EPA/600/R-97/147
March 1998

Environmental Technology Verification Report

Field Portable X-ray Fluorescence Analyzer

Scitec MAP Spectrum Analyzer

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

Notice

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

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

ENVIRONMENTAL TECHNOLOGY VERIFICATION PROGRAM

ENVIRONMENTAL TECHNOLOGY VERIFICATION PROGRAM
VERIFICATION STATEMENT

TECHNOLOGY TYPE:	**FIELD PORTABLE X-RAY FLUORESCENCE ANALYZER**
APPLICATION:	**MEASUREMENT OF METALS IN SOIL**
TECHNOLOGY NAME:	**MAP SPECTRUM ANALYZER**
COMPANY:	**SCITEC CORPORATION**
ADDRESS:	**415 N. QUAY** **KENNEWICK, WA 99336**
PHONE:	**(800) 466-5323**

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 Scitec MAP Spectrum Analyzer.

PROGRAM OPERATION

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

DEMONSTRATION DESCRIPTION

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

The demonstration was designed to detect and measure a series of inorganic analytes in soil. The primary target analytes were arsenic, barium, chromium, copper, lead, and zinc; nickel, iron, cadmium, and antimony were secondary analytes. The demonstration sites were located in Iowa (the RV Hopkins site) and Washington (the ASARCO site). These sites were chosen because they exhibit a wide range of concentrations for most of the target metals and are located in different climatological regions of the United States; combined, they exhibit three distinct soil types: sand, clay, and loam. The conditions at these sites are representative of those environments under which the technology would be expected to operate. Details of the demonstration, including a data summary and

discussion of results, may be found in the report entitled "Environmental Technology Verification Report, Field Portable X-ray Fluorescence Analyzer, Scitec MAP Spectrum Analyzer." The EPA document number for this report is EPA/600/R-97/147.

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

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

The MAP Spectrum Analyzer was originally designed to detect lead on painted surfaces using a cobalt-57 excitation source. It is now marketed for detecting lead and other metals in soil, especially when equipped with a cadmium-109 source. Two other sources, americum-241 and cobalt-57, are also available. The MAP Spectrum Analyzer was empirically calibrated by the developer prior to the demonstration using site-specific calibration standards. The instrument designed to be portable, is composed of two parts, the scanner which weighs 3.5 pounds and an 11-pound control console. In this demonstration, the MAP Spectrum Analyzer was configured to report four of the primary target analytes: arsenic, copper, lead, and zinc. It was operated only in the *in situ* mode. At the time of the demonstration, the cost of the MAP Spectrum Analyzer with the cadmium-109 source was $32,000, or it could be leased for $4,675 per month.

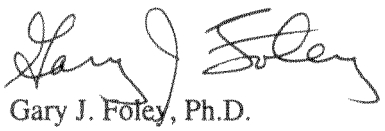
VERIFICATION OF PERFORMANCE

The performance characteristics of the MAP Spectrum Analyzer include the following:

- **Detection limits:** Precision-based detection limits were determined by collecting 10 replicate measurements on site-specific soil samples with metals concentrations 2 to 5 times the expected MDLs. Results ranged from 25 milligrams per kilogram (mg/kg) for zinc to 525 mg/kg for copper. Corresponding values were 225 mg/kg for arsenic and 165 mg/kg for lead.
- **Throughput:** Average throughput was 9 - 12 samples per hour using a live count time of 240 seconds. This rate only represents the analysis time since different personnel were used to prepare the samples.
- **Drift:** Based on a periodic analysis of a calibration check sample, drift was the greatest for copper and least for zinc. The drift values for the mean recovery of copper varied from -25 to +35 percent; arsenic was ±15 percent; lead was -15 to +25 percent; and zinc was ±5 percent.
- **Completeness:** The MAP Spectrum Analyzer produced results for 628 of the 630 *in situ* samples for a completeness of 99.7 percent, above the demonstration objective of 95 percent.
- **Blank results:** Three of the four reported analytes were not detected above the field-based method detection limits in the blanks. Anomalous readings were reported for copper but were considered to be an artifact of the blank measurement process.
- **Precision:** The goal of the demonstration was to achieve relative standard deviations (RSD) of less than 20 percent at analyte concentrations of 5 to 10 times the method detection limits. The RSD values for arsenic, lead, and zinc were less than 9 percent RSD. Copper had an RSD of less than 15 percent.

- **Accuracy:** Accuracy was assessed by using site-specific soil PE samples and soil SRMs. The data showed that 5 of 17 results (29.4 percent) of the analytes in these samples had recoveries within a quantitative acceptance range of 80 - 120 percent. This analyzer showed the greatest accuracy for lead with 50 percent of the samples within the 80 - 120 percent recovery range. The instrument underestimated arsenic and copper in the site-specific PE samples, especially at low concentrations. Recovery values for zinc were inconsistent but overall were underestimated.
- **Comparability:** This demonstration showed that the MAP Spectrum Analyzer produced data that exhibited a $\log_{10}$-$\log_{10}$ linear correlation to the reference data. The coefficient of determination (r^2) which is a measure of the degree of correlation between the reference and field data was 0.85 for lead, 0.80 for copper, 0.76 for arsenic, and 0.67 for zinc.
- **Data quality levels:** Using the demonstration derived precision RSD results and the coefficient of determination as the primary qualifiers, the MAP Spectrum Analyzer produced definitive level data for lead; data of quantitative screening level for copper and arsenic; and data of qualitative screening level for zinc.

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



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

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

Foreword

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

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

The EPA's Super-fund Innovative Technology Evaluation (SlTE) Program evaluates technologies for the characterization and remediation of contaminated Superfund and Resource Conservation and Recovery Act (RCRA) corrective action sites. The SlTE 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 ORDs 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 FPXRF 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 allows 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 evaluated, 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 regarding the performance of the Scitec MAP Spectrum Analyzer. Separate ETVRs have been published for the other analyzers demonstrated.

Quantitative data were provided by the MAP Spectrum Analyzer on a real-time basis. This FPXRF analyzer was configured to report arsenic, copper, lead, and zinc. The analyzer used a count time of 240 live-seconds, which resulted in a throughput of 9 to 12 samples per hour. The analyzer used one radioactive source, cadmium-109 coupled to a solid-state silicon detector. The MAP Spectrum Analyzer provided definitive level data (equivalent to reference data) for lead; quantitative screening level data (not equivalent to reference data, but correctable by collecting confirmatory samples) for copper and arsenic; and qualitative screening level data (identifies presence or absence only) for zinc. The analyzer exhibited precision at 5 to 10 times the method detection limits of less than 15 percent relative standard deviation (RSD) for all four of the reported analytes. The analyzer generally exhibited a precision similar to the reference method.

The analyzer's quantitative results were based on an empirical calibration using site-specific calibration samples. Field-based method detection limits (MDL) for this analyzer were slightly lower than the precision-based MDLs for arsenic, copper, and lead, but much higher for zinc. Data correction had limited effect on the analyzer's average relative bias and accuracy. Except for copper, the precision-based and field-based MDLs were below the developer's projected MDL of 250 mg/kg. The site variable did not affect data comparability. The soil variable showed a slight trend of poorer comparability in loam soils. This study showed that the MAP Spectrum Analyzer produced data that exhibited $\log_{10}$-$\log_{10}$ linear correlation for all four of the reported analytes.

This demonstration found that the MAP Spectrum Analyzer was simple to operate in the field. This FPXRF analyzer is used only in the *in situ* mode which means it analyzed samples in minimally disturbed soil. The operator required no specialized training or experience to operate the analyzer. Ownership and operation of this instrument may require specific licensing by state nuclear regulatory agencies. There are special radiation safety training requirements and costs associated with this type of licensing.

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

Table of Contents

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-l 09
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, Frank Bryant (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 Kevin Dorow, Scitec Corporation.

Section 1
Executive Summary

In April 1995, the U.S. Environmental Protection Agency (EPA) sponsored a demonstration of field portable X-ray fluorescence (FPXRF) analyzers. The primary objectives of this demonstration were to evaluate these analyzers for: (1) their analytical performance relative to standard analytical methods, (2) the influence of sample matrix variations (texture, moisture, heterogeneity, and chemical composition) on performance, (3) the logistical and economic resources needed to operate these technologies in the field, and (4) to test and validate an SW-846 draft method for FPXRF analysis. Secondary objectives for this demonstration were to evaluate FPXRF analyzers for their reliability, ruggedness, cost, range of usefulness, 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) plan, implement, and report on a demonstration of FPXRF analyzers. This demonstration was conducted under the EPA's Superfund Innovative Technology Evaluation (SITE) Program and managed by the National Exposure Research Laboratory-Environmental Sciences Division (NERL-ESD) under the Monitoring and Measurement Technologies Program (MMTP), Las Vegas, Nevada.

The FPXRF analyzers tested in this demonstration were designed to provide rapid, real-time analysis of metals concentrations in soil samples. This information will allow investigation and remediation decisions to be made on-site more efficiently, and it should reduce the number of samples that need to be submitted for confirmatory analysis. Of the seven commercially available analyzers evaluated, 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 a portion of the demonstration. This environmental technology verification report (ETVR) presents information regarding the Scitec MAP Spectrum Analyzer. Separate ETVRs will be published for the other analyzers that were demonstrated.

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

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

This demonstration found that the MAP Spectrum Analyzer was simple to operate in the field. It was designed to be used in the in situ mode; that is to analyze samples in minimally disturbed soil. The developer provided a training course for the technology operator which was similar to that provided to a purchaser of the equipment. The training encompassed enough FPXRF theory and hands-on use to allow the operator to manipulate the data collection software, calibrate the analyzer, and adjust instrument parameters such as count times and target analytes. In addition, the developer provided radiation safety training, required for the use of this analyzer. A license was obtained from the State of Kansas, which has reciprocal licensing agreements with States of Iowa and Washington. The Scitec technical staff provided accessible and timely field support. The analyzer itself was portable and was operated continuously more than a 10 to 12-hour work day with appropriate battery changes. The rainy weather conditions encountered during the demonstration caused no operational downtime for the analyzer.

The analyzer used one radioactive source, cadmium-109, coupled to a solid-state silicon detector. The count times used in this demonstration (240 live-seconds) resulted in a sample throughput of 9 - 12 samples per hour. The MAP Spectrum Analyzer produced data meeting definitive level criteria (equivalent to reference data) for lead; data meeting quantitative screening level criteria (not equivalent to reference data, but correctable with confirmatory sample analysis) for copper and arsenic; and data meeting qualitative screening level criteria (identifies the presence or absence of contamination) for zinc.

The analyzer generally exhibited precision similar to that of the reference methods. Field-based method detection limits (MDL) for this analyzer were lower than the precision-based values for arsenic, copper, and lead, but much higher for zinc. Most of the precision-based and field-based MDLs were below the developer's projected MDL of 250 mg/kg. The site variable did not appear to affect data comparability. The soil variable showed a slight trend of poorer comparability in loam soils. Data correction had limited effect on the analyzer's average relative bias and accuracy.

Based on the performance of the analyzer, this demonstration found the MAP Spectrum Analyzer to be an effective tool for characterizing the concentration of target metals in 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 MAP Spectrum Analyzer. This analyzer was developed by Scitec Corporation to perform elemental analyses (metals quantitation) in the field, most commonly lead in soil and paint. This analyzer uses a solid-state silicon detector and a cadmium-109 (Cd^{109}) source to detect metals in the test sample. The analyzer is designed to operate in the in situ mode; this is commonly referred to as "point-and-shoot." In this mode of operation, the point of measurement on the soil surface is cleared of loose debris and organic matter, the analyzer's probe is then placed directly on the soil surface, and a measurement is taken.

This section provides general information about the demonstration including the purpose, objectives, and design. Section 3 presents and discusses the quality of data produced by the reference methods against which the analyzer was evaluated. Section 4 discusses the MAP Spectrum Analyzer's capabilities, reliability, throughput, accuracy, precision, comparability to reference methods, and other evaluation factors. Section 5 discusses the potential applications of the analyzer, presents a method for data correction, and suggests a framework for a standard operating procedure (SOP). Section 6 lists 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 the performance of FPXRP analyzers to identify and quantify concentrations of metals in soils. The primary objectives were to evaluate FPXRP analyzers in the following areas: (1) their accuracy and precision relative to conventional analytical methods; (2) the influence of sample matrix variations (texture, moisture, heterogeneity, and chemical composition) on their performances; (3) the logistical and economic resources necessary to operate these analyzers; and (4) to test and validate an SW-846 draft method for FPXRP analysis.

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

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

Reference Methods

To assess the performance of each analyzer, FPXRF data were compared to reference data. The reference methods used for this assessment were EPA SW-846 Methods 3050A/6010A, which are considered the standards for metals analysis in soil for environmental applications. For purposes of these discussions, the term "reference" was substituted for "confirmatory" since the data were used as a baseline for comparison. MRI was awarded the subcontract to analyze soil samples using the reference methods in accordance with Federal Acquisition Regulations. 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 sample analyses in the required timeframe.

Method 3050A is the standard acid extraction procedure used 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 (Section 3).

