



EPA

Innovative Technology Verification Report

Field Measurement Technology for Mercury in Soil and Sediment

Milestone Inc.'s Direct Mercury Analyzer (DMA)-80



EPA/600/R-04/012
May 2004

Innovative Technology Verification Report

Milestone Inc.'s Direct Mercury Analyzer (DMA)-80

Prepared by

Science Applications International Corporation
Idaho Falls, ID

Contract No. 68-C-00-179

Dr. Stephen Billets
Characterization and Monitoring Branch
Environmental Sciences Division
Las Vegas, Nevada 89193-3478

National Exposure Research Laboratory
Office of Research and Development
U.S. Environmental Protection Agency

Notice

The U.S. Environmental Protection Agency through its Office of Research and Development funded and managed the research described here under contract to Science Applications International Corporation. It has been subjected to the Agency's peer and administrative review and has been approved for publication as an EPA document. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

Office of Research and Development
Washington, DC 20460

MEASUREMENT AND MONITORING TECHNOLOGY PROGRAM VERIFICATION STATEMENT

TECHNOLOGY TYPE: **Field Measurement Device**

APPLICATION: **Measurement for Mercury**

TECHNOLOGY NAME: **Milestone Inc.'s Direct Mercury Analyzer (DMA)-80**

COMPANY: **Milestone Inc.**
ADDRESS: **160B Shelton Rd.**
Monroe, CT 06468

WEB SITE: **<http://www.milestonesci.com>**

TELEPHONE: **(203) 261-6175**

VERIFICATION PROGRAM DESCRIPTION

The U.S. Environmental Protection Agency (EPA) created the Superfund Innovative Technology Evaluation (SITE) and Measurement and Monitoring Technology (MMT) Programs to facilitate deployment of innovative technologies through performance verification and information dissemination. The goal of these programs is to further environmental protection by substantially accelerating the acceptance and use of improved and cost-effective technologies. These programs assist and inform those involved in design, distribution, permitting, and purchase of environmental technologies. This document summarizes results of a demonstration of the Direct Mercury Analyzer (DMA)-80 developed by Milestone Inc.

PROGRAM OPERATION

Under the SITE and MMT Programs, with the full participation of the technology developers, the EPA evaluates and documents the performance of innovative technologies by developing demonstration plans, conducting field tests, collecting and analyzing demonstration data, and preparing reports. The technologies are evaluated under rigorous quality assurance(QA) protocols to produce well-documented data of known quality. The EPA National Exposure Research Laboratory, which demonstrates field sampling, monitoring, and measurement technologies, selected Science Applications International Corporation as the verification organization to assist in field testing five field measurement devices for mercury in soil and sediment. This demonstration was funded by the SITE Program.

DEMONSTRATION DESCRIPTION

In May 2003, the EPA conducted a field demonstration of the DMA-80 and four other field measurement devices for mercury in soil and sediment. This verification statement focuses on the DMA-80; a similar statement has been prepared for each of the other four devices. The performance of the DMA-80 was compared to that of an off-site laboratory using the reference method, "Test Methods for Evaluating Solid Waste" (SW-846) Method 7471B (modified). To verify a wide range of performance attributes, the demonstration had both primary and secondary objectives. The primary objectives were:

- (1) Determining the instrument sensitivity with respect to the Method Detection Limit (MDL) and Practical Quantitation Limit (PQL);

-
- (2) Determining the analytical accuracy associated with the field measurement technologies;
 - (3) Evaluating the precision of the field measurement technologies;
 - (4) Measuring the amount of time required for mobilization and setup, initial calibration, daily calibration, sample analysis, and demobilization; and
 - (5) Estimating the costs associated with mercury measurements for the following four categories: capital, labor, supplies, and investigation-derived waste (IDW).

Secondary objectives for the demonstration included:

- (1) Documenting the ease of use, as well as the skills and training required to properly operate the device;
- (2) Documenting potential health and safety concerns associated with operating the device;
- (3) Documenting the portability of the device;
- (4) Evaluating the device durability based on its materials of construction and engineering design; and
- (5) Documenting the availability of the device and associated spare parts.

The DMA-80 analyzed 59 field soil samples, 13 field sediment samples, 42 spiked field samples, and 59 performance evaluation (PE) standard reference material (SRM) samples in the demonstration. The field samples were collected in four areas contaminated with mercury, the spiked samples were from these same locations, and the PE samples were obtained from a commercial provider.

