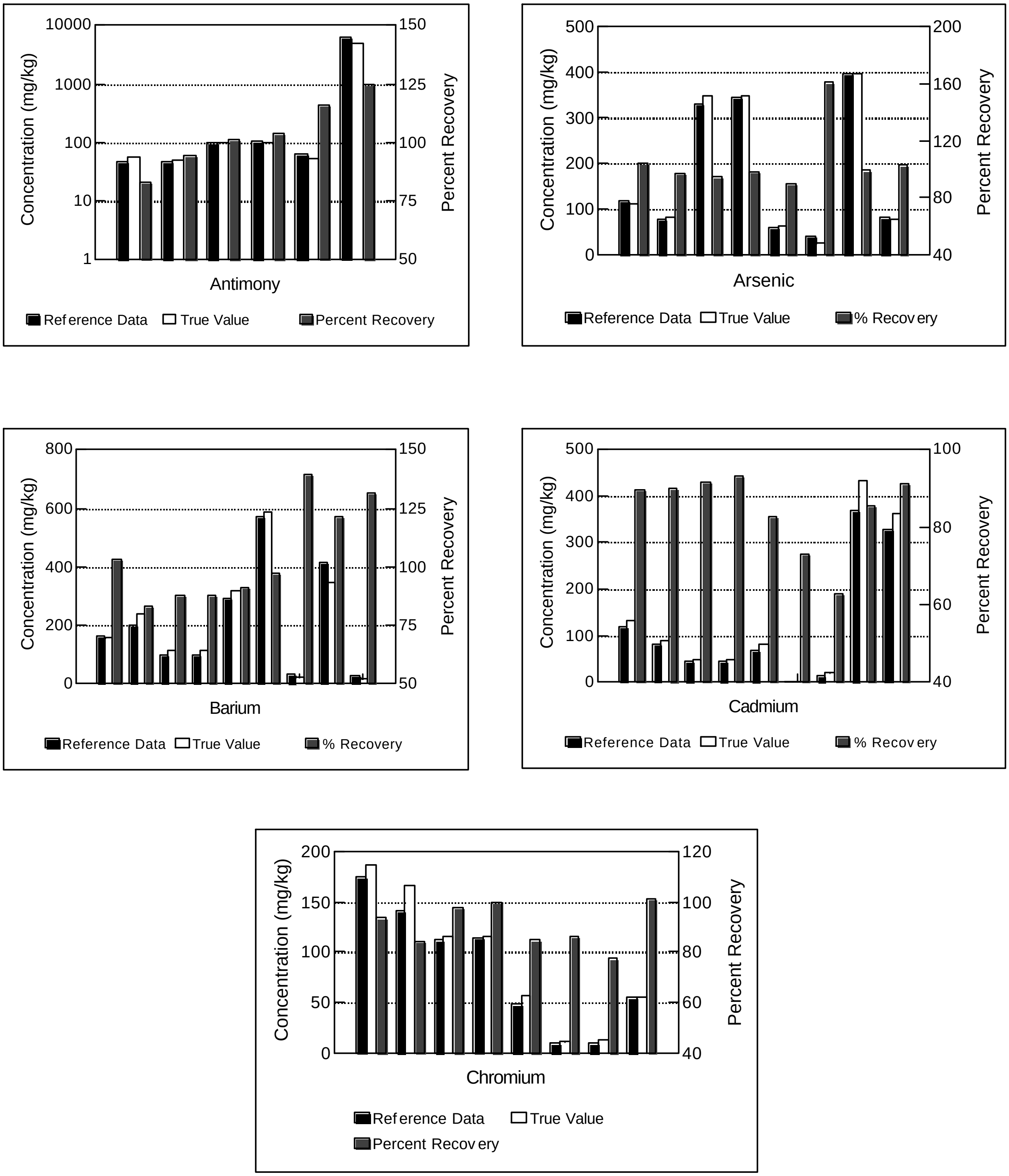

Figure 3-2. Reference Method PE and CRM Results: These graphs illustrate the relationship between the reference data and the true values for the PE or CRM samples. The gray bars represent the percent recovery for the reference data. Each set of three bars (black, white, and gray) represents a single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

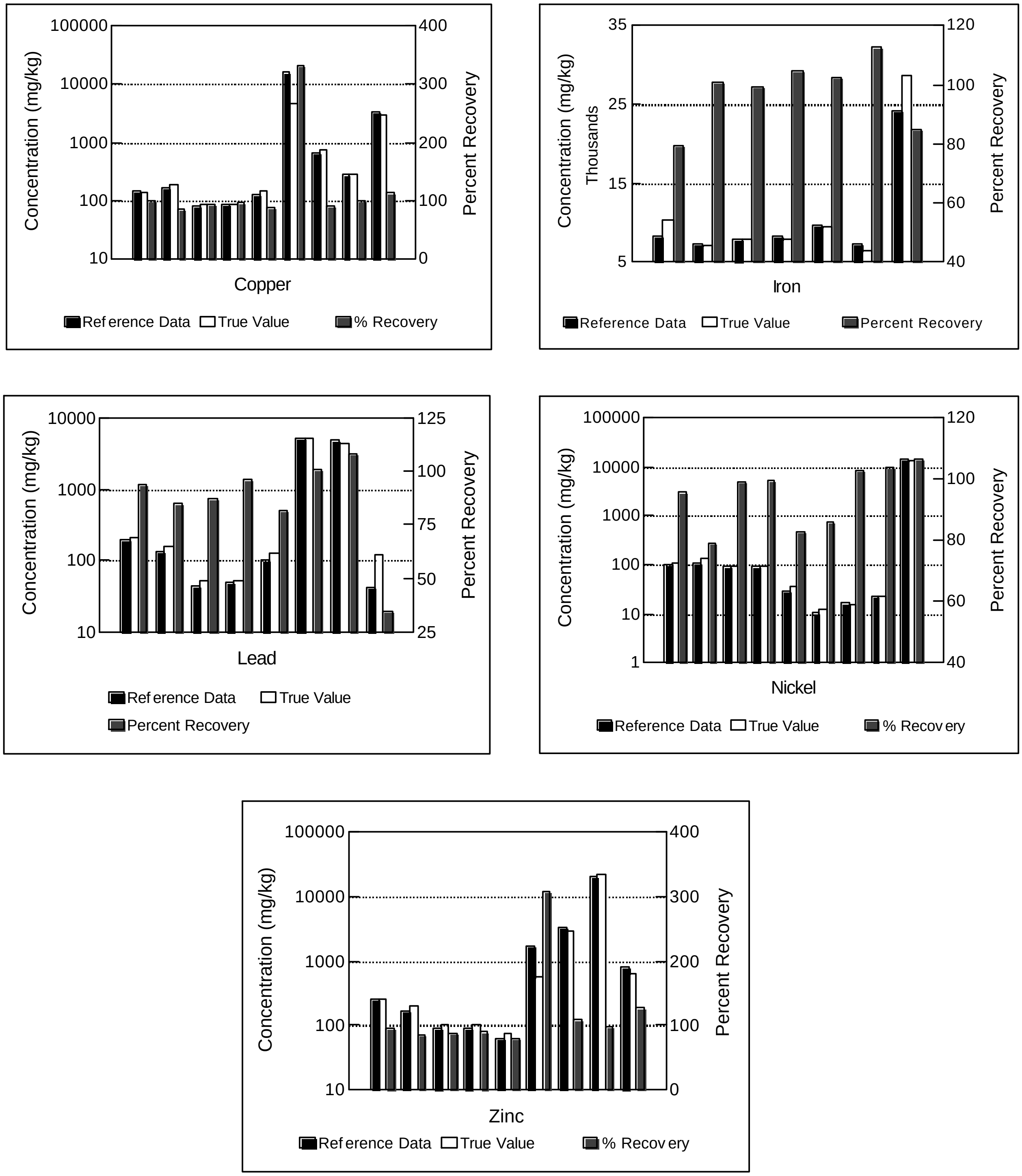

Figure 3-2 (Continued). Reference Method PE and CRM Results: These graphs illustrate the relationship between the reference data and the true values for the PE or CRM samples. The gray bars represent the percent recovery for the reference data. Each set of three bars (black, white, and gray) represents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

Comparability

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

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

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

Table 3-4. SRM Performance Data for Target Analytes

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

Notes:

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

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

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

Notes:

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

NIST National Institute of Standards and Technology.

SRM Standard reference materials.

– Analyte not present above the method LRL.

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

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

Use of Qualified Data for Statistical Analysis

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

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

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

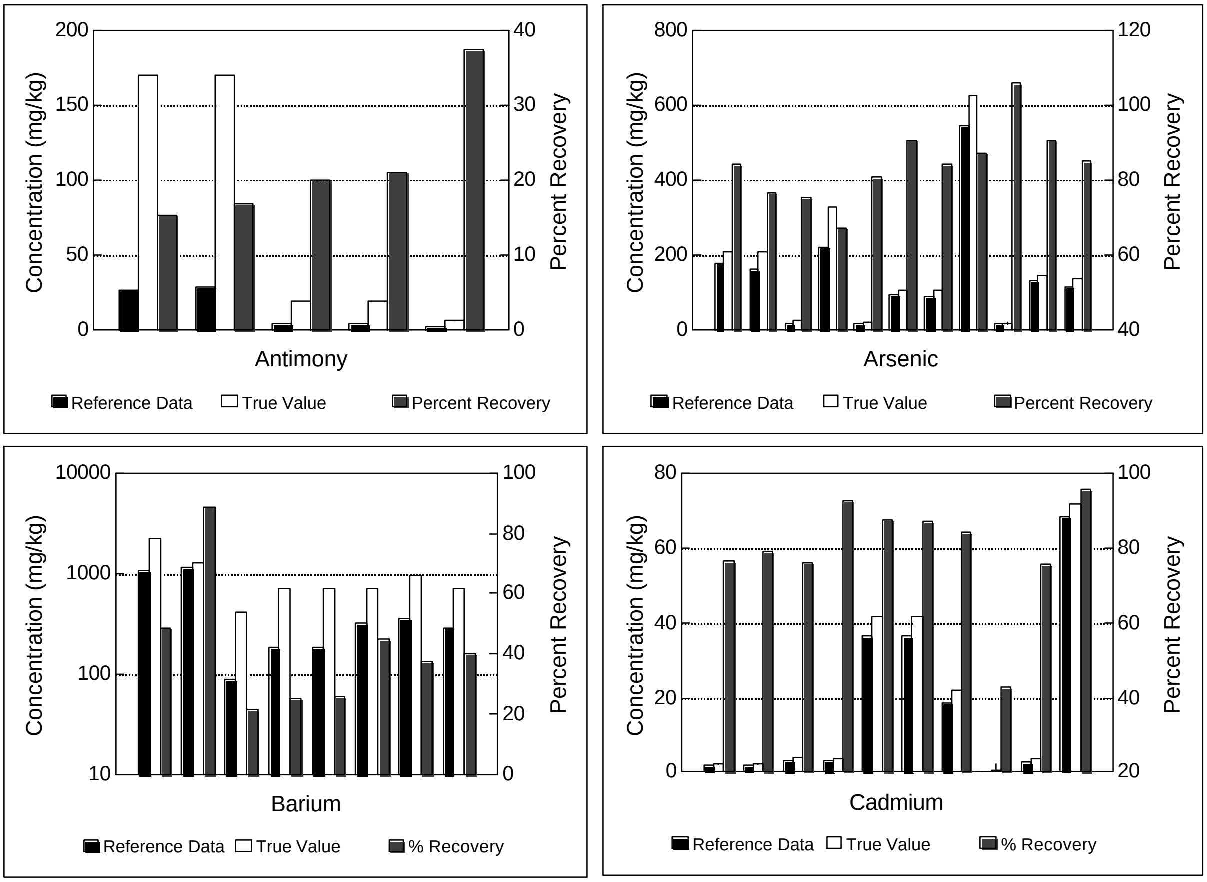

Figure 3-3. Reference Method SRM Results: These graphs illustrate the relationship between the reference data and the true values for the SRM samples. The gray bars represent the percent recovery for the reference data. Each set of three bars (black, white, and gray) represents a single SRM sample.