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

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

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

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

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

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

Site Selection

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

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

PRC contacted NERL-ESD, regional EPA offices, state environmental agencies, metals fabrication, and smelting contacts to create an initial list of potential demonstration sites. PRC received considerable assistance from the EPA RCRA and Superfund Branches in Regions 4,6,7,8,9, and 10. PRC also contacted the Montana Department of Health and Environment, the Nevada Bureau of Mines and Geology, the Oklahoma Department of Environmental Quality, the Arizona Department of Environmental Quality, the Missouri Department of Natural Resources, the Arizona Bureau of Geology, and the New Mexico Bureau of Mines and Mineral Resources. PRC surveyed its offices in Kansas City,

Kansas; Atlanta, Georgia; Denver, Colorado; Dallas, Texas; Albuquerque, New Mexico; Helena, Montana; Chicago, Illinois; Seattle, Washington; and San Francisco, California, for information regarding potential sites. These PRC offices have existing RCRA, Superfund, or Navy environmental contracts that allow access to regional, state, and federal site information. PRC also used the Record of Decision Scan database (Morgan and others 1993) to search for appropriate sites.

PRC screened 46 potential sites based on the site-selection criteria with the assistance of the various contacts listed above. Based on this screening effort, PRC and EPA determined that the RV Hopkins and ASARCO sites met most of the site-selection criteria, and therefore, would be the 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 targeted four general areas of the site: 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 the site: a granular fill and a massive slag fill. The composition of the granular fill material ranges from sand to silt with demolition debris and slag debris mixed throughout. The massive 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 is 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.

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 1800s. 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 this 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/60l0A. Twenty-nine of these samples were split and sent to the developers. Nine field duplicates were collected and submitted for reference method analysis to assess proposed sample homogenization procedures. One purchased PE sample also was submitted to the reference laboratory to provide an initial check of its accuracy.

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

Experimental Design

The experimental design for this demonstration was developed to meet the primary and secondary objectives stated above, and was approved by all 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 sample measurements 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 MAP Spectrum Analyzer is designed to operate in the *in situ* mode. The sampling and analysis procedure was designed to test the common application of FPXRF analyzers. The sampling procedure used is illustrated in Figure 2-l. Since the MAP Spectrum Analyzer operates in the *in situ* mode only, the discussion of the experimental design will be limited to *in situ* sample preparation and analysis.

For *in situ* analysis, an area 4 inches by 4 inches square was cleared of all vegetation, debris, and gravel larger than 2 millimeters (mm) in diameter. The analyzer took one *in situ* measurement in each sample area. These data represented FPXRF *in situ* measurements for unprepared soils *(in situ*-unprepared). Replicate measurements were taken at 4 percent of these locations to assess analyzer precision. Figure 2-l depicts the sample analysis chain for *in situ analyses. The* MAP Spectrum Analyzer only reported *in* situ-unprepared and *in* situ-prepared samples.

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

The homogenized sample was then spread out inside a l-inch-deep petri dish. The FPXRF analyzer then took one measurement of this homogenized material. This represented the homogenized sample analysis for the *in situ* analyzers *(in* situ-prepared). This process represents 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 soil samples that were used for the unprepared precision determination.

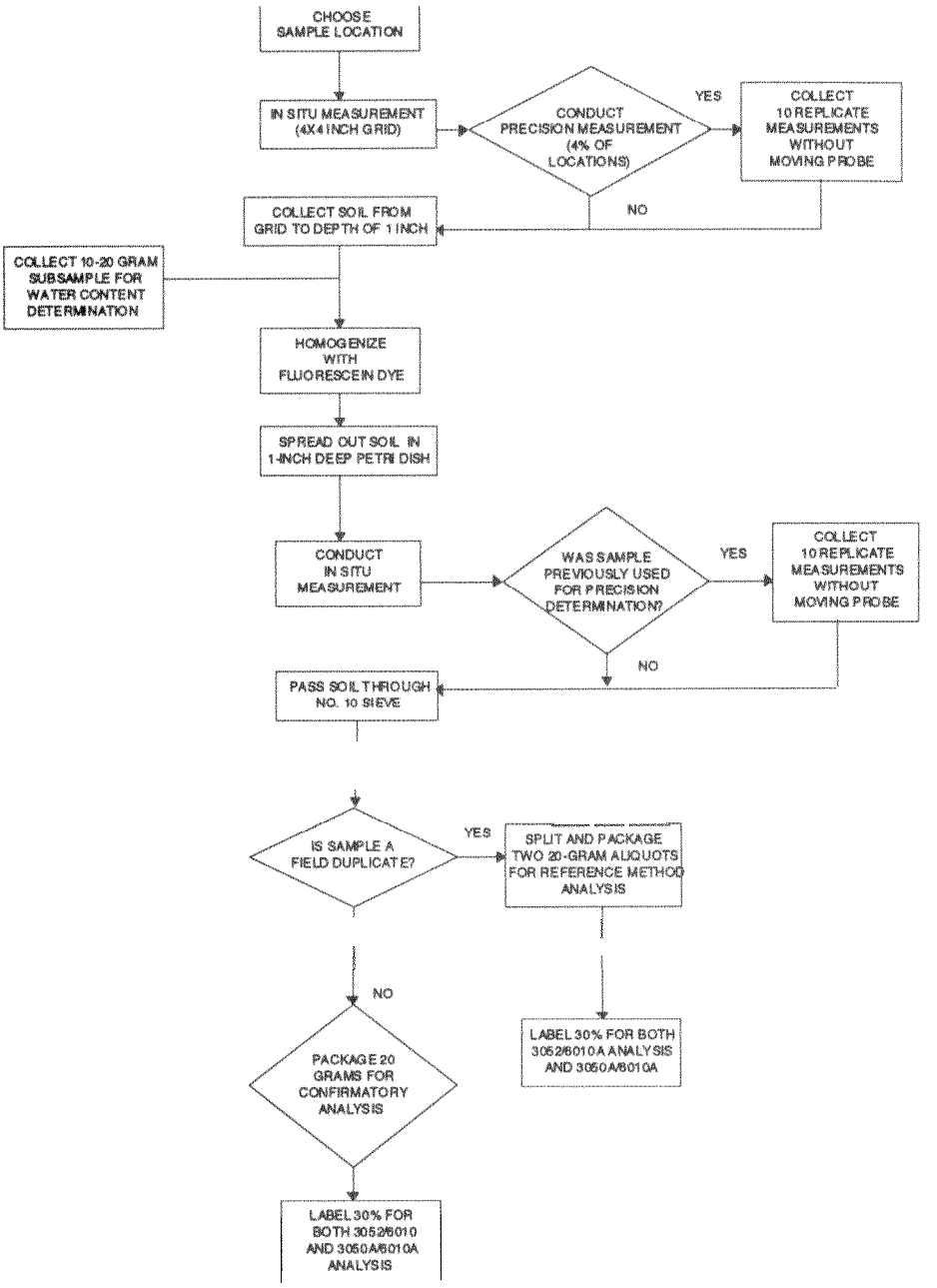


Figure 2-1. Sample Preparation and Analysis: This flowchart depicts the handling procedures for each sample collected for analysis by the MAP Spectrum Analyzer.

Qualitative Factors

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

Many analytical methods exhibit significant "operator effects," in which individual differences in sample preparation or operator technique result in a significant effect on the numerical results. To reduce the possible influence of operator effects, a single operator was used to operate each FPXRP analyzer. While this reduced some potential error from the evaluation, it did not allow the analyzers to be evaluated for their susceptibility to operator-induced error. A single operator was used to analyze all of the samples at both sites during this demonstration. Sample preparation variation effects were minimized in the field by using the same personnel to prepare samples. To eliminate the influence of operator effects on the reference method analysis, only one reference laboratory was used to analyze the samples. Based on this design, there is 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 FPXRP analyzers were to be 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 analytes determined by the MAP Spectrum Analyzer were arsenic, copper, lead, and zinc.

Evaluations of analyzer data comparability involved examining the effects of each site, soil texture, and sample preparation technique (Table 2-l). Two sites were sampled for this demonstration. Thus, 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). The demonstration plan identified four sample preparation steps: (1) in situ-unprepared, (2) *in situ*-prepared, (3) intrusive-unprepared, and (4) intrusive-prepared (samples generated in steps 3 and 4 were not analyzed by the MAP Spectrum Analyzer). 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, while the ASARCO data were divided into sand and loam soil textures; then each soil texture was subdivided by the soil preparations. This design allowed for the examination of particle size and homogenization effects on data comparability. These effects were believed to have the greatest impact on data comparability.

Table 2-1. Performance and Comparability Variables Evaluated

Variables		
Site Name (315)	**Soil Texture (315)**	**Preparation Step [630]**
ASARCO (215)	Sand (100)	in situ-unprepared [100] in situ-prepared [100]
	Loam (115)	in situ-unprepared [115] in situ-prepared [115]
RV Hopkins (100)	Clay (100)	in situ-unprepared [100] in situ-prepared [100]

Notes:

( ) Total number of sample points.

[ ] Total number of measurements taken.

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

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

Another quantitative variable evaluated was the count time used to acquire data. During the formal sample quantitation and precision measurement phase of the demonstration, the count times were set by the developers and remained constant throughout the demonstration. Count times can be tailored to 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. Quantitative radiation readings were made with a gamma ray detector near each analyzer to assess the potential for exposure to radiation.

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

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

- Heterogeneity: For in situ-unprepared measurements, heterogeneity was partially controlled by restricting measurements within a 4-by-4-inch area. For measurements after the initial point-and-shoot preparation, heterogeneity was minimized by sample homogenization. This effect was evaluated through the sample preparation data.
- Particle Size: Since no intrusive samples were analyzed, the effect of particle size was not determined for this analyzer.
- 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 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 data. EPA SW-846 Methods 3050A/6010A were selected as the reference methods because they represent the regulatory standard against which FPXRF is generally compared. In comparing the FPXRF data and reference data, it is important to recognize that, while similar, the process by which the data are obtained is not identical. While there is significant overlap in the nature of the analysis, there are also major differences. These differences, or "perspectives," allow the user to characterize the same sample in

slightly different ways. Both have a role in site characterization and remediation monitoring. It is important to consider these differences and the measurement error intrinsic to each method when comparing the FPXRF method against a reference method.

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

The comparison of FPXRF data and the reference data used 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 to consider include: (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 FPXRF data and reference data are, in fact, related linearly and that this assumption is correct.

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

The use of least squares linear regression has certain limitations. Least squares regression provides a linear equation, which minimizes the squares of the differences between the dependent variable and the regression line. For data sets produced in this demonstration, the variance was proportional to the magnitude of the measurements. That is, a measurement of 100 parts per million (ppm) may exhibit a 10 percent variance of 10 ppm, while a 1,000 ppm measurement exhibits a 10 percent variance of 100 ppm. For data sets with a large range in values, the largest measurements in a data set exert disproportionate influence on the regression analysis because the least squares regression must account for the variance associated with the higher valued measurements. This can result in an equation that has minimized error for high values, but almost neglects error for low values because their influence in minimizing dependent

variable error is small or negligible. In some cases, the resulting equations, biased by high-value data, may lead to inappropriate conclusions concerning data quality. The range of the data examined for the analyzers spanned between 1 and 5 orders of magnitude (e.g., 10 - 100,000 ppm) for the target analytes. This wide range in values and the associated wide range in variance (influenced by concentration) created the potential for this problem to occur in the demonstration data set. To provide a correlation that was equally influenced by both high and low values, logarithms (log_{10}) of the dependent and independent variables were used, thus, scaling the concentration measurements and providing equal weight in the least squares regression analysis to both small and large values (Figure 2-2). All statistical evaluations were carried out on log,, transformed data.