Collectively, the field and PE samples provided the different matrix types and the different concentrations of mercury needed to perform a comprehensive evaluation of the DMA-80. A complete description of the demonstration and a summary of the results are available in the Innovative Technology Verification Report: "Field Measurement Technology for Mercury in Soil and Sediment—Milestone Inc.'s Direct Mercury Analyzer (DMA)-80"(EPA/600/R-04/012).

TECHNOLOGY DESCRIPTION

The DMA-80 is an atomic adsorption spectrophotometer based on mercury vaporization, amalgamation, desorption, and analysis of samples using an adsorbance spectrophotometer. Mercury samples are heated to 750° to 800°C, causing organic materials to be decomposed and mercury to be vaporized in a carrier gas of oxygen. The oxygen flow carries the vaporized mercury to the amalgamator, where it deposits on gold-covered molecular sieves. Potential interferences are carried out of the system with the continuous gas stream. The mercury deposits are then desorbed as the amalgamator is heated; vaporized mercury is transported to the spectrophotometer for analysis. The spectrophotometer uses a mercury vapor lamp as its light source. Light from the lamp is directed through an excitation filter before it irradiates the vaporized mercury contained in a quartz cuvette. The detector utilizes two sequential cuvettes: one for low concentration samples and the other for high concentration samples. Light which is not absorbed by the mercury vapors, then passes through an emission filter before being measured by the detector. Results are transmitted to the system controller, where concentrations are calculated based on sample mass and the detector response relative to a calibration curve.

During the demonstration, no extraction or sample digestion was required. Individual samples were mixed manually using a stainless steel spatula. (Note that samples were already considered to be homogeneous based upon the standard operating procedure used by SAIC to homogenize and aliquot all samples.) This same spatula was used to transfer the sample to a nickel weigh boat designed to fit the auto sampler. The sample was then weighed on a digital balance and placed on the 40-slot, auto sampler tray. The sample weight was automatically relayed to the DMA-80 controller; sequential sample numbers were automatically entered by the software in the data table in the location corresponding to the auto sampler location (1 - 40). Site-specific sample identification numbers were entered manually. The sample was analyzed, and the device displayed the mercury concentration in parts per million, which is equivalent to a soil concentration in milligrams per kilogram.

ACTION LIMITS

Action limits and concentrations of interest vary, and are project specific. There are, however, action limits which can be considered as potential reference points. The EPA Region IX Preliminary Remedial Goals for mercury are 23 mg/kg in residential soil and 310 mg/kg in industrial soil.

VERIFICATION OF PERFORMANCE

To ensure data usability, data quality indicators for accuracy, precision, representativeness, completeness, comparability, and sensitivity were assessed for the reference method based on project-specific QA objectives. Key demonstration findings are summarized below for the primary objectives.

Sensitivity: The two primary sensitivity evaluations performed for this demonstration were the MDL and PQL. Both will vary dependent upon whether the matrix is a soil, waste, or aqueous solution. Only soils/sediments were tested during this demonstration, and therefore, MDL calculations and PQL determinations for this evaluation are limited to those matrices. By definition, values measured below the PQL should not be considered accurate or precise and those below the MDL are not distinguishable from background noise.

Method Detection Limit - The evaluation of an MDL requires seven different measurements of a low concentration standard or sample. Following the procedures established in the 40 Code of Federal Regulations (CFR) Part 136, the MDL is estimated between 0.049 and 0.068 mg/kg. The equivalent calculated MDL for the referee laboratory is 0.0026 mg/kg. The calculated MDL is only intended as a statistical estimation and not a true test of instrument sensitivity.

Practical Quantitation Limit - The PQL for this instrument is approximately 0.082 mg/kg (the concentration of a SRM used during the demonstration) for soil and sediment materials. It is possible that the PQL may be as low as the MDL but there were no SRMs tested at this lower concentration. The referee laboratory PQL confirmed during the demonstration is 0.005 mg/kg, with a %D < 10%.