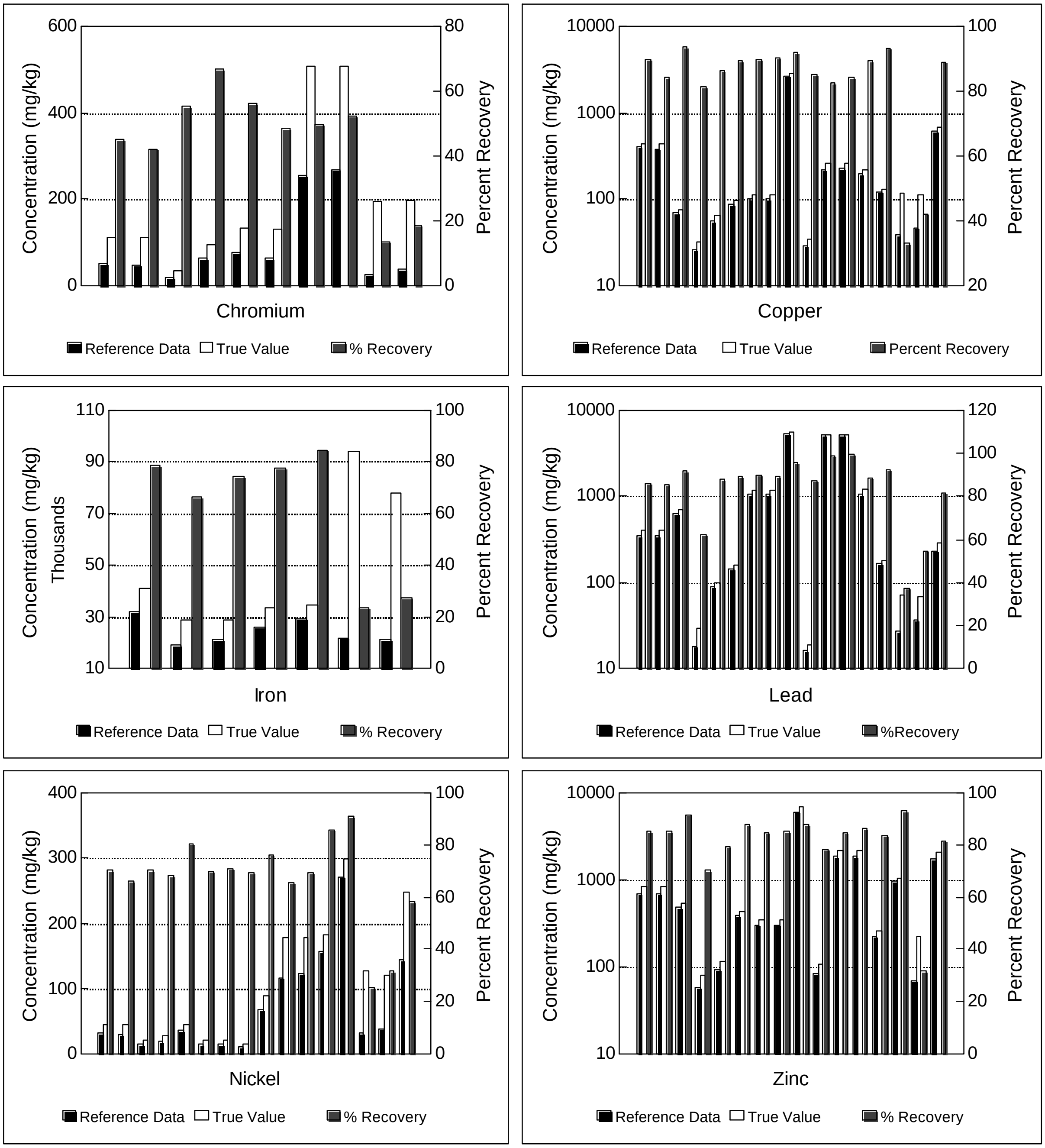

Figure 3-3 (Continued). Reference Method SRM Results: These graphs illustrate the relationship between the reference data and the true values for the SRM samples. The gray bars represent the percent recovery for the reference data. Each set of three bars (black, white, and gray) represents a single SRM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

Section 4
X-MET 920-MP Analyzer

This section provides information on Metorex's X-MET 920-MP Analyzer, including the theory of FPXRF, operational characteristics, performance factors, a data quality assessment, and a comparison of its 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 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 to each element. This remission of X-rays is termed XRF.

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

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.

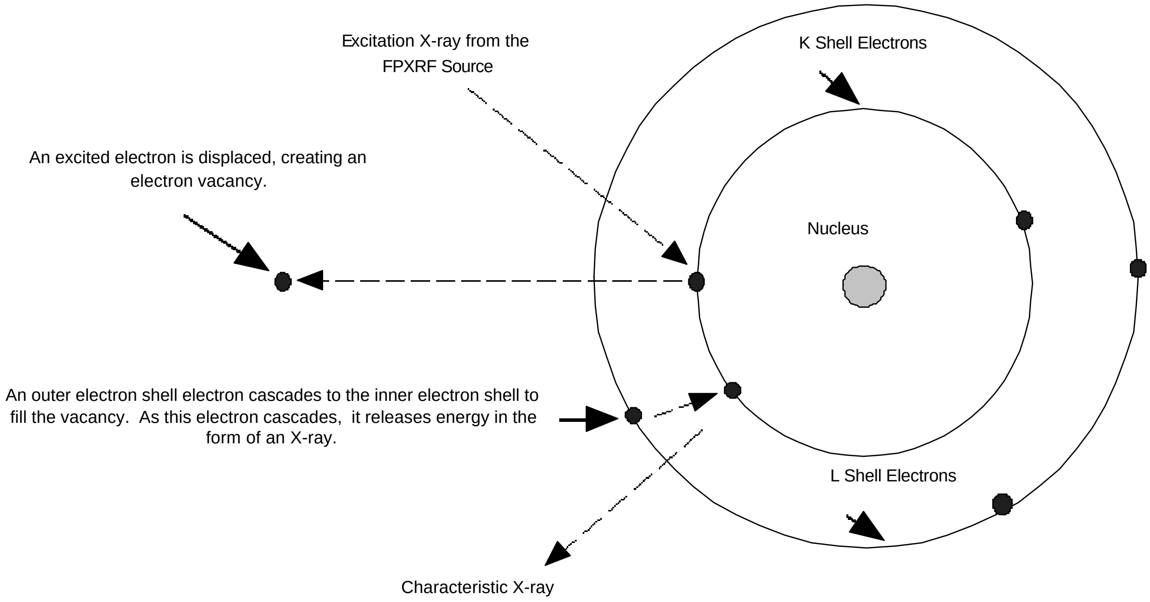

Figure 4-1. Principle of Source Excited X-ray Fluorescence: This figure illustrates the dynamics of source excited X-ray fluorescence.

For a given element, the X-rays emitted from L transitions are always less energetic than those emitted from K transitions. Unlike the K-lines, the L-lines (L_α and $L_ß$) 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.

An X-ray source can excite characteristic X-rays from an analyte only if its energy is greater than the electron binding energies for the target analyte. The electron binding energy, also known as the absorption edge energy, represents the amount of energy an electron has to absorb before it is displaced. The absorption edge energy is somewhat greater than the corresponding line energy. Actually, the K-absorption edge energy is approximately the sum of the K-, L-, and M-line energies of the particular element, and the L- absorption edge energy is approximately the sum of the L- and M-line energies. FPXRF analytical methods are more sensitive to analytes with absorption edge energies close to, but less than, the excitation energy of the source. For example, when using a Cd^{109} source, which has an excitation energy of 22.1 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 can be configured with the hardware and software needed to address a specific analysis problem. With the X-MET 920 line, Metorex offers several analysis probes. The probe contains 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 line of instruments to perform elemental analysis in the petroleum and petrochemical industry, the mining and minerals industry, and the environmental field. The X-MET 920-MP is a field portable technology that can be operated in the *in situ* or intrusive mode. The X-MET 920-MP is Metorex's miniportable version on the X-MET 920 line. The analyzer uses energy dispersive XRF spectroscopy to determine elemental composition of soils and other solid waste materials as well as liquids and slurries. The X-MET 920-MP can identify and quantify 32 elements simultaneously, and 70 elements can be identified and quantified when all available sources are used. Metorex offers four excitation sources, iron-55 (Fe^{55}), cadmium-109 (Cd^{109}), americium-241 (Am^{241}), and curium-244 (Cm^{244}), and two detectors, [Si(Li) and gas-filled proportional counter] in its probes. For this demonstration, the analyzer used the surface analysis probe system (SAPS) equipped with a Cd^{109} source and a gas-filled proportional detector. The SAPS is designed to only house one excitation source. Metorex also offers a dual source version of the SAPS probe called the double source probe system (DOPS). The choice of the source is determined by the user depending on the target analytes.

For *in situ* analyses, the probe is pointed downward against the soil surface to allow the source-detector 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 inside the protective sample cover for analysis. In either mode, sample measurement is initiated by depressing a trigger built into the handle of the probe. This exposes the sample to primary radiation from the excitation source. Fluorescent and backscattered radiation reenters the probe through the probe window and is counted by the detector, which then transmits the electronic signal to the MCA for processing. The X-MET 920-MP is operated and calibrated using the X-MET software, which uses an empirical calibration.

Operational Characteristics

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

Equipment and Accessories

The X-MET 920-MP comes with all the equipment necessary for both *in situ* and intrusive operation. Metal, foam-padded carrying cases are provided for storage of the XPCS unit and the analysis probe. Specifications for the X-MET 920-MP are provided in Table 4-1.

Three main components comprise the analytical system: a laptop computer, the electronics unit or portable XPCS, and the probe. The XPCS contains a 2,048-channel MCA that is used to collect, analyze, and display the data. It is housed in a rugged, weatherproof, self-contained unit. The portable XPCS is about the size of a laptop computer and weighs approximately 5 pounds. It can be powered from 110 or 220V AC electricity or from 10 internal C-sized rechargeable nickel-cadmium batteries. With a full charge, the batteries last 7 to 8 hours before a recharge is needed. The XPCS comes standard with 10 rechargeable batteries and a charger unit and is equipped with a cable that interfaces to an RS-232 port of a laptop computer. It also contains a port for the cable that plugs into the probe.

Table 4-1. Analyzer Instrument Specifications

Characteristic	Specification
Resolution	760 eV (Manganese-K_α)
Source	10 mCi Cd^{109} (Fe^{55}, Am ^{241}, and Cm^{244} also available)
Detector	Gas-filled Proportional (Si(Li) also available)
Probe Size	20 cm x 25 cm x 5 cm
Probe Weight	2.3 kilograms (kg)
XPCS Size	28 cm x 23 cm x 4.5 cm
XPCS Weight	2.3 kg (with batteries)
XPCS Data Acquisition	2,048-channel MCA
XPCS Communication	Via RS-232 port at baud rate of 38.4 kilobytes
Computer Requirements	IBM compatible PC with MS DOS, 486 DX CPU, 4 megabyte random access memory, 170 MB hard disk, VGA graphics, RS-232 port and RS-232 cable, and parallel printer
Power Source	120V or 220V alternating current or batteries
Operational Checks	Gain check and site-specific check sample once per hour
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 920-MP uses the laptop computer's RS-232 port to interface directly with the portable XPCS. The requirements of the laptop computer are also listed in Table 4-1. The laptop computer can be supplied by Metorex or by the user. If Metorex supplies the laptop computer, the X-MET software will already be installed. If the user supplies the laptop computer, Metorex will provide instructions to install the software. During this demonstration, Metorex supplied a laptop computer with the X-MET software already installed. PRC provided two more laptop computers, one owned and one rented. Metorex helped the operator install the X-MET software on both. All three computers were used during the demonstration. As the data is collected and processed by the portable XPCS, the spectra and results files are saved on the hard drive of the laptop computer. The data can also be sent to a printer for a hard copy printout.