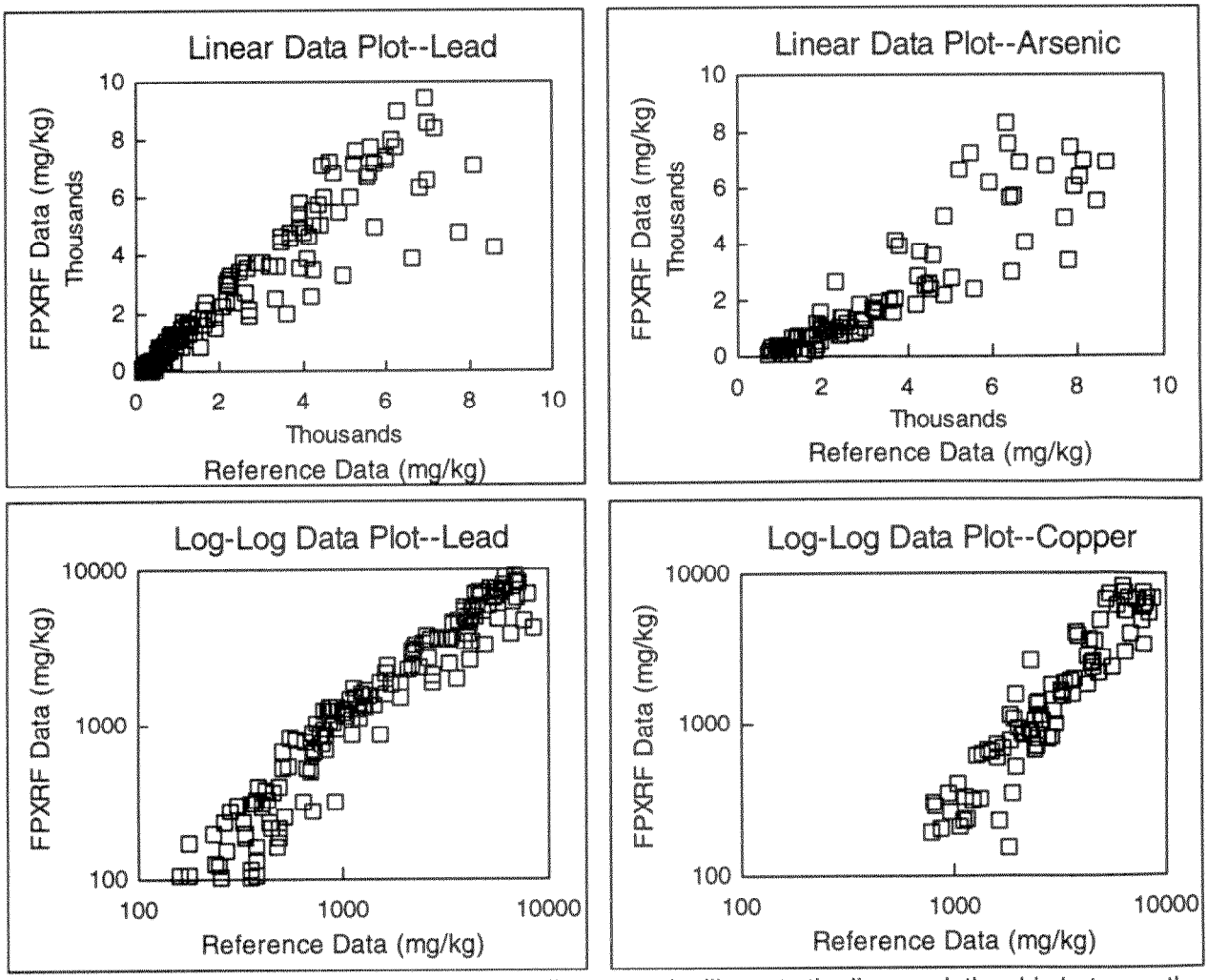


Figure 2-2. Linear and Log-log Data Plots: These graphs illustrate the linear relationship between the MAP Spectrum Analyzer's data and the reference data. The linear data plots illustrate the concentration dependence of this relationship with increased scatter at higher concentrations. The log-log plots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

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

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

Definitive level data are considered the highest level of quality. These data are usually generated by using well-defined, rigorous analytical 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. Data may be generated in the field, as long as the 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 the screening data be confirmed using analytical methods and QA/QC procedures and criteria associated with definitive data. The quality of unconfirmed screening data cannot be determined.

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

At the time of this demonstration, an approved EPA method for FPXRF did not exist. As part of this demonstration, PRC prepared a draft Method 6200 "Field Portable X-Ray Fluorescence Spectrometry for the Determination of Elemental Concentrations in Soil and Sediment." The draft method has been subsequently submitted for inclusion in Update 4 of SW-846 scheduled for approval in FY-97. For the

purposes of this demonstration, the lack 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 MAP Spectrum Analyzer consisted of 630 matched pairs of FPXRF and reference data for each target analyte. This data set was analyzed as a whole and then subdivided and analyzed with respect to each of the variables listed in Table 2-l. This nesting of variables allowed the independent assessment of the influence of each variable on comparability.

For the performance evaluation of this analyzer, a total of 315 soil samples was analyzed by the reference methods. These samples were analyzed by the MAP Spectrum Analyzer for the in situ preparation steps and produced 630 data points. 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 variables were evaluated for all target analytes. If the evaluation of the influence of a given variable did not result in a better correlation, as exhibited by a higher coefficient of determination (r^2) and smaller standard error of the estimate (using $\log_{10}$ transformed data), then the influence was considered to be insignificant. However, if the correlation worsened, the cause was examined and explained. If the correlation improved, resulting in an improved 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. Therefore, 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-l). 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.

$$Y = m x + b \qquad (2\text{-}1)$$

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

The matched pairs t-test was also used to evaluate whether the two sets of $\log_{10}$ transformed data sets 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.

Table 2-2. Criteria for Characterizing Data Quality

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

Notes:
[a] The statistical tests and parameters are discussed later in the "Intermethod Assessment" 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.

Slope Test for Significant Differences (2-2)

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

where

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

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

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

where

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

If the r^2 was between 0.70 and 1, the precision was between 10 and 20 percent RSD, and the slope or intercept were not statistically equivalent, then the analyzer was considered to produce quantitative screening level data quality. However, the linear regression was deemed sufficient so 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 from 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 RSD greater than 20 percent. An analyzer producing data at this level is considered capable of detecting the presence or lack of contamination, above its detection limit, with at least a 90 percent accuracy rate, but it is not considered suitable for reporting of concentrations.

MDLs for the analyzers were determined in two ways. One approach followed a standard SW-846 protocol. In this approach, standard deviations (SD) from precision measurements for samples exhibiting contamination 5 to 10 times the estimated detection levels of the analyzers were multiplied by 3. The result represented the precision-based 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. The value of the MDL was assigned by determining the point where the linear relationship disintegrates and assigning the MDL at two SDS above this concentration.

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 would be dried in a convection oven at 150 °C for 2 hours. Through visual observation, it was found that the samples were completely dried in 1 hour with samples heated to only 110 "C. Therefore, to conserve time, and to reduce the potential volatilization of metals from the samples, 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 the drying process had a significant impact on the analytical results. The mean RPD for the microwaved and nonmicrowaved raw data were significantly different at a 95 percent confidence level. This suggests that the microwave drying process somehow increases error and sample concentration variability. This difference may be due to the extreme heat 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 the concentration of metals. The RPDs for the microwaved samples were 2 to 3 times worse than the RPDs from the field duplicates. This further supports the hypothesis that microwave drying increases variability.

The fifth deviation involved reducing the percentage of analyzer precision measuring points. The demonstration plan called for 10 percent of the samples to be used for assessment of analyzer precision, Due to the time required to complete analysis of an analyzer precision sample, only 4 percent of the samples were used to assess analyzer precision. This reduction in samples was approved by the EPA technical advisor and the PRC field demonstration team leader. This eliminated 720 precision measurements and saved 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. In addition, 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 were used for comparative purposes for the reference methods as were PE data for the FPXRF data.

Sample Homogenization

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

Sodium fluorescein was used as an indicator of complete sample homogenization. Approximately one-quarter teaspoon of dry sodium fluorescein salt was added to each sample prior to homogenization.

After the homogenization was completed, the sample was examined under an ultraviolet light to assess the distribution of sodium fluorescein throughout the sample. If the fluorescent dye was evenly dispersed in the sample, homogenization was considered complete. If the dye was not evenly distributed, the mixing was continued and checked until the dye was evenly distributed throughout the sample.

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

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

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

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.

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

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-l. 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 pre-demonstration 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 Ho/ding Times

The maximum allowable holding time from the date of sample collection to the date of extraction and analysis using EPA SW-846 Methods 305OA/6OlOA 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 60l0A 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 60l0A requires an initial calibration verification (ICV) standard to be analyzed with each ICAL. The method control limit for the ICV is ±10 percent. An interference check sample (ICS) and a high level calibration check standard is required to be analyzed with every ICAL to assess the accuracy of the ICAL. The control limits for the ICS and high level calibration check standard were ±20 percent recovery and ±10 percent of the true value, respectively. All ICALs, ICVs, and ICSs met the respective QC requirements for all target analytes.

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

Detection Limits

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

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

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

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

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

Method Blank Samples

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

Laboratory Control Samples

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

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

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

Predigestion Matrix Spike Samples

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

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

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

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

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

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

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

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

Postdigestion Matrix Spike Samples

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

Secondary analytes, antimony, and iron were observed outside the control limits. However, no corrective action was taken for secondary analytes as stated in the demonstration QAPP. All

postdigestion spike recoveries for target analytes met the QA/QC requirements of the QAPP and were considered acceptable.

Predigestion Laboratory Duplicate Samples

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

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

Postdigestion Laboratory Duplicate Samples

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

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

Performance Evaluation Samples

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

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

to the reference laboratory for analysis by EPA SW-846 Methods 3050A/60l0A. Thereference 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, XRP, 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 (NISI), 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/60l0A as were the ERA PE or RTC CRM samples. Comparison of EPA SW-846 Methods 3050A/60l0A 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/60l0A. 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-l 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-l). 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-l. 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-l). 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.

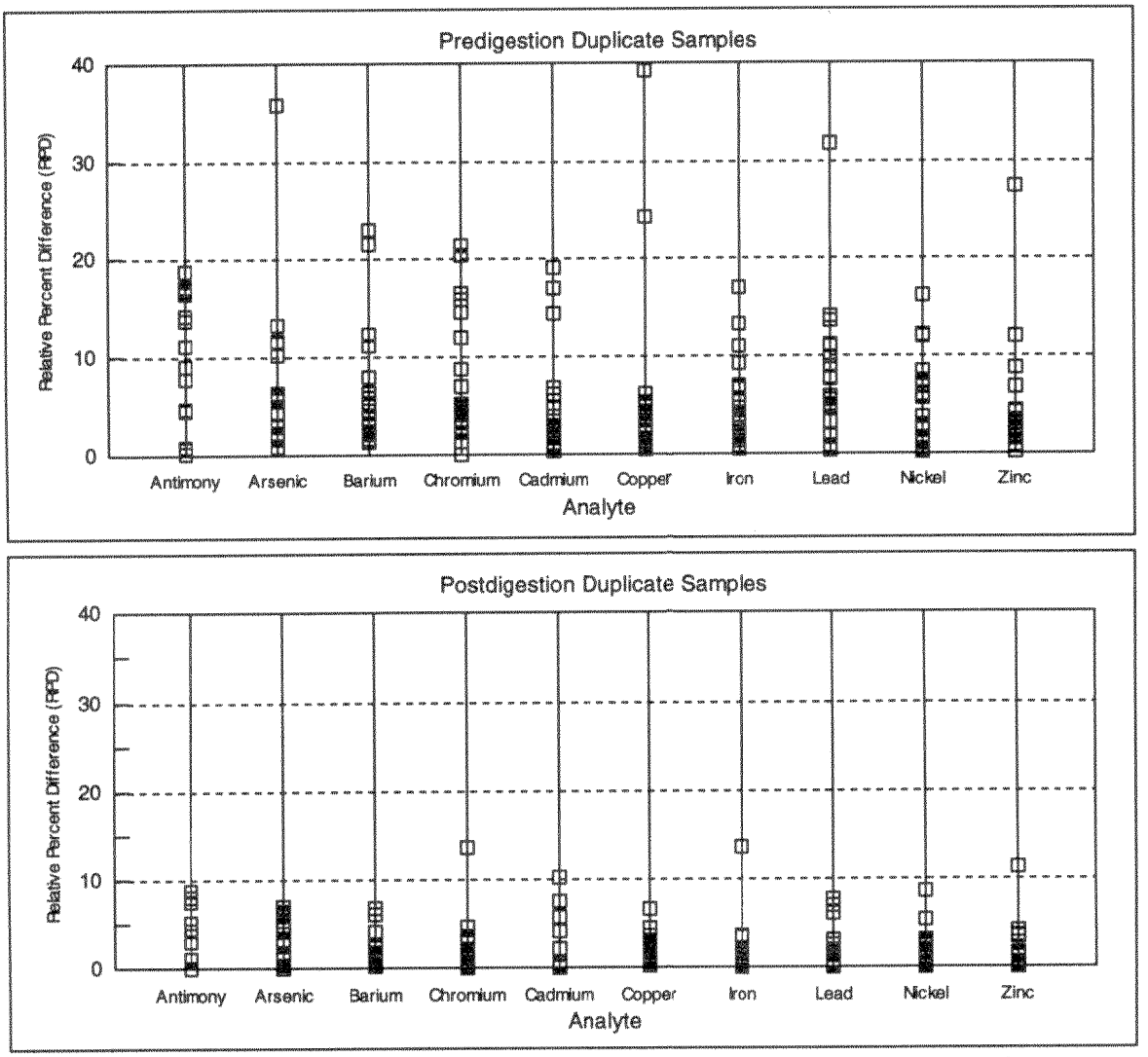


Figure 3-1. Pre- and Postdigestion Duplicate Samples: The top graph illustrates the reference laboratory's performance on analyzing predigestion duplicate samples. Twenty percent RPD represents the predigestion duplicate control limits defined in the demonstration QAPP. Two points were deleted from this 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/60l0A. 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.

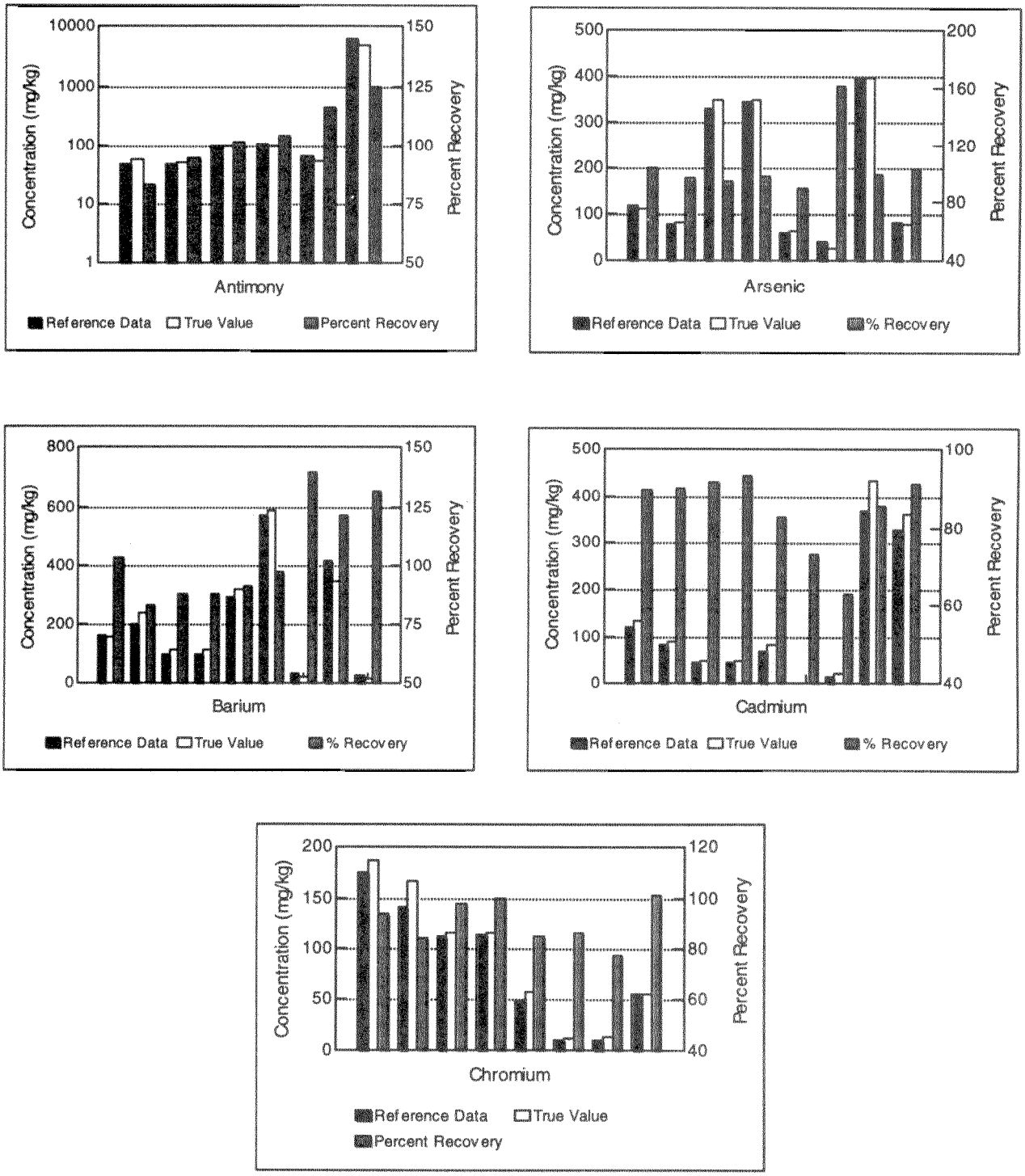
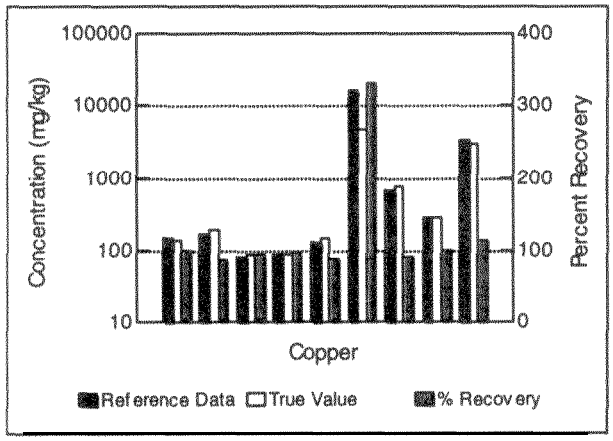
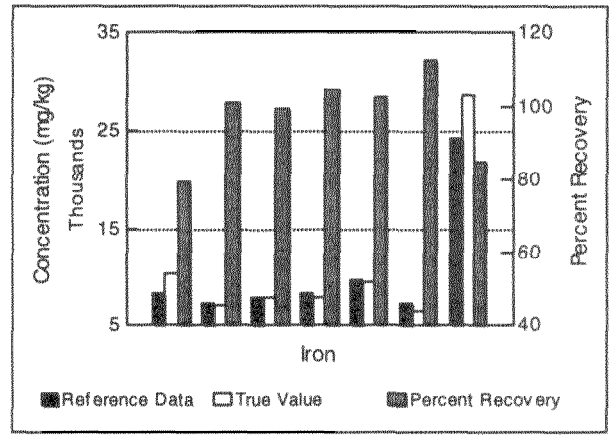
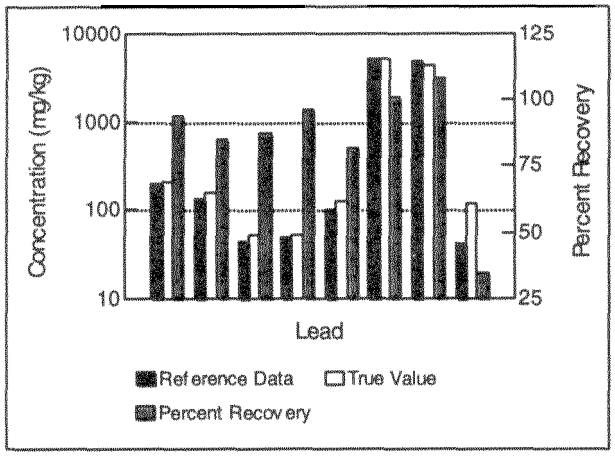
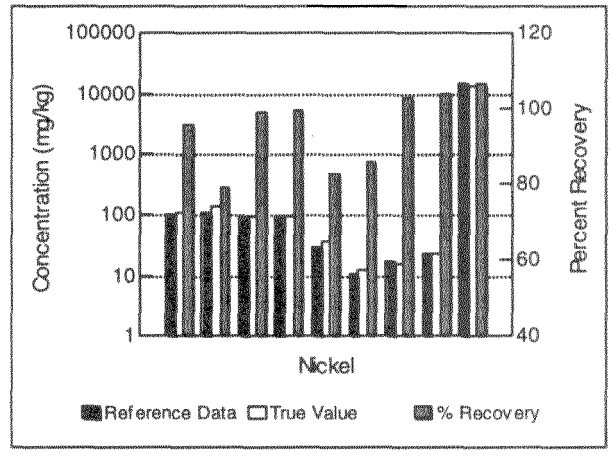


Figure 3-2. Reference Method PE and CRM Results: These graphs illustrate the relationship between the reference data and the true values for the PE or CRM samples. The gray bars represent the percent recovery for the reference data. Each set of three bars (black, white, and gray) represents a single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









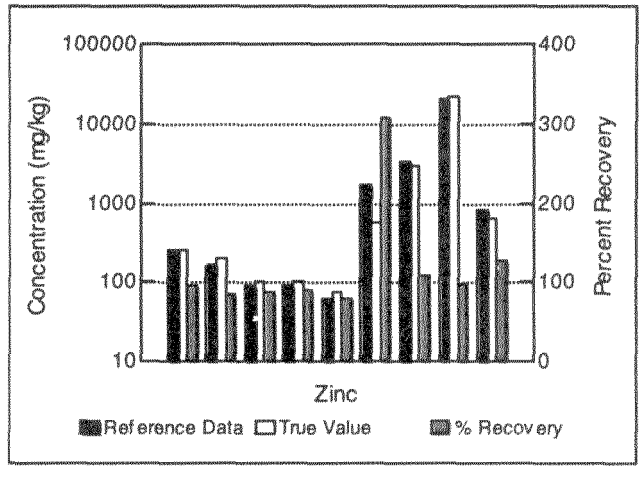


Figure 3-2 (Continued). Reference Method PE and CRM Results: These graphs illustrate the relationship between the reference data and the true values for the PE or CRM samples. The gray bars represent the percent recovery for the reference data. Each set of three bars (black, white, and gray) represents 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/60l0A. 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/60l0A. 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/60l0A 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/60l0A. 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/60l0A 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/60l0A 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.

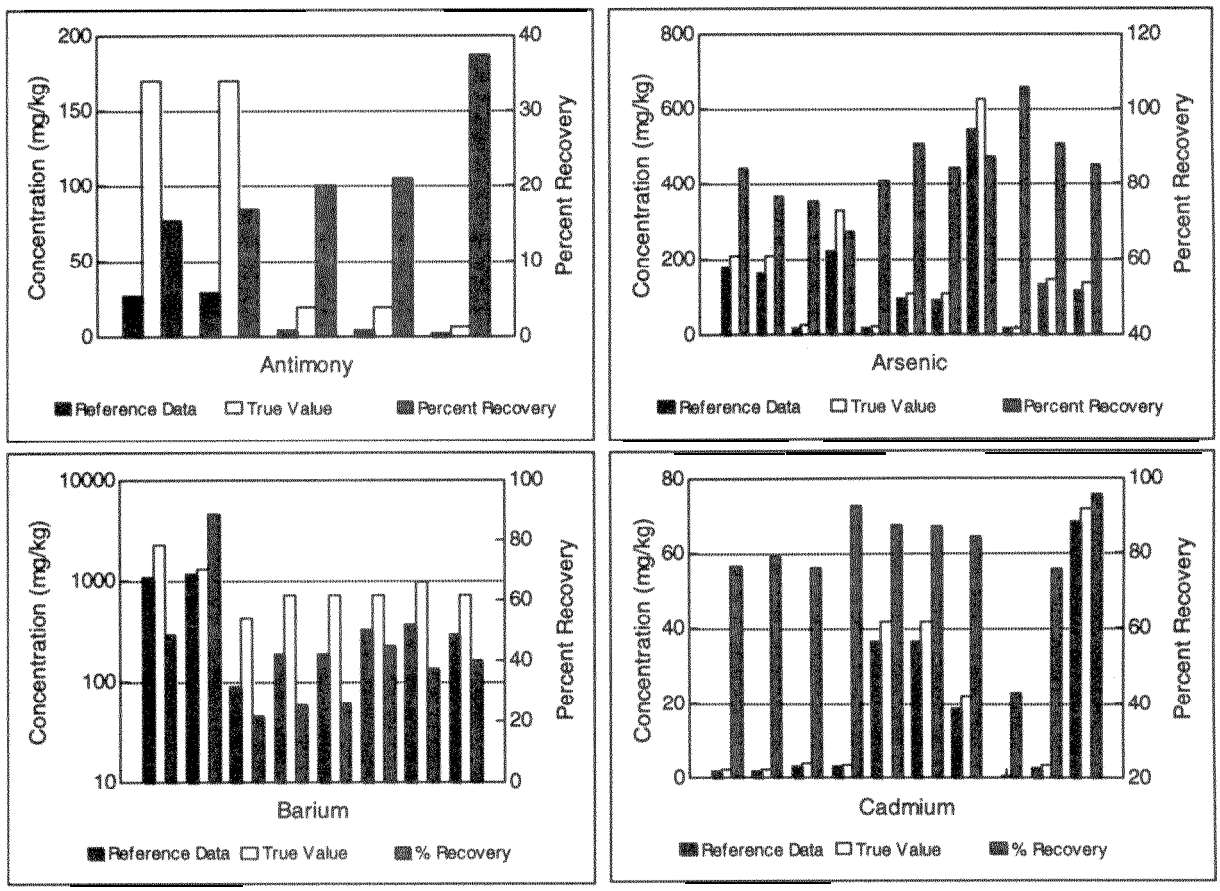


Figure 3-3. Reference Method SRM Results: These graphs illustrate the relationship between the reference data and the true values for the SRM samples. The gray bars represent the percent recovery for the reference data. Each set of three bars (black, white, and gray) represents a single SRM sample.