The probe used with the X-MET 920-MP Analyzer during this demonstration was the SAPS probe. It weighed approximately 3 pounds and used a 10 millicurie (mCi) Cd^{109} radioisotope source for excitation of the target analytes. The Cd^{109} source was installed in March 1995, 1 month before the demonstration. The radioisotope source was in the form of an 8-mm-diameter by 5-mm-thick capsule. The probe's beryllium sample window is environmentally sealed by a 25-mm-diameter window of clear Kapton™ film.

The SAPS probe contains a gas-filled proportional detector. This detector is a chamber filled with a noble gas and fitted with a central wire electrode biased to several hundred volts. An X-ray photon entering the chamber ionizes the gas atoms. The electric charge produced is collected and provides an electric pulse whose magnitude (amplitude) is directly proportional to the energy of the X-ray photon absorbed by the gas in the detector. This detector is rugged and lightweight, it operates at ambient temperatures from -13 to 140 °F. The detector achieves a manganese K_α X-ray resolution of 760 eV. This detector has a lower resolution than the Si(Li) detector (170 eV); however, it has a higher efficiency because it has a larger collection area than the Si(Li) detector. The sensitivity of the detector is the sum of the resolution and efficiency.

The SAPS 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 attaches to the probe over the sample window. Sample measurements are initiated by depressing a trigger built into the handle of the probe. The probe is connected to the XPCS with a coiled, flexible cord that is about 6 feet long.

Metorex used pure element samples for the calibrations conducted at each site. To perform either an empirical or a fundamental parameters (FP) calibration, a pure element standard for each target analyte is required. These pure element standards are in the shape of a coin that is about the size of a $0.50 piece. They are large enough to completely cover the probe's Kapton™ window. Metorex also brought known calibration standards to the test sites. These standards included several NIST SRMs and soil samples that were analyzed by Metorex. Metorex also brought the predemonstration split samples to each site. Other equipment and supplies that were helpful when using the X-MET 920-MP, and are not supplied by the developer, include protective gloves, paper towels, a marking pen, a portable generator, and an extension cord. The generator and extension cord were used because of battery problems associated with the laptop computers.

Operation of the Analyzer

The operation of the X-MET 920-MP can be described in three steps: (1) performing the calibration, (2) taking measurements, and (3) managing the data. The X-MET 920-MP can be calibrated empirically or through the use of FPs. In this demonstration, the X-MET 920-MP was used with the SAPS probe containing a gas-filled proportional detector. When using this detector, only the empirical calibration can be performed. The FP calibration relies on the ability of the detector to fully resolve the peaks from neighboring elements. The gas-filled proportional detector cannot adequately resolve these peaks; therefore, the empirical calibration must be used. A separate empirical calibration was performed at each site.

The empirical calibration is performed using the "Calibration" menu of the X-MET software. Metorex recommends that 10 - 20 site-specific calibration standards (SSCS) be used to perform the empirical calibration. The SSCSs have known analyte concentrations usually determined by analytical methods such as inductively coupled plasma or atomic absorption spectroscopy. If SSCSs are not available, site-typical calibration samples that closely approximate the site's soil matrix with respect to particle size distribution, mineralogy, and contaminant analytes can be used.

The first step in the empirical calibration is to analyze the pure element samples for the target elements plus any elements that may cause interferences at a measurement time of 200 live-time seconds. This enables the X-MET software to establish channel limits or windows for each element for spectral deconvolution. The X-MET software automatically chooses the K_α or L_α peak to set the window for the analytes. However, it is possible for the operator to manually choose a different window or peak. For example, in this demonstration, the L_β peak for lead was chosen because it is almost as intense as the L_α

lead peak but is not influenced by the interference from arsenic. The copper and zinc windows also were manually edited to reduce the amount of spectral overlap.

The next step in the empirical calibration is to enter the analyte concentrations for each SSCS into the X-MET software. After this is completed, each SSCS was analyzed using a count time of 200 live-time seconds. At the ASARCO site, 20 SSCSs were used and 14 SSCSs were used at the RV Hopkins site for the empirical calibration. The SSCSs samples were collected during the predemonstration and sent to Metorex. The analyte concentrations in these samples had been determined by EPA SW-846 Methods 3050A/6010A.

The final step in the empirical calibration is to develop the calibration curves for each analyte of interest. The calibration equation is developed using a multivariable, linear least squares fit regression analysis. After the regression terms to be used in the equation is defined, the software is designed to provide a mathematical equation to calculate the analyte concentration in an unknown sample. The equation is then stored in memory and used to quantitate an unknown sample when needed.

Once the calibration is complete, it is permanently stored as a method in the software. To analyze samples, the operator opens the stored method with the correct calibration and sets the measurement time. For this demonstration, it was suggested by Metorex to use shorter count times when analyzing the *in situ* samples and longer count times when analyzing the intrusive samples. This reflected the developer's opinion that increased sample preparation should be mirrored by increased analyzer accuracy or precision. The count times used at the ASARCO site were 30 seconds for *in situ* samples, 60 seconds for intrusive-unprepared samples, and 120 seconds for intrusive-prepared samples. The count times used at the RV Hopkins site were 60 seconds for the *in situ* samples, 120 seconds for the intrusive-unprepared samples, and 180 seconds for the intrusive-prepared samples. The count times were longer at the RV Hopkins site because chromium was an analyte of interest at RV Hopkins and Metorex was concerned that longer count times would be necessary to quantitate for chromium. The analysis is started by depressing the trigger in the handle of the SAPS probe. When the analysis is complete the X-MET 920-MP beeps and the analytical results are displayed on the screen.

The third step in the operation of the X-MET 920-MP is managing the data. All the analytical data files are automatically saved under a file name specified in the method. The raw spectrum of each analysis can also be saved individually. The operator noted it would be helpful if the software had a function to automatically save the raw spectra as it did the data files so that it would be unnecessary to manually save each one. All results files and spectra files were backed up on a floppy disk and were printed from the DOS editor.

Background of the Technology Operators

Metorex made no recommendation as to the qualifications and background necessary to operate its instrument, leaving the decision to the discretion of the client. Two PRC employees operated the X-MET 920-MP during the demonstration. Both operators have received all required Occupational Safety and Health Administration training to work on sites containing hazardous wastes.

The primary operator during this demonstration was employed as a geologist by PRC for 4 years. He earned a bachelor's degree in geology in 1982 and a master's degree in geology in 1987. While earning his masters, he worked for an environmental testing laboratory. He has extensive experience in field analysis; he worked on the Field Investigative Team contract for 3 years as a geologist and field analyst and has managed several comprehensive site investigations. At the conclusion of the field work at the RV Hopkins site, some intrusive samples from both sites remained for analysis, but the original operator was not able to finish the samples. At this time, the original operator trained another person to finish the analyses.

Approximately 80 of the intrusive-prepared samples from the ASARCO site and 50 of the intrusive-prepared samples from the RV Hopkins site were analyzed by the second operator. The replacement operator had been employed by PRC for 6 months as an environmental scientist prior to the demonstration. Prior to coming to PRC, he worked for 2.5 years in hazardous waste management for a Kansas municipality. He served as a nuclear, biological, and chemical (NBC) defense specialist in the United States Marine Corps for 6 years, specializing in monitoring and surveying of NBC attacks using field screening technologies, such as mobile mass spectrometers. He earned a bachelor's degree in 1992.

Training

The original operator received 2 days of training by the developer at the start of the demonstration. Approximately one-half day of training was dedicated to the theoretical background of XRF technology and the other one-and-a-half days were spent on specific operation and hands-on training for the X-MET 920-MP. The hands-on training covered performing the test measurements using the pure element samples, empirical calibrations and 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 and stated that the PRC operator received more one-on-one instruction in 2 days than typical students in a standard 3-day course. Metorex tailors its training course to match the level of operator 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 downloading the spectra and results files to floppy disk and printing out the data. They departed the next morning. Another Metorex representative also came to the RV Hopkins site for 2 days prior to the initiation of the field work to assist the operator with the empirical calibration of the X-MET 920-MP.

Reliability

More than 1,800 individual measurements were collected with the X-MET 920-MP Analyzer using the SAPS probe. This included the measurement of soil samples using the four sample preparation techniques, 10 replicate measurements on 48 samples for a precision assessment, and the measurement of QC samples such as blanks, check samples, PE samples, and SRMs. During the demonstration, there were frequent light to moderate rains while the analyzers were performing the *in situ* measurements. The temperatures fluctuated between 5 and 16 °C at the ASARCO site and 6 and 22 °C at the RV Hopkins site. Despite the less than ideal weather conditions, there were no mechanical problems experienced with the X-MET analyzer during the demonstration. The only maintenance necessary was to wipe the detector window with a paper towel between readings during the *in situ* analyses and occasionally to wipe dust off the window during the intrusive analyses. The operator did note that most commercially available and economically priced laptop computers are not weatherproof or ruggedized. The computer was covered with a plastic sheet while conducting the *in situ* measurements. Other than the inconvenience, this procedure proved effective and the computers operated without problems.

The primary problems encountered with the X-MET 920-MP were with battery lifetimes and operator familiarity with the software. The operator found that the batteries in the XPCS unit provided 7 to 8 hours of power which was adequate for *in situ* analyses. The limiting factor, however, was the battery life of the laptop computers. The operator was only able to get about 1.5 to 2 hours of power from the computer batteries, despite using three different computers from two manufacturers and obtaining three new batteries and a charger for one computer. The need to constantly change computers resulted in significant downtime. To keep pace with the demonstration schedule, a 35-pound portable generator was used at both sites to supply power to the laptop computer. The use of the generator decreased throughput and ease of use of the X-MET 920-MP. The limited PC battery lifetime and nonweatherproof design of the laptop computers diminished the effectiveness of the X-MET 920-MP as an *in situ* FPXRF analyzer.

The operator noted that several days were required to become familiar with the various menu options. The operator did feel that some aspects of the software were not self evident and were somewhat awkward.

A series of minor software problems was encountered by the operator. For example, every time batteries were switched, the operator was locked out of the software. After having this problem two or three times, the solution was to turn off the entire system and reinitialize the XPCS and the computer when changing the battery.

It was also discovered that attempting to use the same data file after reinitialization caused the file to be corrupted. The operator attempted to solve the problem with the assistance of Metorex but without success. A large portion of the data collected on the third day of analysis was never retrieved. The operator failed to back up this data from the hard drive. The rental laptop computer was returned before another attempt was made to retrieve the data. Ultimately, 78 data points from the *in situ*-unprepared sample measurements were lost. The operator learned that a new file had to be created for the data each time a lockout occurred to avoid corrupting the original data file.

Viewing and saving spectra also proved to be a problem at the ASARCO site. The operator was locked out of the software several times while attempting to save a spectrum. He had to reboot the computer and start a new data file when this occurred. This problem often occurred when the operator tried to type while the software was changing prompt messages.

On the final day of field analysis at the ASARCO site while analyzing the intrusive-prepared samples, the operator encountered an error message while saving data. The operator was concerned about losing data, so he copied all the spectra onto a floppy disk. He failed to back up the data files on floppy disk and did not print out a hard copy of this data. This rental laptop computer was sent back to the developer before the data was retrieved from the hard drive. This caused the loss of 13 data points from the intrusive-prepared sample measurements at the ASARCO site.

Health and Safety

Exposure to radiation from the excitation source was the largest health and safety consideration while using the analyzer. Radiation was monitored with a gamma ray detection 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. The permissible occupational exposure in Kansas 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), a maximum radiation value of 0.60 mrem/hr at the probe face was obtained with the Cd^{109} source exposed. The radiation values dropped off to 0.03 to 0.04 mrem/hr at the probe handle when the Cd^{109} source was exposed. Metorex designed the SAPS probe primarily for *in situ* measurements.

Two different protective covers were used on the SAPS probe with the probe pointing upward to conduct intrusive measurements. The SAPS was not designed to perform analysis of samples in a cup; however, for this demonstration the developer was interested in evaluating the SAPS in the intrusive mode. The first protective cover simply sat over the probe window but did not attach to the probe. With the protective cover in place and a sample cup over the window, radiation values of 2.0 to 3.0 mrem/hr were obtained at the sides of the cover with the source exposed. The radiation values dropped off to 0.10 mrem/hr at 1 foot from the probe and at the handle on the probe. Metorex sent the operator a second protective cover that was designed to attach to the SAPS probe. However, this protective cover was open on one side so that a sample cup could be placed over the probe window. With the source exposed, radiation values of 60 - 80 mrem/hr were obtained 1 inch from the open side of the protective cover. One foot away from the open side of the protective cover, radiation values of 2.0 to 4.0 mrem/hr were recorded. At the probe handle, 0.20 - 0.25 mrem/hr was measured with this protective cover. Based on these radiation values, the operator chose to use the first protective cover. At the RV Hopkins site, the operator placed a cardboard box lined with lead foil around the probe to offer additional protection. Radiation values of 0.016 to 0.018 mrem/hr were recorded behind the lead foil with the Cd^{109} exposed.

Cost

At the time of the demonstration, the X-MET 920-MP and SAPS probe with the Cd^{109} radioisotope source cost $36,325. This includes the X-MET software, eight pure element standards, and 3 days of training for two people at Metorex. Travel and accommodation costs for the training are not included. The purchase price is $3,000 less if the user supplies the laptop computer. The cost of the miniportable XPCS unit is $14,745. This price does not include the computer, probe, or source. Periodic maintenance includes replacement and disposal of the 10 mCi Cd^{109} source every 2 to 3 years at a cost of $3,725 with an additional $500 source disposal fee.

The X-MET 920-MP can be leased from Metorex. There is a 1-month minimum rental. The cost is 10 percent of the purchase price per month, and all shipping costs. Metorex mandates that the users be trained on the X-MET 920-MP. Users have a choice of training options. A 3-day class is offered at Metorex’s facility for $685 per person plus travel and lodging expenses. On-site training classes are also available.

The primary cost benefit of field analysis is the quick access to analytical data. This allows the process dependent on the analytical results to move efficiently onto 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 lowers the per sample cost. This is a key advantage that field analysis has over a conventional laboratory. Furthermore, more samples are usually taken for field analysis since questions raised in the preliminary findings may be resolved completely without the need to return for another sample collection event.

A representative list of costs associated with the X-MET 920-MP is presented in Table 4-2. Also included in this table is the measured throughput and the per sample charge of the reference laboratory. Given the special requirements of this demonstration, it was not considered reasonable to report a per sample cost for the field analysis. However, some estimates can be derived from the data provided in this table.

Table 4-2. Instrument and Field Operation Costs

Item	Amount	
X-MET 920-MP	$ 36,325	Purchase Price
	3,650	Per Month Lease
Replacement Source	3,725	For Cd^{109}
Operator Training (Vendor Provided)	685	—
Radiation Safety License (State of Kansas)	500	—
Field Operation Costs		
Supplies and Consumables (Sample cups, window film, sieves, standards)	300 - 500	(Varies, depending on sample load)
Field Chemist (Labor Charge)	100 - 150	Per day
Per diem	80 - 120	Per day
Travel	200 - 500	Per traveler
Sample Throughput	8 - 14	Samples per hour
Cost of Reference Laboratory Analysis	150	Per sample

Performance Factors

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

Detection Limits

MDLs, using SW-846 protocols, were determined by collecting 10 replicate measurements on site-specific soil samples with metals concentrations 2 to 5 times the expected values. These data were obtained during the precision evaluation. 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. The precision-based MDLs were calculated for soil samples that had been homogenized, dried, ground, and the intrusive-prepared level placed in a sample cup. The precision-based MDLs for the X-MET 920-MP are shown in Table 4-3. The precision-based MDLs for all analytes except chromium were obtained using a 120-second count time for the Cd^{109} source. The chromium MDL was determined using a 180-second count time. The developer did not provide specific MDLs for the analytes in Table 4-3; however, it did indicate the MDLs should be similar to those for the X-MET 920-P with the Si(Li) detector.

Another method of determining MDLs involved the direct comparison of the FPXRF and reference data. When these sets of data are plotted against each other, the results were linear. As the plotted line approached zero, there was a point at which the slope of the curve changes. Figure 4-2 illustrates this effect for arsenic. This point was determined by observation and was somewhat subjective; however, an analysis showed that even a 25 percent error in identifying this point resulted in only a 10 percent change in MDL. By determining the values of this FPXRF concentration and multiplying it by two standard deviations around the mean, it was possible to determine a field or performance-based MDL for the analyzer. This result is also shown in Figure 4-3.

Table 4-3. Method Detection Limits

Analyte	Precision-based MDL (mg/kg)	Field-based MDL (mg/kg)
Arsenic	50	175
Barium	330	Not determined
Chromium	115	130
Copper	50	Not determined
Lead	30	105
Nickel	Not determined	Not determined
Zinc	30	210

Note: mg/kg Milligrams per kilogram.

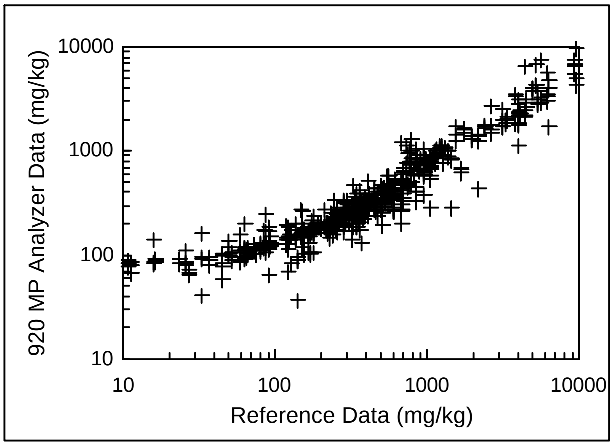

Figure 4-2. Critical Zone for the Determination of a Field-based Method Detection Limit for Copper: Between 100 and 200 mg/kg for the reference data, the linear relationship between the two data sets change. This point of change was used to determine the field-based MDL.

As would be expected, the precision-based MDLs generally improved when the absorption energy of the analyte was close to the excitation energy of the Cd^{109} source. Detection limits could not be determined for nickel because the nickel concentrations were too low. A field-based MDL for barium and copper could not be determined because the regression plots were linear through the y-intercept at zero. The field-based MDLs for the remaining three analytes were 3 to 7 times higher than the precision-based MDLs. Detection limits for chromium were consistent using either technique.

Throughput

Throughput for the X-MET 920-MP varied during the demonstration since differing count times were used. The count times varied from 30 seconds at the ASARCO site to 180 seconds at the RV Hopkins site. The following discussion includes operator time only and takes into consideration the time spent for daily set up, QC checks, and for data handling. The throughput does not take into consideration sample preparation time.

At the ASARCO site, 30-second count times were used for all *in situ*-unprepared and *in situ*-prepared analyses. On the first day of *in situ*-unprepared analysis, the operator was only able to achieve a throughput of three samples per hour due to problems with computer batteries and software lockouts. By the third day of *in situ*-unprepared analyses, a throughput of eight samples per hour was achieved. When the work moved indoors at the ASARCO site and the X-MET 920-MP was run from alternating current power, the operator was able to analyze 14 *in situ*-prepared samples per hour.

At the RV Hopkins site, the count times were increased to 60 seconds for the *in situ*-unprepared and prepared samples for better chromium resolution and quantitation. Despite the increase in count time, throughput for the *in situ*-unprepared and *in situ*-prepared samples was 8 and 14 samples per hour, respectively.

Using a 60-second count time for the intrusive-unprepared samples at ASARCO, the operator analyzed 176 samples in a 12.5-hour day for a throughput of 14 samples per hour. This was the maximum number of samples analyzed in one day. When the count time was increased to 120 seconds and 180 seconds for the intrusive-prepared samples at the two sites, the throughput dropped to 11 and 8 samples per hour, respectively. Metorex claimed that 200 - 400 samples could be analyzed in an 8- to 10-hour day based on 10- to 100-second count times. The throughput during this demonstration averaged between 70 and 150 samples per day, depending on count times.

Drift

Drift is a measurement of an analyzer's variability in quantitating a known amount of a standard over time. Drift was evaluated by reviewing results from the periodic analysis of a mid-level SSCS from each site.

Metorex recommended that the operator run the SSCS as a check sample once per hour during the demonstration. The ASARCO SSCS check sample was analyzed 57 times and the RV Hopkins SSCS check sample 42 times during the demonstration. These data were reduced to RSDs for the target analytes and the percent drift from the mean recovery of the true value. The percent drift from the mean recovery was averaged for each day and is shown in Figure 4-3 for the target analytes reported by the X-MET 920-MP.

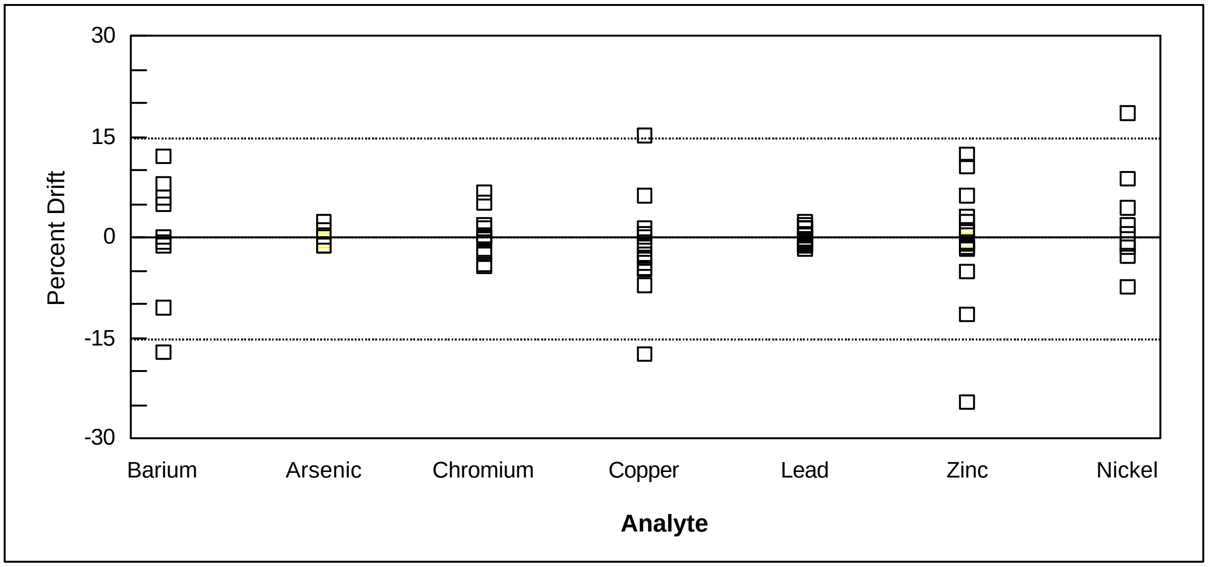

Figure 4-3. Drift Summary: This graph shows the general drift of the analyzer's results in measuring a check sample. Each point represents a different day's analysis of the same sample. The daily fluctuations exhibited for each analyte are a direct representation of drift.