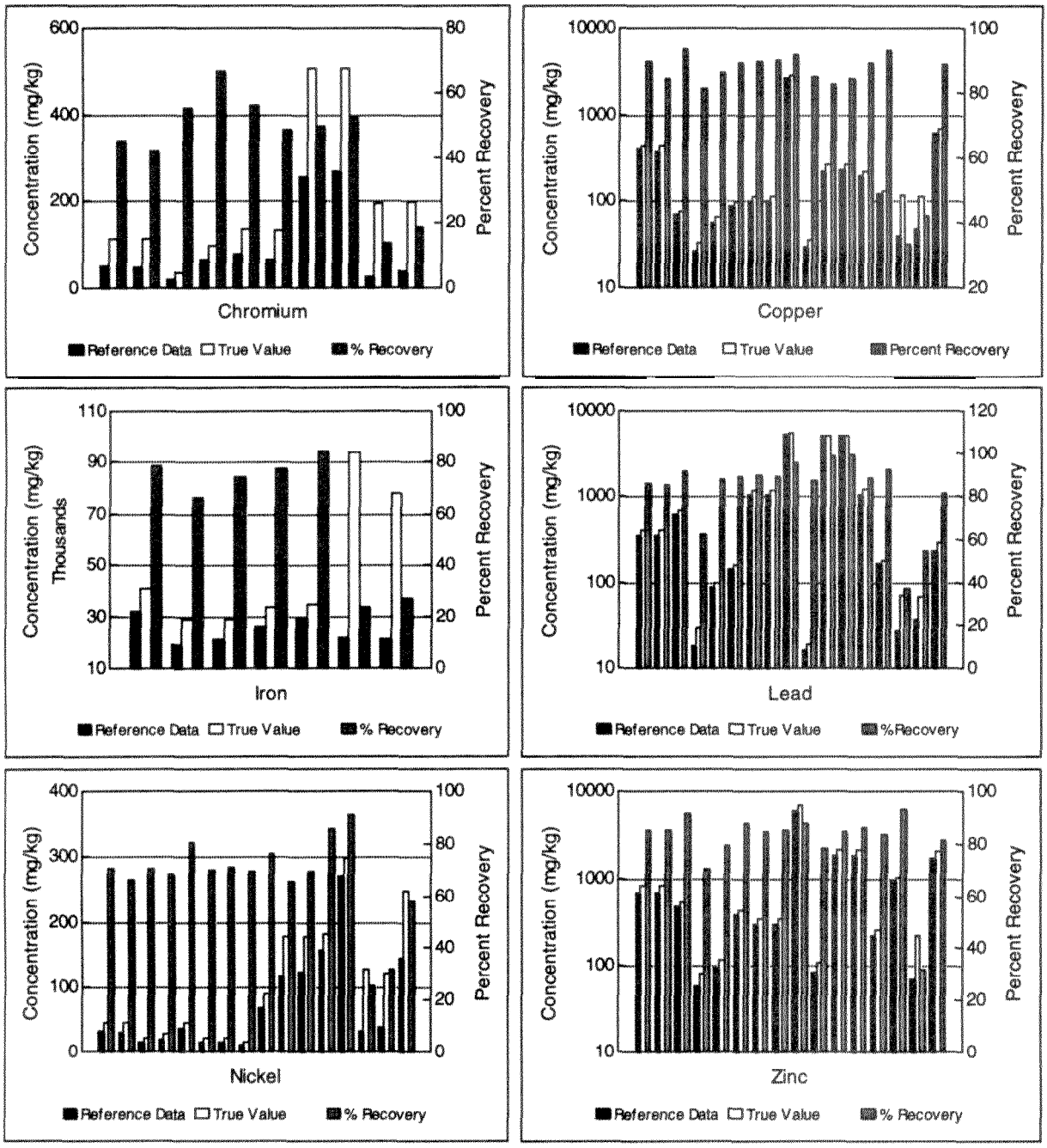


Figure 3-3 (Continued). Reference Method SRM Results: These graphs illustrate the relationship between the reference data and the true values for the SRM samples. The gray bars represent the percent recovery for the reference data. Each set of three bars (black, white, and gray) represents a single SRM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

Section 4
MAP Spectrum Analyzer

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

Theory of FPXRF Analysis

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

Because each element has a unique electron shell configuration, each will emit unique X-rays at set energies 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 characteristic X-rays emitted is proportional to the concentration of a given element and can be used to quantitate each target element.

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

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

the K-lines, the L-lines (L_α and L_β) for an analyte are of nearly equal intensity. The choice of which one to use for analysis depends on the presence of interfering lines from other analytes.

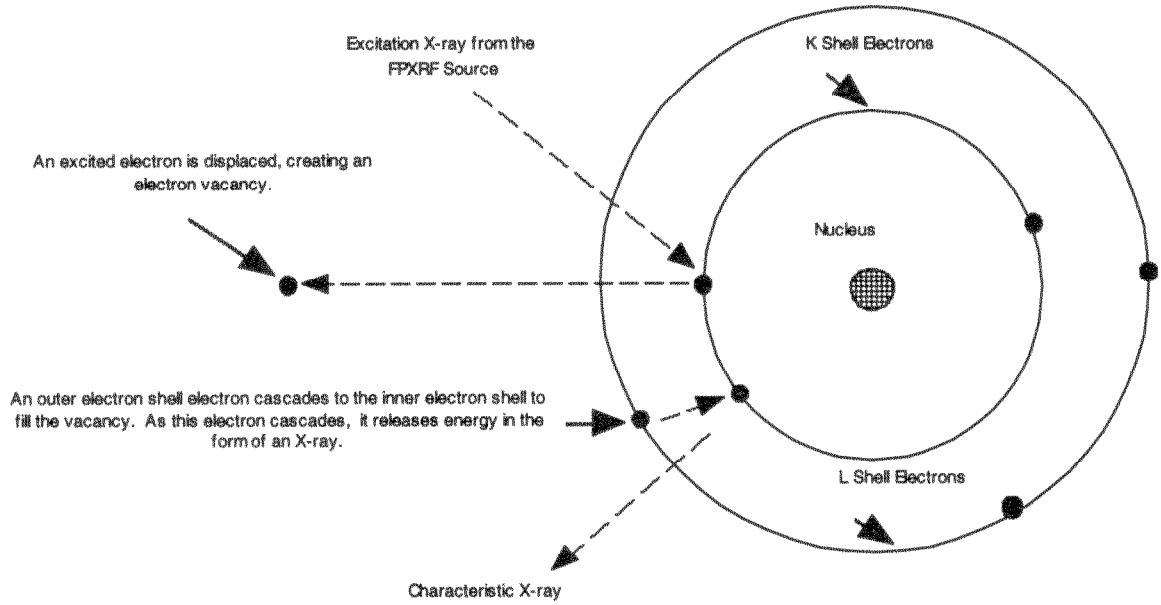


Figure 4-1. Principle of Source Excited X-ray Fluorescence: This figure illustrates the dynamics of source excited X-ray fluorescence.

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

The MAP Spectrum Analyzer was originally developed by Scitec to detect lead in paint using a cobalt-57 (Co57) excitation source. It is a lightweight, portable technology that collects ***in situ*** readings by placing the scanner in direct contact with the surface to be measured. Scitec currently markets the MAP Spectrum Analyzer as capable of detecting lead as well as other metals in soil, when equipped with a Cd^{109} source. The analyzer uses energy dispersive XRF spectroscopy to determine elemental composition of paint, soil, and other solid materials. The specific analyzer demonstrated during this evaluation was a third generation analyzer known as the MAP 3 Spectrum Analyzer. Since this demonstration was completed, Scitec has produced a fourth generation analyzer known as the MAP 4 Spectrum Analyzer.

Operational Characteristics

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

Equipment and Accessories

The primary components of the MAP Spectrum Analyzer are the control console and the ambient scanner. The control console is connected to the ambient scanner with a lo-foot cable. The basic system also includes a carry pack, rechargeable batteries, battery charger, operator's manual, site-specific standard, and a shipping case. For this demonstration, the scanner, control console, battery charger, and cords were contained in a carrying case. Additional equipment, such as the calibration check standards and spare batteries, did not fit in the carrying case and were shipped in a separate box. Specifications for the MAP Spectrum Analyzer used during this demonstration are provided in Table 4-l.

The equipment used in the demonstration included:

- One MAP 3 control console calibrated to detect arsenic, lead, copper, and zinc
- One MAP 3 scanner including a 55 millicuries (mCi) Cd^{109} sealed source
- Six 12-volt direct current (DC) lead-gel batteries (the console requires two batteries leaving two sets of two as spares)
- Two battery chargers
- One clip adapter for charging two batteries outside the console
- Two 10-foot cables for attaching the scanner to the console (one for use and one spare)
- One cable for connecting the console to a personal computer (PC)
- One computer port adapter
- AcuTransfer software
- One ring stand with clamps (the ring stand was used to hold the scanner in place to allow the calibration check standards to be assayed and to hold the scanner on slopes in the field)
- Three calibration check standards in 4-ounce jars
- One painted wooden block check standard
- One carrying case
- Plastic wrap used to cover the face of the scanner
- One portable computer and printer supplied by PRC

Table 4-1. Analyzer Instrument Specifications

Characteristic	Specification
Resolution	1.5 keV (Manganese-K_α)
Source	55 mCi Cd^{109} (Am^{241} and Co^{57} also available)
Detector	Solid-state silicon
Scanner Size	33.7 cm long
Scanner Weight	1.6 kilograms (kg) (3.5 pounds)
Scanner Operating Temperature	-6 to 43 °C
Control Console Size	19.3 cm x 20 cm x 7.6 cm
Control Console Weight	5 kg with batteries (11 pounds)
Control Console Operating Temperature	-6 to 43 °C
Control Console Data Acquisition	256 MCA
Control Console Data Storage Capacity	1 megabyte or 325 spectra and analyses
Power Source	120-volt or 220-volt alternating current, or 12-volt DC rechargeable batteries
Operational Checks	Calibration check sample once per hour
Contact: Bill Boyce or Kevin Dorow 415 N. Quay Kennewick, WA 99336 (800) 466-5323 FAX: (509) 735-9696	

The control console is a 256-multichannel analyzer (MCA) contained in a high-impact plastic case. It has a liquid crystal display (LCD) that can provide readouts of operation menus, measurement values, calibration menus, count rates, time clock, analysis identification numbers, and a graphic spectrum display. The keyboard is weatherproof and has a 14-key keypad. The two lead-gel type batteries necessary to supply power to the control console are capable of 10 hours of continuous use without recharging. Each lead-gel battery has an approximate useful life of 12 - 18 months or 150 recharges. The operator noted during this demonstration that the battery life ranged from 6.3 hours to 8.9 hours with an average of 7.7 hours. The control console also has an output port for downloading data to a PC with the use of the AcuTransfer software.

The ambient scanner is shaped like a pistol and contains the excitation source and the solid-state silicon detector. It has a 0.5-mm-thick beryllium window and a 0.5-mm-thick aluminum front face plate. The source shutter is constructed of tungsten and is designed to house one of three sources: cobalt-57 (Co57), cadmium-109 (Cd^{109}), or americium-241 (Am^{241}). The Cd^{109} source was used in this demonstration. The Cd^{109} source was assayed as 55 mCi on September 24, 1993. Based on a half life of 462 days, the source was 24mCi at the beginning of the demonstration and was 23 mCi at the end. The ambient scanner contained a solid-state silicon detector with a resolution of 1.5 keV at the K_a manganese line. It has a **Breech-lok™** connector that connects the scanner to the control console via a l0-foot cable.

Other equipment and supplies that are helpful when using the MAP Spectrum Analyzer, but are not supplied by the developer, include paper towels, protective gloves, a marking pen, an umbrella to shield the control console and scanner from rain, and lead foil to shield the operator from radiation during the calibration checks.

Operation of the Analyzer

Analysis with the MAP Spectrum Analyzer requires placing the scanner in direct contact with the sampling medium and opening the shutter with a key. The shutter exposes the sample to X-rays from the radioisotope source. Emission X-rays are then counted (measured) over an operator-specified period of time (source exposure time) by a counting circuit. This data is recorded by the MCA and produces a spectrum characteristic of the metals in the sample. The intensities for each target analyte are calculated by software deconvolution of the characteristic spectra and converted to concentration values by means of a calibration model. This model is derived empirically by measuring the intensities of the target analytes in a set of calibration standards and fitting a linear function that relates these values to concentration by a multiple regression procedure. The MAP Spectrum Analyzer measures a surface area of about 20 mm in diameter.

An empirical calibration of the MAP Spectrum Analyzer was performed by the developer prior to the demonstration using the predemonstration soil samples as site-specific calibration standards (SSCS). Calibration involved measuring the SSCSs and incorporating the data from the resultant spectra into a mathematical function. This function, which is a component of proprietary software, is used to calculate concentrations of the target analytes in the field samples.

Scitec states that to minimize enhancement or adsorption and spectral interference errors, calibration standards should be collected from the specific site in question. The SSCSs should closely match the matrix of the routine samples. Scitec recommends that characterization of the SSCSs be done by using a total digestion procedure, rather than a partial extraction because X-ray fluorescence is most closely related to a total extraction or digestion-type analysis. However, for this demonstration, the concentration of analytes in the SSCSs was determined using EPA SW-846 Methods 3050A/6010A because these were the methods used for the reference method analyses.

The in situ analysis with the MAP Spectrum Analyzer does not require that a sample be physically removed from the ground. The probe is placed on the ground and the analysis mode is activated by turning on a key. Acquisition time can be preset at any desired length: "Screen," "Test," or "Confirm" is the most common. The measurement times for the three options are 15 seconds, 60 seconds, and 240 seconds, respectively. In this demonstration, the "confirm" assay with a 240-second count time was used. These times are automatically corrected to account for the age of the source. Scitec points out that the precision of the analysis will improve as the measurement time increases.