The X-MET 920-MP displayed the least drift for arsenic and lead with RSD values less than 3.6 percent for both SSCS check samples. Greater drift was shown for barium, chromium, copper, nickel, and zinc. The drift for barium, chromium, and nickel was only assessed using the RV Hopkins SSCS check sample. Barium and nickel showed the greatest drift as supported by RSD values of 19.5 and 16.2, respectively. Chromium had an RSD value of 6.7. Copper and zinc had RSD values of 9.1 and 5.0, respectively, for the ASARCO SSCS check sample, but RSD values of 23.5 and 24.9, respectively, for the RV Hopkins SSCS check sample. This may indicate a difference in the empirical calibration between the two sites, and specifically, in defining the overlap tables. All data points for barium, copper, and zinc in Figure 4-3 outside of 10 percent drift were for the RV Hopkins SSCS check sample.

Intramethod Assessment

Intramethod measures of each analyzer's performance included results on analyzer blanks, completeness, precision, and accuracy. The following narrative discusses these characteristics.

Blanks

Analyzer blanks for the X-MET 920-MP consisted of pure lithium carbonate placed directly in a sample cup and lithium carbonate that had been put through 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. Six blanks were analyzed at the ASARCO site using the ASARCO empirical calibration. Four blanks were analyzed at the RV Hopkins site using the RV Hopkins empirical calibration.

The X-MET 920-MP reported values for arsenic and copper above the precision-based MDLs for the ASARCO blanks and values for chromium, lead, and zinc above the precision-based MDLs for the RV Hopkins blanks. None of the other analytes were detected in the blanks. An examination of the data showed the results for the analytes found above their respective MDLs were an artifact of the regression equations used for calibration. It was apparent that the regression equations developed using SSCSs did not apply to a much different matrix such as lithium carbonate. The blank data for the X-MET 920-MP demonstrated that a "clean" matrix matching the SSCSs should be used for a blank. It was therefore concluded that the lithium carbonate blanks did not show a contamination problem.

Completeness

The X-MET 920-MP produced data for 1,168 out of the 1,260 samples for a completeness of 92.7 percent, slightly below the demonstration objective of 95 percent. The 92 samples for which no data were obtained were all from the ASARCO site. There were 78 *in situ*-unprepared, one *in situ*-prepared, and 13 intrusive-prepared samples for which no data were obtained. The loss of data for these samples was a combination of software problems, operator oversight, and a failure of the lead chemist to realize the data set was not complete prior to returning the rental computer to the manufacturer. It should be noted that none of the lost data was caused by mechanical or electronic malfunctions of the analyzer.

Precision

Precision was expressed in terms of the percent RSD between replicate measurements. The percent RSD is defined as the SD divided by the mean concentration times 100. The precision data for the target analytes detectable by the analyzer are shown in Table 4-4. The precision data reflected in the range of 5 to 10 times the MDL reflects the precision generally referred to in analytical methods such as SW-846 and represents general method precision.

The analyzer performed 10 replicate measurements on 12 soil samples that had analyte concentrations ranging from less than 50 mg/kg to tens of thousands of milligrams per kilogram. Each of the 12 soil samples underwent 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.

In this demonstration, the analyzer's precision or RSD for a given analyte had to be less than or equal to 20 percent to be considered quantitative screening level data and less than or equal to 10 percent to be considered definitive level data. The analyzer's precision data, reflected by its precision data in the 5 to 10 times MDL range, one measure in the overall data quality evaluation, were all below the 10 percent RSD required for definitive level data quality classification. Chromium and nickel did not have sufficient data to allow data quality conclusions based on precision. The lower precision for nickel in the 50 - 500 mg/kg range may be an artifact of the low concentrations of nickel in the soil samples and the influence of iron interference in the samples.

Table 4-4. Precision Summary

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	1.91 (16)	3.08 (20)	1.32 (8)	0.73 (4)
Barium	5.91 (4)	ND	9.52 (8)	10.08 (8)
Chromium	ND	8.19 (16)	ND	ND
Copper	7.92 (12)	6.81 (16)	15.93 (4)	4.19 (8)
Lead	7.60 (16)	4.16 (16)	6.85 (12)	5.43 (12)
Nickel	ND	20.92 (4)	ND	ND
Zinc	2.10 (20)	3.06 (36)	1.00 (4)	ND

Notes:

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

The precision data in Table 4-4 shows there was little effect of concentration on the precision. The precision samples were purposely chosen to span a large concentration range to test the effect of analyte concentration on precision. It was expected that as the analyte concentration increased, the precision would increase. However, the precision was so good even for the low concentration samples, that little or no increase in precision was seen as the concentrations increased. With the exception of nickel, all mean RSD values were less than 10 percent for samples containing less than 500 mg/kg of an analyte. The good precision for this instrument can be attributed to the high efficiency of the gas proportional detector and to the manner in which the spectrum is deconvoluted by the software.

Table 4-5 shows the precision data as affected by sample preparation and count time. It was not possible to separate the two factors as both were changing throughout the demonstration. For the analytes barium, chromium, nickel, and zinc, there was no consistent improvement in precision with increased sample preparation and count time. There was a slight increase in arsenic precision with increased sample preparation and count time. The most dramatic increase in precision was seen for copper and lead, especially for the intrusive samples analyzed at longer count times. A paired t-test was used to determine if the average RSD values were significantly different for copper and lead for the different sample preparation techniques and count times. A significance level of 0.05 was used to test for a difference in mean RSD values. For copper, the precision for the intrusive samples was significantly different (better) than for the *in situ* samples. There was no statistical difference for the copper precision between the intrusive-prepared and intrusive-unprepared samples. The precision for lead was found to be statistically equal for the *in situ*-unprepared and *in situ*-prepared samples. The precision for lead for the intrusive samples was found to be statistically significantly better than for the *in situ* samples.

Table 4-5. Precision as Affected by Sample Preparation and Count Times

Analyte	Mean % RSD Values			
	In Situ-Unprepared (30 Seconds)	In Situ-Prepared (30 Seconds)	Intrusive-Unprepared (60 Seconds)	Intrusive-Prepared (120 Seconds)
Arsenic	3.45	3.11	2.29	1.69
Barium[a]	8.17	11.15	6.77	12.16
Chromium[a]	6.77	8.04	9.06	8.87
Copper	13.91	14.62	11.51	7.44
Lead	11.74	7.32	1.84	1.81
Nickel[a]	17.45	24.65	ND	20.67
Zinc	2.24	2.72	2.57	3.59

Notes: [a] RSD values obtained for RV Hopkins precision samples only. Count times were 60 seconds for all in situ samples, 120 seconds for intrusive-unprepared, and 180 seconds for intrusive-prepared samples.

ND No data.

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. Originally, it was intended that accuracy would be assessed for the X-MET 920-MP using six site-specific PE samples and 14 SRMs. However, since the X-MET 920-MP was calibrated empirically using SSCSs whose true values were determined using EPA SW-846 Methods 3050A/6010A, the analysis of the SRMs would not be a true assessment of accuracy since the matrix (particle size, contaminant analytes, and interferant analytes) of the SRMs did not match the matrix on which the calibration was based. Also, the true values in the SRMs were determined using “total metals” methods unlike the SSCSs, which had analyte concentrations determined using EPA SW-846 Methods 3050A/6010A, a partial digestion method.

The site-specific PE samples consisted of three samples 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. Because the PE samples matched the matrix at each site and had true values determined by a method closely matching FPXRF, these PE samples would serve as a good accuracy check. However, because the X-MET 920-MP used Methods 3050A/6010A-based data for calibration, the analytical method used to determine the PE sample's concentrations did not match the analytical method used to calibrate the analyzer. Therefore, the site-specific PEs could not be used as an ideal assessment of accuracy.

Ultimately, it was concluded that none of the site-specific PEs or SRMs could be used as a true accuracy check for this analyzer. However, during this demonstration, the site-specific PEs were submitted to the reference laboratory for analysis by EPA SW-846 Methods 3050A/6010A. It was decided that the reference laboratory data would be used as the true value for the site-specific PEs. However, only one analysis was performed on each site-specific PE sample and that there is a degree of error associated with any single measurement. The only alternative would have been to send the site-specific PE samples to multiple laboratories for analysis by EPA SW-846 Methods 3050A/6010A and then determine the mean analyte concentrations and the standard error around those means. This was not considered prior to the demonstration and, therefore, was considered an option.

Although not a perfect assessment of accuracy, the reference laboratory data was used as the true values for the site-specific PE samples. The PEs 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. Table 4-6 summarizes the accuracy data for the target analytes for the X-MET 920-MP. Figure 4-4 shows the true value, the measured value, and percent recovery for the individual PEs. No figure is presented for the nickel data because there were only two detects.

Table 4-6. Accuracy Summary for Site-Specific PE Results

Analyte	n	Percent Within Acceptance Range	Mean Percent Recovery	Range of Percent Recovery	SD of Percent Recovery	Concentration Range (mg/kg)
Site-Specific Performance Evaluation Samples						
Arsenic	4	0	57	44 - 68	12.3	105 - 20,885
Barium	4	25	151	88 - 226	59.5	195 - 3,817
Chromium	3	0	46	22 - 69	23.4	139 - 1,761
Copper	6	50	143	113 - 209	38.0	69 - 5,869
Nickel	3	0	469	369 - 569	141	129 - 321
Lead	6	0	64	52 - 79	11.3	242 - 13,210
Zinc	6	50	74	38 - 102	21.3	116 - 3,497