The operator found the MAP Spectrum Analyzer easy to use. The console has only 14 keys and prompts the operator through the steps necessary to conduct each assay. The scanner and console are relatively lightweight. The console is usually equipped with a shoulder strap to keep the hands free to operate the scanner.

To operate the MAP Spectrum Analyzer, the operator set the console and scanner on the ground which allowed hands-free operation. On slopes, the operator either stabilized the scanner by securing the cord uphill or locking the scanner in the ring stand. The operator was able to maintain hands-free operation for all samples except two, which were on very steep slopes.

The console produced readings in ppm on the LCD for arsenic, lead, copper, and zinc. The readings appeared about 10 seconds after completing each assay and were automatically stored when the next assay was started. The MAP Spectrum Analyzer was operated by battery power while used outdoors, but it was connected to a battery charger while analyzing samples indoors. Downloading and printing data

was accomplished using the AcuTransfer software provided. Downloading and printing required just a few keystrokes on the console and computer keyboard.

QC procedures for the MAP Spectrum Analyzer included a calibration check sample. The calibration check sample was used to assess the accuracy of the technology. The calibration check was analyzed by placing the scanner in a specially built wooden jig with the scanner pointing up. A 4-ounce sample jar covered with a plastic wrap was then placed upside down over the source and scanner. The sample jar was filled with a soil sample collected during the predemonstration that was provided to Scitec. Scitec instructed the operator to conduct five confirmatory assays of the calibration check standard each morning and then one each hour during each day in the field. The five results each morning were averaged and compared to the average values determined in the factory calibration check performed by Scitec on this same soil sample. Scitec said that if the values differed by more than 250 ppm, the operator should contact the company. The operator found conducting calibration checks to also be relatively easy.

Background of the Technology Operator

The operator was an environmental scientist with more than 9 years experience in the environmental field. He earned a Master of Science degree from the University of Tulsa in 1986. He had worked at PRC for more than 3 years prior to the demonstration. While at PRC, he has managed and worked on many projects involving solid and hazardous waste and risk assessments.

Training

Training for safety and operation of the MAP Spectrum Analyzer was conducted by Scitec. The operator attended radiation safety training in St. Louis, Missouri, on December 15, 1994, and attended training on the operation of the MAP Spectrum Analyzer on April 3 and 4, 1995.

Radiation safety training was conducted at the AC Lead Testing and Training Center in St. Louis, Missouri. The training was adequate to address the level of exposure expected from the MAP Spectrum Analyzer. The operator received a certificate for the course.

The operator attended training to operate the MAP Spectrum Analyzer on April 3 and 4, 1995. This training was conducted at the Scitec facility in Kennewick, Washington. Mr. Kevin Dorow of Scitec conducted the training, which was sufficient to allow operation of the instrument under the conditions expected during the demonstration.

The training included a description of the calibration procedure and a hands-on demonstration of the process. The discussion of the AcuTransfer software included an extended session in the collection and downloading of data. The second day of training was dedicated to field use where a number of analyses were conducted and data were collected, as would be expected during the demonstration.

Reliability

Overall, 1,025 assays were conducted during this demonstration. This included 630 soil sample assays, 240 precision measurements, 145 calibration check assays, 3 PE sample assays, 2 SRM assays, and 5 blank assays. During the demonstration, there was frequent light to moderate rain while the FPXRF 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, no mechanical problems were experienced with the MAP Spectrum Analyzer. The only maintenance necessary was to periodically wipe the plastic covering on the face of the scanner or replace the plastic if it became too dirty to clean. The operator did encounter a few problems with the MAR Spectrum Analyzer that are discussed below.

After the third day of operation, the console "locked up" while the operator was reviewing data. The console would not come out of the "recall" mode. When this happened, the operator turned off the console, re-entered the data, and started the next assay. This appeared to correct the problem, no data were lost and the unit functioned normally for the remainder of the demonstration.

The operator experienced problems when downloading data on four occasions during the demonstration. On two attempts, the computer indicated data errors were detected. Both errors were solved by reinitiation of the download sequence. The third downloading problem was a failure to download all assay data without an indication of a problem on the computer screen. This could have been a significant problem resulting in lost data except that the MAP Spectrum Analyzer console did not return to the main menu as it should have. When the operator observed that the console did not return to the main menu, he initialized a new download again with success.

The fourth downloading problem occurred late in the demonstration. The computer indicated data errors in the transfer of assay data as occurred before. The operator retransmitted the data with apparent success. However, when he attempted to print the assay data, the values for all four metals were zero. The operator attempted to transfer data to the computer twice more with no success. The operator then exited and reentered the AcuTransfer software, turned off the printer and renamed the file, and attempted to download the data. Following this attempt, the computer indicated data errors were detected in the assay transfer. The operator attempted to transfer the assay data again this time with success. The source of the problem was not identified. The difficulties encountered in downloading resulted in about 40 minutes of lost time.

In the training provided by Scitec, the operator was instructed to call the manufacturer if the calibration check standard varied from the factory check by more than 250 ppm. Both lead and copper values varied by more than 250 ppm in the first set of five calibration check standards run in the demonstration. As instructed, the operator called Scitec regarding the discrepancies. Scitec thought the difference was due to heterogeneity in the calibration check standard used. Scitec told the operator to continue operating the MAP Spectrum Analyzer and that it would send other calibration check standards for the operator to run.

Data were sent two 4-ounce glass jars of ASARCO soil samples and one painted wooden block to be used as calibration check standards at the RV Hopkins site. The instructions were to use the wooden block as the primary calibration check standard. Scitec said to conduct five assays of the wooden block with the scanner placed upside down in the ring stand prior to any soil sample analysis. The developer stated that the check standard assay would be equivalent to the factory check test. The new standards meet the calibration requirements and data collection was resumed.

The most common operator error was forgetting to turn the key on the scanner to the "on" position prior to starting an assay. This occurred nine times during the evaluation. Each time this occurred, several minutes were lost restarting the test. However, the operator felt the safety considerations of the key outweighed the inconvenience of forgetting to turn the key.

The operator did not notice a low battery indication at any time during the demonstration. Each time the batteries died, the operator was in the process of running an assay. Therefore, each of those assays had to be restarted. The operator noted that if a low battery indicator had been observed, he would have changed batteries between assays to prevent the need for reanalysis. The operations manual for the MAP Spectrum Analyzer states that a "Low Batt" indicator will flash at the bottom right comer of the LCD approximately 45 minutes prior to complete battery discharge. It is possible the operator simply failed to notice the "Low Batt" indicator.

The operator observed that the MAP Spectrum Analyzer permitted identical sample numbers to be entered. In this demonstration, this feature was not a problem because the operator kept careful notes regarding sample numbers. However, if this were not the case, duplicate sample numbers could result in confusion.

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 radiation survey meter. Background radiation at the two sites was between 0.006 and 0.012 millirems per hour (mrem/hr). Radiation was monitored 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 in the field, a maximum radiation value of 1.2 mrem/hr at the probe face was obtained with the Cd^{109} source exposed. The radiation values dropped off to 0.40 mrem/hr at the key and 0.05 mrem/hr at the handle of the scanner. While taking ***in situ*** measurements indoors with the scanner pointed down at the sample, radiation values of 4.0 to 6.0 mrem/hr at the probe face were obtained with the Cd^{109} exposed. The radiation values dropped off to 0.10 - 0.20 mrem/hr at 2 feet from the probe face and were 0.07 to 0.08 mrem/hr at the scanner handle. The operator placed a lead shield and a row of bricks around the scanner while conducting the ***in situ*** measurements indoors. The radiation behind the lead shield and bricks was measured at 0.03 to 0.05 mrem/hr with the Cd^{109} exposed. With the exception of the radiation values right at the probe face, all radiation values were below the permissible 2.0 mrem/hr.

A greater radiation hazard was experienced while measuring the calibration check sample. In this mode, the scanner was pointed upward with the soil sample placed on top of the scanner. The scanner was held motionless in a ring stand. While analyzing the calibration check standard, radiation values of greater than 100 mrem/hr above the scanner and 10 - 20 mrern/hr at the side of the scanner were encountered. At head height, these radiation values dropped to 0.20 mrem/hr at 1 foot from the scanner and to 0.05 mrem/hr at 3 feet from the scanner.

Cost

At the time of this demonstration, the cost of a new MAP Spectrum Analyzer standard package was $32,000 with a Cd^{109} source. The standard package includes the control console, the ambient scanner, a Cd^{109} radioisotope source, auto source decay time correction, carry pack, rechargeable batteries, spectrum display software, 256-kilobyte memory, battery charger, operator's manual, shipping case, a l0-foot

cable, and a lead-check standard. Periodic maintenance includes replacement and disposal of the Cd^{109} source every 2 years. A new Cd^{109} source costs $6,000 with a disposal cost of $75. A wipe test must be performed every 6 months at a cost of $50. A replacement Co^{57} source costs $3,695. The long half-life of the Am^{241} source precludes the need for replacement. The MAP Spectrum Analyzer can be rented for $4,675 per month plus a $4,225 deposit.

A basic radiation safety and operator training course is offered by Scitec for $245 per person plus travel expenses. Costs to obtain the specific license for the MAP Spectrum Analyzer also were incurred for this demonstration. It cost $500 to obtain the license for ownership and operation of a sealed radioactive source in the State of Kansas. Since the demonstration sites were in Washington and Iowa, reciprocal agreements were required from both states to operate the instrument in those states. The reciprocal agreements cost $585 for Washington and $700 for Iowa. Operator training time may vary depending on the technical knowledge of the operator. Scitec claims the MAP Spectrum Analyzer can be used by individuals with no more than a high school education and a minimal amount of technical training.

The primary cost benefit of field analysis is the quick access to analytical data. This allows the process dependent on the testing to move efficiently onto the next stage. Costs associated with field analysis are 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 MAP Spectrum Analyzer 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 estimate can be derived from the data provided in this table.

Table 4-2. Instrument and Field Operation Costs

Item	Amount	
MAP Spectrum Analyzer	$ 32,000	Purchase Price
	4,675	Per Month Lease
Replacement Source	6,000	For Cd^{109}
	3,695	For Co^{57}
Operator Training (Vendor Provided)	245	—
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	9 - 12	Samples per hour
Cost of Reference Laboratory Analysis	150	Per sample

Performance Factors

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

Detection *Limits*

MDLs, using SW-846 protocols, were determined by collecting 10 replicate measurements on site-specific soil samples with metals concentrations 2 to 5 times the expected MDL value. 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 SD for each target analyte. All the precision-based MDLs were calculated for the measurements on the in situ-prepared soil samples. The precision-based MDLs for the MAP Spectrum Analyzer are shown in Table 4-3. The precision-based MDLs for all analytes were obtained using a 240-second count time for the Cd^{109} source.

Another method of determining MDLs involved the direct comparison of the FPXRF data and the reference data. When these sets of data were plotted against each other the resultant plots were linear. As the plotted line approached zero for either method, there was a point at which the FPXRF data appeared to respond to the same reading for decreasing concentrations of the reference data. Figure 4-2 illustrates this effect for zinc. 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 calculation. By determining the mean values of this FPXRF data and subsequently two SDs around this mean, it was possible to determine a field or performance-based MDL for the analyzer. For the MAP Spectrum Analyzer, these field-based MDLs also are shown in Table 4-3.

Table 4-3. Method Detection Limits

Analyte	Precision-based MDL (mg/kg)	Field-based MDL (mg/kg)
Arsenic	225	150
Copper	525	270
Lead	165	160
Zinc	25	180

Note: mg/kg Milligrams per kilogram.

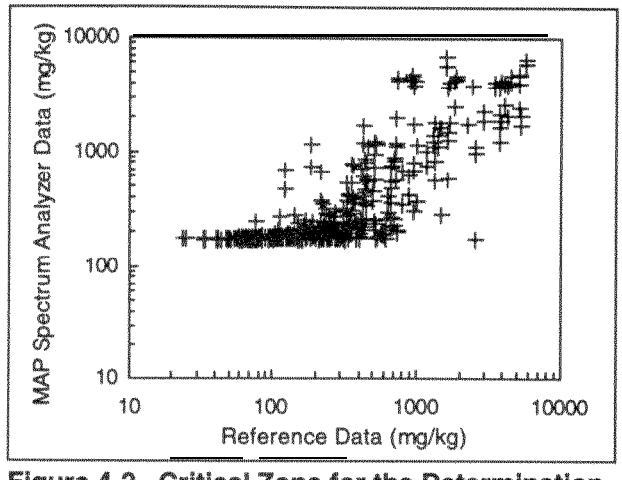


Figure 4-2. Critical Zone for the Determination of a Field-based Method Detection Limit for Zinc: At approximately 180 mg/kg, the linear relationship between the field and reference data changed. This point of change was used to determine the field-based MDLs.