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

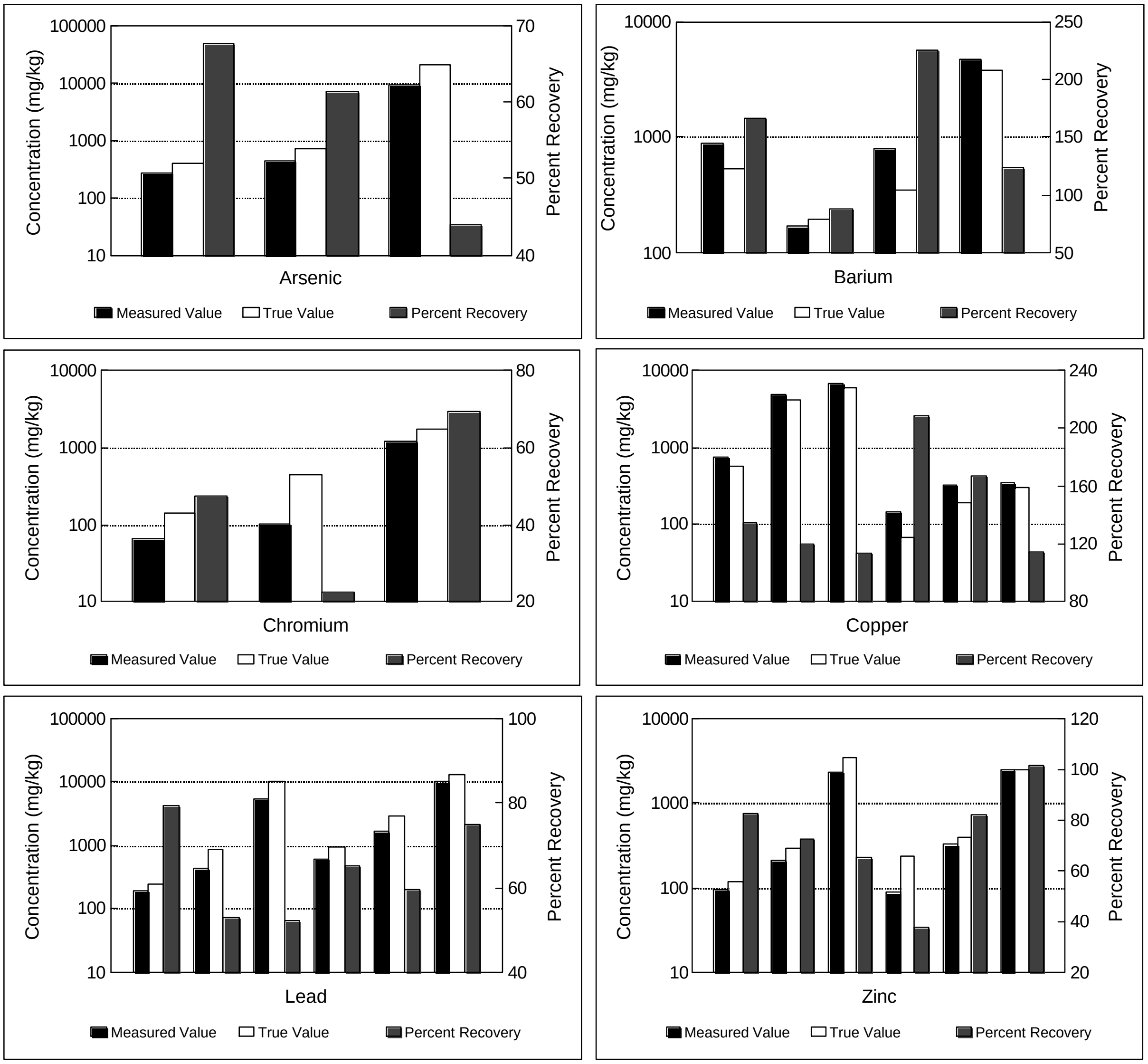

Figure 4-4. Site-specific PE Sample Results: These graphs illustrate the relationship between the analyzer's data (measured values) and the true values for the site-specific PE samples. The gray bars represent the percent recovery for the analyzers. Each set of three bars (black, white, and gray) represents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

Based on the 80 - 120 percent recovery acceptance criteria, the analyzer produced 7 out of 32 results or 21.9 percent within the acceptance range for all analytes in the six PEs. The X-MET 920-MP produced 0 percent of the results within the acceptance range for arsenic, chromium, nickel, and lead; 25 percent for barium; and 50 percent for copper and zinc. The mean percent recovery and range of percent recoveries demonstrate that the X-MET 920-MP results were biased high for barium, copper, and nickel. The X-MET 920-MP nickel results showed the most disparity from the true values, possibly due to interferences from iron in these samples. The copper and zinc X-MET 920-MP results showed the closest comparison to the true values. The X-MET 920-MP underestimated arsenic, chromium, and lead. In comparing the results of the check sample analysis discussed earlier in the drift assessment section, it is interesting to note the mean percent recoveries for the target analytes in the check sample were similar to the mean percent recoveries listed in Table 4-6. This indicates that the low or high biases for specific analytes were consistent throughout the demonstration.

Intermethod Assessment

The comparison of the X-MET 920-MP results to those of the reference methods was performed using the statistical methods detailed in Section 2. The purpose of this evaluation was to determine the comparability between data produced by the analyzer 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 or less), the data met the definitive level criteria. If the data did not meet the definitive level criteria, but could be mathematically corrected to be equivalent to the reference data, 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 analyzer was configured to report concentrations for all of the target analytes except iron, cadmium, and antimony. Due to a variety of technical problems, some analyzer data were considered not valid and subsequently dropped from the comparability assessment.

The comparability assessment begins with a regression analysis of the entire data set for the primary analytes, arsenic and lead which had r^2 values above 0.85. Barium, chromium, copper, and zinc had r^2 values ranging from 0.69 to 0.55. Based on the slope of the regressions, the analyzer tended to overestimate barium and underestimate arsenic, copper, zinc, and chromium concentrations relative to the reference methods. (The slope values were determined with the FPXRF data plotted on the y-axis and the reference data plotted on the x-axis.) The r^2 value for nickel, 0.32, showed poor comparability; however, the number of data points was limited.

The next step in the data evaluation involved assessing the potential impact of the site, soil type, and sample preparation variables on the regression analysis (Table 4-7). 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 type variable was also limited. Copper and arsenic did not exhibit a wide concentration distribution in the clay soil, 10 - 250 mg/kg and 4 - 28 mg/kg, respectively. Barium, chromium, and nickel did not exhibit a wide concentration distribution in either the sand or loam soil, 30 - 300 mg/kg (95 percent of data), 10 - 130 mg/kg, and 8 - 550 mg/kg, respectively. Based on this evaluation, there was no apparent impact of either the site or soil type variables on the regression.

The sample preparation variable was confounded by the increased count times used for each successive sample preparation step. For this reason, any increase in comparability cannot be wholly attributed to either sample preparation or count times. The precision study indicated that only copper and lead data showed increases in precision with increased count times. The lead exhibited a 10-fold increase in precision over the four sample preparation-increased count time steps, and copper exhibited a twofold increase in precision over the same intervals. Based on this, for all target analytes, except lead and copper, any increase in comparability with increased sample preparation was attributed to sample preparation. Changes in lead and copper comparability remain confounded, including both sample preparation and count time effects.

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

Arsenic					Variable	Barium				
n	r^2	Std. Err.	Y-Int.	Slope[b]		n	r^2	Std. Err.	Y-Int.	Slope[b]
746	0.946	0.14	0.49	0.80	All Data	113	0.620	0.47	0.37	1.08
746	0.946	0.14	0.49	0.80	ASARCO Site	ND	ND	ND	ND	ND
ND	ND	ND	ND	ND	RV Hopkins Site	113	0.550	0.57	0.15	1.15
368	0.926	0.17	0.63	0.74	Sand Soil	ND	ND	ND	ND	ND
383	0.957	0.11	0.42	0.82	Loam Soil	ND	ND	ND	ND	ND
ND	ND	ND	ND	ND	Clay Soil	116	0.550	0.57	0.15	1.15
135	0.898	0.21	0.62	0.75	In Situ-Unprepared	ND	ND	ND	ND	ND
208	0.969	0.10	0.42	0.82	In Situ-Prepared	ND	ND	ND	ND	ND
209	0.964	0.11	0.44	0.81	Intrusive-Unprepared	41	0.580	0.51	0.13	1.14
197	0.937	0.15	0.47	0.80	Intrusive-Prepared	74	0.660	0.46	0.20	1.16

Chromium					Variable	Copper				
n	r^2	Std. Err.	Y-Int.	Slope[b]		n	r^2	Std. Err.	Y-Int.	Slope[b]
928	0.680	0.23	0.20	0.82	All Data	747	0.690	0.46	0.17	0.88
709	0.006	0.06	1.27	0.04	ASARCO Site	618	0.730	0.43	-0.38	1.06
219	0.650	0.30	0.61	0.69	RV Hopkins Site	129	0.390	0.35	0.59	0.81
348	0.001	0.02	1.30	0.01	Sand Soil	321	0.761	0.22	0.19	0.82
385	0.003	0.20	1.30	0.08	Loam Soil	288	0.660	0.58	-0.86	1.12
187	0.650	0.30	0.61	0.69	Clay Soil	129	0.390	0.35	0.59	0.81
124	0.180	0.04	1.17	0.11	In Situ-Unprepared	106	0.550	0.34	0.20	0.82
214	0.260	0.04	1.09	0.16	In Situ-Prepared	142	0.700	0.36	0.24	0.76
298	0.793	0.18	0.18	0.80	Intrusive-Unprepared	233	0.734	0.36	0.40	0.74
282	0.639	0.29	0.43	0.76	Intrusive-Prepared	184	0.716	0.29	0.66	0.73

Lead					Variable	Nickel				
n	r^2	Std. Err.	Y-Int.	Slope[b]		n	r^2	Std. Err.	Y-Int.	Slope[b]
849	0.878	0.25	-0.06	1.00	All Data	130	0.320	0.47	0.17	0.79
700	0.885	0.24	-0.19	1.05	ASARCO Site	ND	ND	ND	ND	ND
149	0.760	0.27	0.46	0.82	RV Hopkins Site	130	0.320	0.47	0.17	0.79
340	0.910	0.21	-0.26	1.06	Sand Soil	ND	ND	ND	ND	ND
363	0.841	0.26	-0.04	1.01	Loam Soil	ND	ND	ND	ND	ND
146	0.760	0.27	0.46	0.82	Clay Soil	130	0.320	0.47	0.17	0.79
127	0.850	0.30	-0.12	1.01	In Situ-Unprepared	ND	ND	ND	ND	ND
193	0.920	0.20	-0.23	1.07	In Situ-Prepared	ND	ND	ND	ND	ND
275	0.840	0.27	0.04	0.96	Intrusive-Unprepared	47	0.250	0.42	0.54	0.58
251	0.890	0.23	-0.05	1.00	Intrusive-Prepared	79	0.470	0.33	0.24	0.77

Table 4-7. ***Continued***

Zinc					Variable
n	r^2	Std. Err.	Y-Int.	Slope[b]	
667	0.550	0.36	0.29	0.85	All Data
619	0.527	0.34	0.39	0.79	ASARCO Site
48	0.730	0.34	0.40	0.98	RV Hopkins Site
271	0.498	0.44	0.13	0.84	Sand Soil
354	0.537	0.25	0.71	0.70	Loam Soil
49	0.730	0.34	0.40	0.98	Clay Soil
101	0.480	0.39	0.10	0.87	In Situ-Unprepared
179	0.640	0.30	0.40	0.81	In Situ-Prepared
196	0.662	0.36	-0.09	0.99	Intrusive-Unprepared
192	0.417	0.39	0.56	0.78	Intrusive-Prepared

Notes:

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

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

n Number of data points.

Y-Int. Y-Intercept.

Std. Err. Standard error.

ND Analytes not present in significant quantities to provide meaningful regression. This data evaluation does not include copper data from ASARCO samples 102 to 201 for intrusive-prepared analyses and RV Hopkins data from both in situ sample preparation analyses.

The sample preparation variable exhibited the greatest impact on the regression analysis. Generally, the largest shift in the r^2 was exhibited between the *in situ*-unprepared and *in situ*-prepared analyses (Figure 4-5). Except for chromium, sample homogenization and increased count times accounted for between 80 and 100 percent of the total increase in the r^2 experienced across all sample preparation steps. Chromium exhibited its greatest increase in comparability between the last *in situ* and first intrusive analysis.

The impact of the soil texture and site variables was also assessed for each of the four sample preparation steps (Tables 4-8 and 4-9). This evaluation was conducted for lead and zinc only. These were the only primary analytes exhibiting relatively even concentration distribution between the site and soil variables. No clear effect on comparability was observed for these variables. A minor trend was observed for the site variable. Generally, the r^2 values for lead at the ASARCO site were higher, relative to the RV Hopkins site. This may have been due to the wider range of analyte concentrations found at the ASARCO site and not a site or soil related effect.