The developer stated a detection limit of 250 mg/kg should be achievable for all four primary analytes. The field-based MDLs were close to or below 250 mg/kg. The precision-based MDLs showed much more variation between analytes than did the field-based MDLs. The precision-based and field-based MDLs were similar for arsenic and lead but different for copper and zinc. The high precision RSD

for copper caused its precision-based MDL to be large. Likewise, the extremely good precision for zinc caused its precision-based MDL to be very low. Given the nature of the detector in the MAP Spectrum Analyzer and based on recommendations by the developer, the field-based MDLs for copper and zinc appear more realistic than the precision-based MDLs.

Throughput

The MAP Spectrum Analyzer used a Cd^{109} source live count time of 240 seconds. With the additional "dead" time of the detector and the time required to label each sample and store data in between sample measurements, the time required to analyze one sample was between 5 and 7 minutes. The average number of samples analyzed was 98 in an 1 l-hour day for a throughput of 8.9 samples per hour throughout the demonstration. The minimum number of samples analyzed was on the first day at ASARCO when 60 samples were analyzed in 9 hours for a throughput of 6.7 samples per hour. As the operator became more familiar with the analyzer, the throughput increased. The maximum number of samples analyzed was 140 in 12 hours at the ASARCO site for a throughput of 11.7 samples per hour. This throughput was achieved while analyzing the in situ-prepared samples indoors.

This throughput included the time necessary to analyze the QC samples, which included five assays of the calibration check, and the subsequent hourly analysis of the calibration check each day. The throughput did not include the time required for sample handling and preparation or for data downloading, printing, and documentation. Data handling required approximately 30 minutes each day. Homogenization for the *in* situ-prepared samples required approximately **5** minutes per sample.

Drift

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

The calibration check was analyzed five times each morning and once per hour each day of analysis. The drift summary is displayed for the four analytes in Figure 4-3. Each box on the figure represents the mean performance for a given analyte for a given day. The drift values were standardized by taking the mean for all calibration check sample measurements and then finding the percent difference between this overall mean and the daily mean concentration. Figure 4-3 shows that the MAP Spectrum Analyzer showed the most drift for copper and the least for zinc. These drift results mimic the reproducibility displayed by the high and low precision-based MDLs for copper and zinc, respectively. The copper drift varied from -25 to +35 percent, while the zinc drift remained between ±5 percent. Arsenic drift was within ±15 percent, and the majority of the lead drift was within ±20 percent.

Intramethod Assessment

Intramethod assessment measures each analyzer's performance on characteristics such as: analyzer blanks, completeness of the data set, intramethod precision, and intramethod accuracy. The following narrative discusses these characteristics.

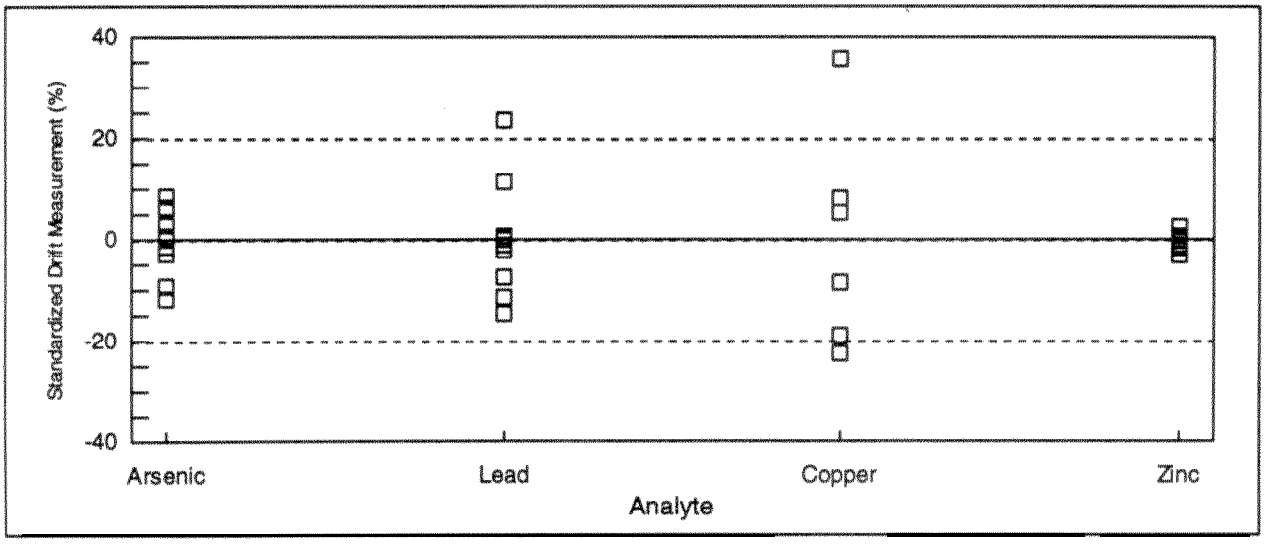


Figure 4-3. Drift Summary: This figure shows the general drift of the analyzer's results in measuring a check sample. Each point represents a different day's analysis of the same sample. The daily fluctuations exhibited for each analyte are a direct representation of drift.

Blanks

Analysis blanks for the MAP Spectrum Analyzer were obtained by shooting ambient air with the scanner. The blanks were used to monitor contamination of the scanner by material such as soil left on the scanner face. Four blanks were analyzed during the demonstration, one using a 240-second count time, and three using a 60-second count time. The results for all four blanks were similar. The blank values for arsenic and lead were all below their precision and field-based MDLs. The zinc blank values ranged from 176 to 179 mg/kg. These values were slightly below the field-based MDL for zinc but well above the precision-based MDL for zinc of 25 mg/kg. The blank values for copper ranged from 477 to 673 mg/kg which are above the precision and field-based copper MDLs. These copper results were surprising because the MAP Spectrum Analyzer gave copper values of 0 - 400 mg/kg for many of the soil samples, so it is not believed that cross contamination caused the high blank results for copper. It may be an artifact of using air as the blank matrix instead of a clean silica sand, which is more similar to a soil matrix.

Completeness

The MAP Spectrum Analyzer produced data for 628 out of the 630 samples for a completeness of 99.7 percent, which is above the demonstration objective of 95 percent. The two samples for which no data were obtained were from the ASARCO site. In one case the operator failed to analyze the sample, For the other sample, it appears a software malfunction in the downloading process caused the loss of data. The lack of data was not caused by a mechanical or an electronic malfunction of the analyzer.

Precision

Precision was expressed in terms of the percent RSD between replicate measurements. 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.

Table 4-4. Precision Summary

Mean % RSD Values by Concentration Range				
Analyte	5 - 10 Times MDL[a] (mg/kg)	50 - 500 (mg/kg)	500 - 1,000 (mg/kg)	>1,000 (mg/kg)
Arsenic	6.68 (2)	31.36 (4)	20.39 (8)	6.68 (2)
Copper	14.86 (4)	ND	23.48 (2)	14.86 (4)
Lead	8.54 (6)	19.03 (2)	6.71 (2)	6.33 (10)
Zinc	0.64 (18)	0.77 (22)	1.52 (2)	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. Numbers do not always add up to 24 precision points because some samples had analyte concentrations below the analyzer's MDL.

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 two ***in*** situ sample preparation steps. Therefore, there was a total of 24 precision points for the analyzer. 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 replicate determinations in the 5 to 10 times MDL range, were below the 10 percent RSD required for definitive level data quality classification for arsenic, lead, and zinc. Copper had an RSD between 10 and 20 percent, placing its precision into the quantitative screening level data quality category.

The precision data in Table 4-4 shows there was an effect of concentration on the precision for arsenic, copper, and lead. The precision samples were purposely chosen to span a large concentration range to test the effect of analyte concentration on precision. As expected, the precision increased as analyte concentration increased for arsenic, copper, and lead. The zinc precision changed little with differing zinc concentrations and did not show the same trend as the other three analytes, possibly because the zinc precision was so much better than the other three analytes that it was difficult to observe changes. There was no observable effect of sample preparation on precision. This was expected because the method used to assess precision during this demonstration was measuring analyzer precision, not total method precision.

Accuracy

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

Intramethod accuracy was assessed for the MAP Spectrum Analyzer by using three site-specific PE samples and two NIST SRMs. There were 12 other SRMs and 3 other site-specific PEs included in this demonstration, but there was not enough material for these other PEs and SRMs to fill a petri dish for

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

Sample	Arsenic			Copper			Lead			Zinc		
	True Conc.	Meas. Conc.	% Rec.	True Conc.	Meas. Conc.	% Rec.	True Conc.	Meas. Conc.	% Rec.	True Conc.	Meas. Conc.	% Rec.
ASARCO Low PE	419	227	54.2	771	108	14.1	292	0	0.0	164	180	109.8
ASARCO Med. PE	837	497	59.4	5,408	2,003	37.0	1,012	963	95.2	378	210	55.6
ASARCO High PE	22,444	21,275	94.8	7,132	4,545	63.7	9,498	10,049	105.8	4,205	4,499	107.0
NIST SRM 2704	23.4	56	NA	98.6	49	NA	161	0	NA	438	175	39.9
NIST SRM 2710	626	799	127.6	2,950	3,740	126.8	5,532	4,175	75.5	6,952	386	5.5
Mean % Recovery			84.0			60.4			92.1			63.6
SD of % Recovery			34.2			48.7			15.4			44.7
% Within 80 to 120 Percent Recovery Acceptance Range			25.0			0.0			50.0			40.0

Notes: % Rec. Percent recovery.

PE Performance evaluation sample.

SRM Standard reference material.

NA Not applicable. Percent recovery not calculated because true concentration was below the method detection limit.

SD Standard deviation.

For this demonstration, the MAP Spectrum Analyzer was configured to report concentrations for lead, arsenic, copper, and zinc. The regression analysis of the entire $\log_{10}$ transformed data set for these target analytes showed that lead had an r^2 value of 0.85. The corresponding value for copper was 0.80, while arsenic and zinc had r^2 values of 0.76 and 0.67, respectively. The arsenic and copper comparability was biased low for the whole data set. The concentrations of these analytes at the RV Hopkins site were generally at or below the analyzer's precision-based MDLs.

The next step in the data evaluation involved the assessment of the potential impact of the variables: site, soil texture, and sample preparation on the regression analysis (Table 4-6).

Based on this evaluation, there was no apparent impact of the site variable on the regression. The soil variable showed a slight decreasing trend in comparability with the sand soils exhibiting the highest comparability and the loam soils exhibiting the lowest comparability. The only exception to this soil trend was for lead. Lead met definitive data quality criteria for the clay soil' s entire data set. This increased comparability for the clay soil data may be due to the origin of the lead contamination. The clay soils were only analyzed at the RV Hopkins site and the source of lead contamination at this site was primarily paint waste. The paint matrix is what this analyzer was originally designed to analyze. The slope values for the soil variable data indicated that the analyzer tended to underestimate lead and copper concentrations and overestimate zinc concentrations. The slope values were determined by plotting the FPXRF data on the x-axis (independent variable) and the reference method data on the y-axis (dependent variable). The data were plotted in this fashion to get an indication of a correction factor to be applied to the FPXRF data to get it to match the reference data.

Table 4-6. Regression Parameters[a] by Primary Variable

Arsenic					Variable	Lead				
n	r^2	Std. Err.	Y-Int.	Slope		n	r^2	Std. Err.	Y-Int.	Slope
449	0.763	0.37	0.62	0.79	All Data	368	0.849	0.28	-0.74	1.21
398	0.897	0.23	-0.03	1.00	ASARCO Site	210	0.865	0.28	-1.22	1.36
55	0.049	0.64	1.43	0.73	RV Hopkins Site	158	0.847	0.24	-0.08	1.02
181	0.925	0.23	-0.21	1.03	Sand Soil	90	0.914	0.26	-1.15	1.33
219	0.871	0.23	0.13	0.97	Loam Soil	119	0.798	0.29	-1.38	1.42
55	0.049	0.64	1.43	0.73	Clay Soil	158	0.847	0.24	-0.08	1.02
230	0.624	0.45	1.02	0.65	In Situ-Unprepared	186	0.770	0.36	-0.66	1.19
219	0.900	0.25	0.20	0.93	In Situ-Prepared	184	0.916	0.20	-0.90	1.26

Copper					Variable	Zinc				
n	r^2	Std. Err.	Y-Int.	Slope		n	r^2	Std. Err.	Y-Int.	Slope
251	0.801	0.36	-0.71	1.13	All Data	613	0.669	0.20	1.05	0.59
232	0.851	0.30	-1.48	1.34	ASARCO Site	420	0.650	0.20	1.03	0.60
20	0.034	0.73	1.13	0.61	RV Hopkins Site	199	0.634	0.26	0.97	0.64
62	0.938	0.19	-2.25	1.57	Sand Soil	195	0.746	0.21	0.95	0.66
172	0.807	0.33	-1.13	1.25	Loam Soil	224	0.457	0.18	1.27	0.48
20	0.034	0.73	1.13	0.61	Clay Soil	199	0.634	0.26	0.97	0.64
125	0.761	0.40	-1.00	1.22	In Situ-Unprepared	305	0.654	0.20	1.10	0.57
126	0.846	0.32	-0.50	1.07	In Situ-Prepared	308	0.673	0.21	1.01	0.61

Notes:

[a] Regression parameters based on $\log_{10}$ transformed data. Since the FPXRF data were used as the dependent variable in calculating these regression parameters, the regression data must be used to correct the FPXRF data. See Section 5.

Y-Int. Y-intercept.

Std. Err. Standard error.

n Number of data points.

The sample preparation variable exhibited the greatest impact on the regression analysis. This sample preparation effect makes sense since the homogenization step assured that the analyzer and the reference methods were analyzing essentially the same sample. The initial sample homogenization (in situ-prepared) improved the comparability for arsenic between the two data sets to the point that the analyzer met the definitive level criteria. The analyzer's lead and copper data met quantitative screening level quality criteria at the initial sample preparation step. The sample homogenization improved the r^2 for copper to 0.85, raising it to the definitive level of data quality. Increasing sample preparation increased comparability; however, the data for zinc never met quantitative screening level quality criteria because the r^2 values for zinc remained below 0.70.

The impact of the site and soil texture variables was then assessed for each of the two sample preparation steps (Tables 4-7 and 4-8). To simplify the analysis, this evaluation was only conducted for lead which exhibited a relatively even concentration distribution between the site and soil variables. No clear effect on comparability was observed for the site variables; however, the soil variable reflected the lowest comparability for the loam soil.

Within the sample preparation steps, the effect of contaminant concentration was also examined. The data sets for the analytes were sorted into the following concentrations ranges: 0 - 100 mg/kg, 100-1,000 mg/kg, and greater than 1,000 mg/kg. The regression analysis for each target analyte and for each sample preparation step was rerun on these concentration-sorted data sets. A review of these results showed general improvement in the r^2 and standard error 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 below the analyzer's MDLs. The analyzer's precision and accuracy are lowest in this concentration range. Generally, the r^2s improved between the 100 and 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.

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 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), or only narrow concentration ranges can be analyzed, or average relative accuracy can be examined. 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 derived from this demonstration involves an assessment of both r^2 and the precision RSD. Using the criteria presented in Table 2-2, a summary of the MAP Spectrum Analyzer's data quality performance in this demonstration is provided in Table 4-9.

Table 4-7. Regression Parameters[a] for the Sample Preparation Variable Sorted by Soil Texture

Arsenic						Lead				
n	r^2	Std. Err.	Y-Int.	Slope	Soil Texture	n	r^2	Std. Err.	Y-Int.	Slope
In Situ-Unprepared						In Situ-Unprepared				
89	0.895	0.26	-0.08	0.98	Sand Soil	46	0.882	0.30	-0.86	1.23
110	0.799	0.28	0.30	0.92	Loam Soil	59	0.731	0.36	-1.41	1.42
32	0.164	0.59	1.02	1.31	Clay Soil	78	0.762	0.27	0.37	0.89
In Situ-Prepared					Soil Texture	In Situ-Prepared				
92	0.955	0.18	-0.34	1.08	Sand Soil	44	0.945	0.21	-1.33	1.39
111	0.918	0.18	0.14	0.96	Loam Soil	57	0.915	0.15	-1.24	1.39
23	0.009	0.66	1.64	0.31	Clay Soil	81	0.923	0.18	-0.47	1.13

Copper						Zinc				
n	r^2	Std. Err.	Y-Int.	Slope	Soil Texture	n	r^2	Std. Err.	Y-Int.	Slope
In Situ-Unprepared						In Situ-Unprepared				
33	0.940	0.20	-2.04	1.50	Sand Soil	98	0.743	0.20	1.03	0.62
89	0.746	0.40	-0.86	1.19	Loam Soil	110	0.340	0.17	1.52	0.36
4	0.868	0.25	10.78	-3.71	Clay Soil	93	0.750	0.17	1.07	0.60
In Situ-Prepared					Soil Texture	In Situ-Prepared				
28	0.959	0.13	-2.31	1.61	Sand Soil	98	0.738	0.23	0.87	0.70
82	0.927	0.19	-1.48	1.32	Loam Soil	114	0.550	0.19	1.09	0.56
15	0.161	0.51	0.29	0.89	Clay Soil					

Notes:

[a] Regression parameters based on $\log_{10}$ transformed data. Since the FPXRF data were used as the dependent variable in calculating these regression parameters, the regression data must be used to correct the FPXRF data. See Section 5.

Y-Int. Y-intercept.

Std. Err. Standard error.

n Number of data points.

Table 4-8. Regression Parameters[a] for the Sample Preparation Variable Sorted by Site Name

Arsenic						Lead				
n	r^2	Std. Err.	Y-Int.	Slope[b]		n	r^2	Std. Err.	Y-Int.	Slope[b]
In Situ-Unprepared					Site Name	In Situ-Unprepared				
195	0.861	0.27	0.08	0.97	ASARCO Site	106	0.807	0.34	-1.09	1.31
32	0.164	0.59	1.02	1.31	RV Hopkins Site	78	0.762	0.27	0.37	0.89
In Situ-Prepared					Site Name	In Situ-Prepared				
203	0.931	0.19	-0.12	1.03	ASARCO Site	102	0.931	0.19	-1.38	1.43
23	0.009	0.66	1.64	0.31	RV Hopkins Site	81	0.923	0.18	-0.47	1.13

Copper						Zinc				
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				
122	0.799	0.37	-1.24	1.29	ASARCO Site	209	0.625	0.20	1.11	0.56
4	0.868	0.25	10.78	-3.71	RV Hopkins Site	93	0.750	0.17	1.07	0.60
In Situ-Prepared					Site Name	In Situ-Prepared				
111	0.917	0.21	-1.57	1.35	ASARCO Site	210	0.678	0.20	0.96	0.63
15	0.161	0.51	0.29	0.89	RV Hopkins Site	99	0.619	0.25	1.02	0.61

Notes:

[a] Regression parameters based on $\log_{10}$ transformed data. Since the FPXRF data were used as the dependent variable in calculating these regression parameters, the regression data must be used to correct the 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.

Y-Int. Y-intercept.

Std. Err. Standard error.

n Number of data points.

Table 4-9. Summary of Data Quality Level Parameters

Target Analytes	MAP Spectrum Analytes	Precision (mg/kg) Mean % RSD 5 - 10 X MDL	Method Detection Limits (mg/kg) (Precision-based)	Coefficient of Determination (r^2 All Data)	Data Quality Level
Arsenic	Arsenic	6.68	225	0.763	Quantitative
Barium	Not Reported	—	—	—	—
Chromium	Not Reported	—	—	—	—
Copper	Copper	14.86	525	0.801	Quantitative
Lead	Lead	8.54	165	0.849	Definitive
Zinc	Zinc	0.64	25	0.669	Qualitative
Nickel	Not Reported	—	—	—	—
Iron	Not Reported	—	—	—	—
Cadmium	Not Reported	—	—	—	—
Antimony	Not Reported	—	—	—	—

Section 5
Applications Assessment and Considerations

The MAP Spectrum Analyzer is designed to analyze for metals in soils, sludges, and other solids. The analyzer uses an empirical site-specific calibration and quantitation procedure to maximize its performance. This calibration accounts for common soil-related matrix interferences. This analyzer is designed for field use in the ***in situ*** mode. The analyzer experienced no hardware failures during this demonstration and the few software malfunctions resulted in little downtime and no lost data during the l-month field demonstration. During this time, more than 630 samples were measured by the analyzer. The training provided by the developer was sufficient to allow basic field operation. Limited developer assistance was required to address the software problems encountered during the demonstration. The developer provided accessible and timely field support. The use of this analyzer requires specific radiation licensing, which adds some cost and training to the use of this analyzer.

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 lead. This indicated that the analyzer' s data were statistically equivalent to the reference data for these analytes. For arsenic and copper, the analyzer produced quantitative screening level data. In addition, this analyzer exhibited instrument precision similar to the reference methods, indicating high measurement reproducibility. The analyzer produced zinc data which met the qualitative screening level data quality criteria. A summary of key operational features is listed in Table 5-l

The analyzer' s probe uses one radioactive source allowing analysis of a limited number of metals in soils. The analyzer used count times of 240 live-seconds. Longer count times generally increase accuracy and lower the detection limits but decrease sample throughput. The throughput for the analyzer was 9 - 12 samples per hour. There were no apparent effects of site or soil texture on performance for any of the analytes; however, lead data did show its highest comparability for the RV Hopkins samples, which were clay soils. This may be due to the fact that the lead in these soils was derived from paint waste, a matrix for which this instrument was originally designed. This demonstration identified sample preparation as the most important variable with regard to analyzer performance.

The analyzer can be applied only in an ***in situ*** mode. The data from this demonstration indicated that when operated in the ***in*** situ-unprepared mode, the results did not show a strong correlation between FPXRF and reference data. This may not be due to instrument error but rather to inherent spatial variability of contamination, even within an area as small as the 4-inch by 4-inch grid sampled during this demonstration. The greatest increase in correlation between the FPXRF data and reference data for the analyzer was achieved after the initial sample preparation step (sample homogenization), which defined the ***in*** situ-prepared sample set.

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

Total weight less than 15 pounds, battery life of 8 hours
Sample throughput of 9 to 12 samples per hour at 240 live-second count times
In situ measurements only
Rugged and reliable - Data completeness of 99.7 percent
Operation requires minimal training
Produces EPA quantitative screening level data for arsenic and copper and EPA definitive level data for lead
Empirical calibration is site-specific
Precision - Percent RSD values less than 15 percent at 5 to 10 times the MDL for all reported analytes
Generally not susceptible to soil matrix effects
Can be used on soils exhibiting up to 30 percent water saturation by weight
A single source limits the number of elements that can be quantified
Empirical calibration requires well characterized site-specific samples
Possible radiation hazard when performing calibration checks with the scanner pointed upward
Produced EPA qualitative screening level data for zinc

Based on this demonstration, the analyzer is well suited for the rapid real-time assessment of metals contamination in soil samples. The ease of operation and minimal training requirements increases the probability that a first-time user will produce reliable data. Although in most cases the analyzer produced data statistically equivalent to the reference data, generally confirmatory analysis will be required or requested for FPXRF analysis. If 10 - 20 percent of the samples measured by the analyzer are submitted for reference method analysis, instrument bias relative to standard methods such as 3050A/60l0A 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. Bias correction allows most FPXRF data to be corrected so that it more closely matches the reference data. The demonstration showed that the analyzer exhibits a strong $\log_{10}$-$\log_{10}$ linear relationship with the reference data over a concentration range of 5 orders of magnitude. A concentration effect on comparability was noted for this analyzer. Measurements near or below the analyzer' s MDLs showed the poorest comparability. As concentrations rise above the MDLs, the data comparability increases. This should be taken into consideration when evaluating the usability of field-generated data. For optimum correlation and bias correction, samples with high, medium, and low concentration ranges from a project should be submitted for reference method analysis.

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

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

5. Conduct a linear regression analysis and determine the r^2, y-intercept and slope of the relationship. The r^2 must be greater than 0.70 to proceed.
6. Place the regression parameters into Equation 5-1:

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

7. Use the above equation with the $\log_{10}$ transformed FPXRF results from Step 4 above and calculate the equivalent $\log_{10}$ corrected FPXRF data.
8. Take the anti-$\log_{10}$ ($10^{[\log_{10}\ transformed\ corrected\ FPXRF\ data]}$) 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 different from 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 are generally small. Alternately, analytes falling into the quantitative and qualitative screening level data quality categories generally have larger average relative bias and accuracy.

In cases where the corrected average relative accuracy is worse than the raw average relative accuracy, such as seen in Table 5-2 for arsenic, the eight-step FPXRF data correction approach presented earlier may not be appropriate. If the data set in question is representative of the entire population of data being characterized, then the raw FPXRF data merely needs to be multiplied by the raw average relative accuracy factor for correction. However, the eight-step regression base approach should be used anytime 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 Scitec MAP Spectrum Analyzer can provide 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 development of Method 6200 will help in the acceptance of FPXRF data for all definitive level applications and most quantitative 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.

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 Accuracy for PE Samples[e]
Arsenic	1.06	1.13	2.19	2.24	1.76
Copper	0.71	1.29	1.73	2.41	1.18
Lead	0.93	1.06	1.39	1.35	1.63
Zinc	1.56	1.23	2.23	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:

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

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:

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

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.

General Operational Guidance

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

General operation of FPXRF instruments will vary according to specific developer protocols. For all environmental applications, confirmatory or reference sampling should be conducted so that **FPXRF** data can be corrected. Before operating any FPXRF instrument, the developer's manual should be consulted.

Most developers recommend that their instruments be allowed to warm up for 15 - 30 minutes before analysis of samples. This will help alleviate drift or energy calibration problems.

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

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

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

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

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

After a sample is prepared, a portion of the sample should then be placed in a 3l-mm polyethylene sample cup (or equivalent) for analysis. The sample cup should be completely filled. The sample cup should be covered with a 2.5-micrometer MylarTM (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 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 Supetfund-Interim Final Guidance." Office of Solid Waste and Emergency Response. Washington, D.C. EPA/540/R-93/071.