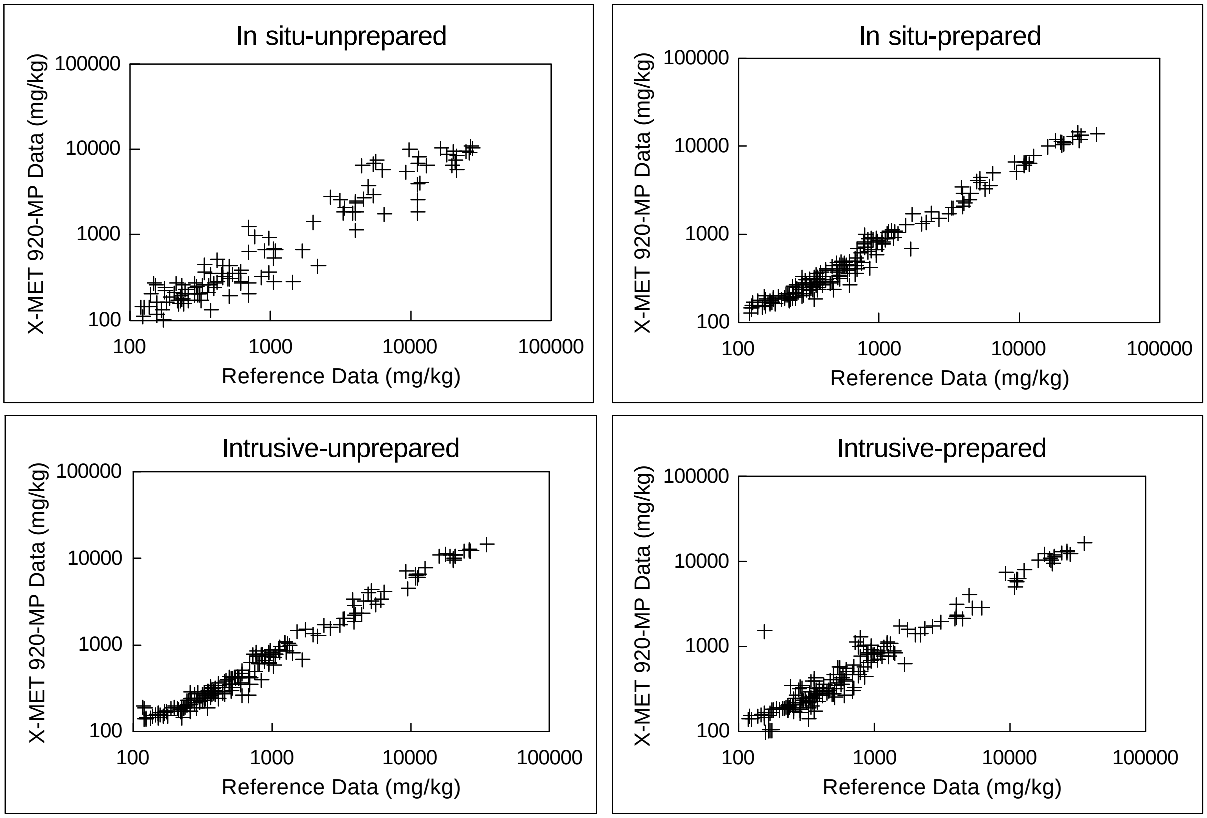

Figure 4-5. Sample Preparation Effects on Arsenic Results: These log-log plots illustrate the change in comparability with changes in sample preparation step.

Within the sample preparation steps, the effect of contaminant concentration was also examined. The 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 m/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 for each target analyte 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 at or below the analyzers MDLs. The analyzer's precision and accuracy are lowest in this concentration range. Generally, the r^2 values improved among the 100 - 1,000 mg/kg and greater than 1,000 mg/kg ranges. This data indicated that there was a concentration effect on comparability. This 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.

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

Arsenic						Barium				
n	r^2	Std. Err.	Y-Int.	Slope[b]		n	r^2	Std. Err.	Y-Int.	Slope[b]
In Situ-Unprepared					**Soil Texture**	**In Situ-Unprepared**				
78	0.904	0.20	0.67	0.72	Sand Soil	ND	ND	ND	ND	ND
57	0.892	0.21	0.54	0.78	Loam Soil	ND	ND	ND	ND	ND
ND	ND	ND	ND	ND	Clay Soil	44	0.574	0.48	0.46	1.05
In Situ-Prepared					**Soil Texture**	**In Situ-Prepared**				
95	0.946	0.15	0.56	0.77	Sand Soil	ND	ND	ND	ND	ND
112	0.981	0.07	0.39	0.84	Loam Soil	ND	ND	ND	ND	ND
ND	ND	ND	ND	ND	Clay Soil	29	0.702	0.29	0.77	0.91
Intrusive-Unprepared					**Soil Texture**	**Intrusive-Unprepared**				
97	0.928	0.17	0.61	0.75	Sand Soil	ND	ND	ND	ND	ND
114	0.973	0.08	0.42	0.82	Loam Soil	ND	ND	ND	ND	ND
ND	ND	ND	ND	ND	Clay Soil	39	0.482	0.49	0.65	0.97
Intrusive-Prepared					**Soil Texture**	**Intrusive-Prepared**				
98	0.926	0.17	0.65	0.75	Sand Soil	ND	ND	ND	ND	ND
101	0.942	0.13	0.30	0.86	Loam Soil	ND	ND	ND	ND	ND
ND	ND	ND	ND	ND	Clay Soil	74	0.660	0.46	0.20	1.16

Chromium						Copper				
n	r^2	Std. Err.	Y-Int.	Slope[b]		n	r^2	Std. Err.	Y-Int.	Slope[b]
In Situ-Unprepared					**Soil Texture**	**In Situ-Unprepared**				
69	0.108	0.03	1.19	0.09	Sand Soil	67	0.646	0.24	0.07	0.88
52	0.000	0.03	1.33	0.00	Loam Soil	38	0.523	0.44	-0.84	1.14
96	0.450	0.31	1.27	0.46	Clay Soil	68	0.236	0.53	0.80	0.80
In Situ-Prepared					**Soil Texture**	**In Situ-Prepared**				
96	0.000	0.02	1.31	0.00	Sand Soil	80	0.638	0.22	0.68	0.59
114	0.520	0.03	1.02	0.21	Loam Soil	61	0.715	0.45	-0.42	0.95
97	0.365	0.25	1.57	0.30	Clay Soil	68	0.084	0.40	1.55	0.33
Intrusive-Unprepared					**Soil Texture**	**Intrusive-Unprepared**				
98	0.012	0.02	1.33	-0.02	Sand Soil	93	0.895	0.17	0.13	0.83
113	0.361	0.03	1.11	0.15	Loam Soil	85	0.828	0.33	-0.84	1.09
93	0.633	0.30	0.64	0.66	Clay Soil	51	0.439	0.36	0.25	0.96
Intrusive-Prepared					**Soil Texture**	**Intrusive-Prepared**				
85	0.000	0.01	1.31	0.00	Sand Soil	81	0.929	0.13	-0.56	1.17
102	0.003	0.23	1.80	-0.07	Loam Soil	25	0.951	0.12	-0.92	1.26
94	0.673	0.31	0.59	0.71	Clay Soil	77	0.413	0.31	0.75	0.74

Table 4-8. ***Continued***

Lead						Nickel				
n	r^2	Std. Err.	Y-Int.	Slope[b]		n	r^2	Std. Err.	Y-Int.	Slope[b]
In Situ-Unprepared					**Soil Texture**	**In Situ-Unprepared**				
72	0.885	0.26	-0.17	1.02	Sand Soil	ND	ND	ND	ND	ND
53	0.840	0.32	-0.11	1.02	Loam Soil	ND	ND	ND	ND	ND
40	0.500	0.47	-0.63	1.12	Clay Soil	23	0.217	0.83	-2.31	2.04
In Situ-Prepared					**Soil Texture**	**In Situ-Prepared**				
86	0.941	0.17	-0.29	1.08	Sand Soil	ND	ND	ND	ND	ND
106	0.875	0.22	0.04	0.99	Loam Soil	ND	ND	ND	ND	ND
35	0.577	0.30	0.03	0.91	Clay Soil	33	0.325	0.68	-0.75	1.08
Intrusive-Unprepared					**Soil Texture**	**Intrusive-Unprepared**				
90	0.920	0.20	-0.36	1.10	Sand Soil	ND	ND	ND	ND	ND
108	0.880	0.22	-0.10	1.02	Loam Soil	ND	ND	ND	ND	ND
74	0.667	0.29	0.86	0.69	Clay Soil	45	0.248	0.41	0.50	0.60
Intrusive-Prepared					**Soil Texture**	**Intrusive-Prepared**				
86	0.938	0.17	-0.18	1.05	Sand Soil	ND	ND	ND	ND	ND
94	0.813	0.29	-0.05	1.04	Loam Soil	ND	ND	ND	ND	ND
72	0.856	0.24	0.07	0.94	Clay Soil	79	0.470	0.33	0.24	0.77

Zinc					
n	r^2	Std. Err.	Y-Int.	Slope[b]	
In Situ-Unprepared					**Soil Texture**
51	0.558	0.35	-0.11	0.86	Sand Soil
50	0.506	0.36	0.27	0.87	Loam Soil
91	0.620	0.19	1.09	0.54	Clay Soil
In Situ-Prepared					**Soil Texture**
75	0.664	0.40	0.03	0.91	Sand Soil
106	0.566	0.25	0.70	0.71	Loam Soil
92	0.607	0.20	1.08	0.52	Clay Soil
Intrusive-Unprepared					**Soil Texture**
71	0.641	0.46	-0.26	0.98	Sand Soil
107	0.630	0.25	0.36	0.82	Loam Soil
23	0.712	0.43	-0.54	1.31	Clay Soil
Intrusive-Prepared					**Soil Texture**
72	0.007	0.42	1.72	0.12	Sand Soil
91	0.452	0.17	1.37	0.42	Loam Soil
31	0.816	0.34	0.29	1.09	Clay Soil

Notes:

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

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

Y-Int. Y-Intercept.

Std. Err. Standard Error.

n Number of data points.

ND Analytes not present in significant quantities to provide meaningful regression. This data does not include copper data from ASARCO samples 102 to 201 for the intrusive-prepared analyses.

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

Arsenic						Barium				
n	r^2	Std. Err.	Y-Int.	Slope[b]		n	r^2	Std. Err.	Y-Int.	Slope[b]
In Situ-Unprepared					Site Name	In Situ-Unprepared				
135	0.898	0.21	0.62	0.75	ASARCO	ND	ND	ND	ND	ND
ND	ND	ND	ND	ND	RV Hopkins	44	0.574	0.48	0.46	1.05
In Situ-Prepared					Site Name	In Situ-Prepared				
208	0.969	0.10	0.42	0.82	ASARCO	ND	ND	ND	ND	ND
ND	ND	ND	ND	ND	RV Hopkins	29	0.702	0.29	0.77	0.91
Intrusive-Unprepared					Site Name	Intrusive-Unprepared				
209	0.964	0.11	0.44	0.81	ASARCO	ND	ND	ND	ND	ND
ND	ND	ND	ND	ND	RV Hopkins	39	0.482	0.49	0.65	0.97
Intrusive-Prepared					Site Name	Intrusive-Prepared				
197	0.937	0.15	0.47	0.80	ASARCO	ND	ND	ND	ND	ND
ND	ND	ND	ND	ND	RV Hopkins	74	0.660	0.46	0.20	1.16

Chromium						Copper				
n	r^2	Std. Err.	Y-Int.	Slope[b]		n	r^2	Std. Err.	Y-Int.	Slope[b]
In Situ-Unprepared					Site Name	In Situ-Unprepared				
121	0.069	0.03	1.22	0.07	ASARCO	104	0.621	0.31	0.10	0.86
96	0.450	0.31	1.27	0.46	RV Hopkins	68	0.236	0.53	0.80	0.80
In Situ-Prepared					Site Name	In Situ-Prepared				
209	0.227	0.03	1.14	0.12	ASARCO	140	0.738	0.33	0.20	0.77
97	0.365	0.25	1.57	0.30	RV Hopkins	68	0.084	0.40	1.55	0.33
Intrusive-Unprepared					Site Name	Intrusive-Unprepared				
211	0.156	0.03	1.18	0.09	ASARCO	180	0.822	0.29	-0.15	0.90
93	0.633	0.30	0.64	0.66	RV Hopkins	51	0.439	0.36	0.25	0.96
Intrusive-Prepared					Site Name	Intrusive-Prepared				
188	0.003	0.26	1.38	0.10	ASARCO	107	0.923	0.15	-0.64	1.19
94	0.673	0.31	0.59	0.71	RV Hopkins	77	0.413	0.31	0.75	0.74

Table 4-9. ***Continued***

Lead						Nickel				
n	r^2	Std. Err.	Y-Int.	Slope[b]		n	r^2	Std. Err.	Y-Int.	Slope[b]
In Situ-Unprepared					**Site Name**	**In Situ-Unprepared**				
126	0.859	0.29	-0.17	1.03	ASARCO	ND	ND	ND	ND	ND
40	0.500	0.47	-0.63	1.12	RV Hopkins	23	0.217	0.83	-2.31	2.04
In Situ-Prepared					**Site Name**	**In Situ-Prepared**				
193	0.915	0.20	-0.23	1.07	ASARCO	ND	ND	ND	ND	ND
35	0.577	0.30	0.03	0.91	RV Hopkins	34	0.346	0.78	-1.26	1.28
Intrusive-Unprepared					**Site Name**	**Intrusive-Unprepared**				
201	0.881	0.24	-0.25	1.06	ASARCO	ND	ND	ND	ND	ND
74	0.667	0.29	0.86	0.69	RV Hopkins	45	0.248	0.41	0.50	0.60
Intrusive-Prepared					**Site Name**	**Intrusive-Prepared**				
181	0.878	0.24	-0.17	1.06	ASARCO	ND	ND	ND	ND	ND
72	0.856	0.24	0.07	0.94	RV Hopkins	79	0.470	0.33	0.24	0.77

Zinc					
n	r^2	Std. Err.	Y-Int.	Slope[b]	
In Situ-Unprepared					**Site Name**
104	0.459	0.43	-0.01	0.91	ASARCO
91	0.620	0.19	1.09	0.54	RV Hopkins
In Situ-Prepared					**Site Name**
179	0.640	0.30	0.40	0.81	ASARCO
92	0.607	0.20	1.08	0.52	RV Hopkins
Intrusive-Unprepared					**Site Name**
174	0.665	0.32	0.11	0.90	ASARCO
23	0.712	0.43	-0.54	1.31	RV Hopkins
Intrusive-Prepared					**Site Name**
161	0.229	0.30	1.21	0.46	ASARCO
31	0.816	0.34	0.29	1.09	RV Hopkins

Notes:

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

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

Y-Int. Y-Intercept.

Std. Err. Standard Error.

n Number of data points.

ND Analytes not present in significant quantities to provide meaningful regression. This data does not include copper data from ASARCO samples 102 to 201 for the intrusive-prepared analyses.

Another way to examine the comparability between the two methods involves measuring the average relative bias and accuracy between the FPXRF data and the reference. These measurements were made by using the raw FPXRF and reference data. The average relative bias indicates the average factor by which the two data sets differ. Concentration effects can affect bias. For example, it is possible for an analyzer to greatly underestimate low concentrations but greatly overestimate high concentrations and have a relative bias of zero. To eliminate this concentration effect, the data can be corrected by a regression approach (see Section 5). Otherwise, only narrow concentration ranges can be analyzed, or average relative accuracy can be analyzed. The average relative accuracy is the average factor by which each individual analyzer measurement differs from the corresponding reference measurement.

A final decision regarding the assignment of data quality levels involves an assessment of both the r^2 and precision results. Using the criteria presented in Table 2-2, a summary of the X-MET analyzer's performance during this demonstration is provided in Table 4-10.

Table 4-10. Summary of Data Quality Level Parameters

Target Analytes	X-MET 920-MP Analytes	Precision Mean % RSD (mg/kg) 5 - 10 X MDL	Method Detection Limits (mg/kg) (Precision-based)	Coefficient of Determination (r^2 all Data)	Data Quality Level
Arsenic	Arsenic	1.901	50	0.946	Definitive
Barium	Barium	5.91	330	0.620	Qualitative Screening
Chromium	Chromium	ND	115	0.680	Insufficient Data
Copper	Copper	7.92	50	0.690	Qualitative Screening
Lead	Lead	7.60	30	0.878	Definitive
Zinc	Zinc	2.10	30	0.550	Qualitative Screening
Nickel	Nickel	ND	ND	0.320	Insufficient Data
Iron	Not Reported	–	–	–	–
Cadmium	Not Reported	–	–	–	–
Antimony	Not Reported	–	–	–	–

Section 5
Applications Assessment and Considerations

The X-MET 920-MP is designed to provide qualitative and quantitative data on the concentration of metals in soils, sludges, and other solids. X-MET software uses empirical site-specific calibration and quantitation to maximize instrument performance and account for common soil-related matrix interferences. This analyzer is designed for field use; however, its physical and power requirements made it cumbersome to use in the *in situ* mode. The analyzer's design is best suited for benchtop applications. The analyzer experienced some software operating problems resulting in downtime and lost data during the 1-month field demonstration. During this time, more than 1,260 samples were measured by the analyzer. The training provided by the developer was sufficient to allow basic field operation of the analyzer; however, frequent developer assistance was required to address software problems throughout the demonstration. The complexity of an empirical calibration suggested that the analyzer's calibration and performance may improve with increased operator familiarity. The developer provided highly accessible and timely field support. Much of the downtime occurred while using the analyzer in the *in situ* mode. The data lost was a result of software problems and involved less than 8 percent of the data.

Comparison of the analyzer's $\log_{10}$ transformed data to the $\log_{10}$ transformed reference data indicated that the analyzer could produce definitive level quality data for arsenic and lead. This indicated that the analyzer's data were statistically equivalent to the reference data for these analytes. For the other target analytes, barium, copper, and zinc, the analyzer produced qualitative screening level data. The X-MET 920-MP exhibited instrument precision similar to the reference methods, indicating high measurement reproducibility.

The analyzer's probe uses one radioactive source with count times ranging from 60 to 180 live-seconds. Longer count times generally increase accuracy and lower the detection limits but decrease sample throughput. The throughput for the analyzer ranged from 8 to 14 samples per hour. A summary of key operational features is listed in Table 5-1.

There were no apparent effects of site or soil type on performance. This demonstration identified sample preparation as the most important variable with regard to comparability of the FPXRF analyzer to the reference method. For copper and lead, this preparation effect was confounded by changing count times. Therefore, increased comparability for lead and copper, with increased preparation, may also be the result of the associated increase in count times. The analyzer can be applied in an *in situ* or intrusive mode. 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 data and reference data for the analyzer 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.

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

Field portable—Total weight of 10 pounds for probe and electronics unit
Sample throughput (8 to 14 samples per hour)
Produces qualitative screening level data for barium, copper, and zinc, arsenic, and lead at the definitive level
Data is linearly related to EPA SW-846 Methods 3050A/6010A data
Precision—Percent RSD values less than 10 percent at 5 to 10 times the MDL for all analytes
Generally not susceptible to soil matrix effects
Can be used on soils exhibiting up to 30 percent water saturation by weight
Operation requires detailed training (1 to 3 days)
The need for an auxiliary computer reduces the sample throughput in the in situ mode. Battery lifetime of only 2 hours on the auxiliary computer
Empirical calibration requires 10 to 20 well characterized and homogenized site-specific samples
A single source limits the number of elements that can be reported

Based on this demonstration, the analyzer is well suited to the real-time assessment of metals contamination in soil samples. This demonstration showed that the use of this analyzer in an *in situ* mode is less efficient than applying it in an intrusive mode. In addition, more extensive training or operator experience was needed to reduce the potential for data loss and downtime associated with software and calibration problems. Although in several cases the analyzer produced data statistically equivalent to the reference data, generally confirmatory analysis is required for FPXRF analysis. If 10 - 20 percent of the samples measured by the analyzer are submitted for reference method analysis, instrument bias relative to standard methods such as 3050A/6010A can be determined. This will only hold true if the analyzer and the reference laboratory measure similar samples. This was accomplished in this demonstration by thorough sample homogenization. The demonstration showed that the analyzer exhibits a strong linear relationship with the reference data more than a 5 orders of magnitude concentration range. For optimum correlation and bias correction, samples with high, medium, and low concentration ranges from a project must be submitted for reference method analysis. Table 5-2 shows the effects of data correction for the *in situ*-prepared data set. Changes in average relative bias and accuracy are used to show the effects of data correction for the *in situ*-prepared samples.

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

1. Conduct sampling and FPXRF analysis.
2. Select 10 - 20 percent of the sampling locations for resampling. These resampling locations can be evenly distributed over the range of concentrations measured or they can focus on an action level concentration range.
3. Resample the selected locations. Thoroughly homogenize the samples and have each sample analyzed by FPXRF and a reference method.

Table 5-2. Effects of Data Correction on FPXRF 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]
Arsenic	0.89	1.03	1.35	1.27	1.76
Barium	4.14	1.19	4.00	1.74	1.36
Chromium	1.01	1.24	1.92	1.98	1.55
Copper	1.90	2.62	1.63	1.97	1.18
Lead	0.87	1.07	1.39	1.46	1.63
Nickel	0.46	1.20	2.72	2.43	1.56
Zinc	1.03	1.21	1.90	1.90	1.64

Notes:

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

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

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

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

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

The average relative bias is calculated as follows:

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

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

The average relative accuracy is calculated as follows:

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

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

4. Tabulate the resulting data with reference data in the x-axis column (independent variable) and the FPXRF data in the y-axis column (dependent 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:

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

7. Use the above equation with the $\log_{10}$ transformed FPXRF results from Step 4 above and calculate the equivalent $\log_{10}$ corrected FPXRF data.
8. Take the anti-$\log_{10}$ ($10^{[\log_{10} \text{ transformed corrected FPXRF data}]}$) of the equivalent $\log_{10}$ corrected FPXRF data calculated in Step 7. These resulting values (in milligrams per kilogram) represent the corrected FPXRF data.

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

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

In cases where the regression derived corrected average relative accuracy is worse than the raw average relative accuracy, such as seen in Table 5-2 for chromium, copper, lead, nickel, and zinc, the eight-step FPXRF data correction approach presented earlier may not be appropriate. For these elements, data at or below the field-based MDLs were responsible for the decreased accuracy. If the data set in question is representative of the entire population of data being characterized, then the raw FPXRF data can simply be multiplied by the raw average relative accuracy factor for correction. However, the eight-step regression approach should be used whenever the performance of the analyzer is strongly concentration-dependent or if the sample population being used for data correction is not representative of the entire data population being characterized.

The Metorex X-MET 920-MP Analyzer can provide a rapid assessment of the distribution of metals contamination 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 creation of draft Method 6200 will help in the acceptance of FPXRF data for all definitive level applications and possibly qualitative screening level applications. The 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.

General Operational Guidance

The following paragraphs describe general operating considerations for FPXRF analysis. This information is based on the SW-846 Method 6200 for FPXRF analysis.

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

An 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. Many 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 sample preparation was found to take less than 5 minutes per sample location. The last requirement is that the soil or sediment not be saturated with water. Most FPXRF instruments will perform adequately for soils with moisture contents of 5 - 20 percent, but will not perform well for saturated soils, especially if ponded water exists on the surface. Data from this demonstration did not indicate 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 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. 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. Microwave oven drying can also melt plastic containers used to hold 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 of the sample should then be placed in a 31-mm polyethylene sample cup (or equivalent) for analysis. The sample cup should be completely filled and covered with a 2.5-micrometer Mylar™ (or equivalent) film for analysis. The rest of the soil sample should be placed in a jar, labeled, and archived. All equipment, including the mortar, pestle, and sieves, must be thoroughly cleaned so that the sample 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.