Accuracy: The results from the DMA-80 were compared to the 95% prediction interval for the SRM materials and to the referee laboratory results (Method 7471B). DMA-80 results were within SRM 95% prediction intervals 93% of the time, which suggests significant equivalence to certified standards. The number of Milestone average values less than 30% different from the referee laboratory results or SRM reference values; however, was 16 of 30 different sample lots. Only 2 of 30 Milestone average results have relative percent differences greater than 100% for this same group of samples. However, when making the comparison between Milestone and ALSI data, and taking into account the possible bias associated with both sets of data, this comparison may be within reasonable expectations for considering these two separate analyses to be equivalent. With the exception of a slight low bias for the referee laboratory and a slight high bias for the DMA-80 (similar to biases observed during other inter-laboratory studies), the data sets for the DMA-80 compared to the referee laboratory were considered to be similar and within expected statistical variation.

Precision: The precision of the Milestone field instrument is very comparable to the referee laboratory precision, and within expected precision variation for soil and sediment matrices. The overall average relative standard deviation (RSD) was 23.7% for the referee laboratory and 19.4% for Milestone. Both the laboratory and Milestone precision results are within the predicted 25% RSD objective for precision expected from both analytical and sampling variance.

Measurement Time: From the time of sample receipt, Milestone required 22 hours and 10 minutes to prepare a draft data package containing mercury results for 173 samples. One technician performed all setup, calibration checks, sample preparation and analysis, and equipment demobilization. Individual analyses took 5 minutes each (from the time the sample was injected until results were displayed), but the total time per analysis averaged nearly 8 minutes when all field activities and data package preparation were included in the calculation.

Measurement Costs: The cost per analysis based upon 173 samples, when renting the DMA-80, is \$35.90 per sample. The cost per analysis for the 173 samples, excluding rental fee, is \$18.55 per sample. Based on the 3-day field demonstration, the total cost for equipment rental and necessary supplies is estimated at \$6,210. The cost breakout by category is: capital costs, 48.3%; supplies, 9.5%; support equipment, 4.5%; labor, 14.5%; and IDW, 23.2%.

Key demonstration findings are summarized below for the secondary objectives.

Ease of Use: Based on observations made during the demonstration, the DMA-80 is easy to operate, requiring one field technician with a basic knowledge of chemistry acquired on the job or in a university and training on the DMA-80. A 1-day training course on instrument operation is offered at additional cost; this training would likely be necessary for most device operators who have no previous laboratory experience.

Potential Health and Safety Concerns: No significant health and safety concerns were noted during the demonstration. The only potential health and safety concerns identified were the generation of mercury vapors and the use of oxygen as the carrier gas. The vendor recommends and can provide a mercury filter; oxygen can be safely handled using standard laboratory procedures.

Portability: The DMA-80 was not easily portable (by hand) due to its size (80 cm by 42 cm by 30 cm high) and weight (56 kg). It was easy to set up and can be taken any place accessible to a small van or SUV. The instrument is better characterized as mobile rather than field portable. It operates on 110 or 220 volt AC current; no battery power supply is available.

Durability: The DMA-80 was well designed and constructed for durability. The auto sampler piston required re-alignment once early in the demonstration, an operation normally required after shipment. In two incidents related to piston alignment, one sample was dropped by the weigh boat injector and the auto-sampler tray later jammed. These problems were easily rectified, requiring less than 5 minutes each to troubleshoot and fix.

Availability of the Device: The DMA-80 is readily available for lease, or purchase. DMA-80 rental is available on a limited basis. Spare parts and consumable supplies can be added to the original DMA-80 order or can be received within 24 to 48 hours of order placement. Supplies and standards not provided by Milestone are readily available from laboratory supply firms.

PERFORMANCE SUMMARY

In summary, during the demonstration, the DMA-80 exhibited the following desirable characteristics of a field mercury measurement device: (1) good accuracy, (2) good precision, (3) high sample throughput, (4) low measurement costs, and (5) ease of use. During the demonstration the DMA-80 was found to have the following limitation: (1) non-portable due to the instrument size and weight. The demonstration findings collectively indicated that the DMA-80 is a reliable field measurement device for mercury in soil and sediment.

NOTICE: EPA verifications are based on an evaluation of technology performance under specific, predetermined criteria and appropriate quality assurance procedures. The EPA makes no expressed or implied warranties as to the performance of the technology and does not certify that a technology will always operate as 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 natural resources. Under the 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 provides data and scientific support that can be used to solve environmental problems, build the scientific knowledge base needed to manage ecological resources wisely, understand how pollutants affect public health, and prevent or reduce environmental risks.

The National Exposure Research Laboratory is the agency's center for 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 (1) develop and evaluate methods and technologies for characterizing and monitoring air, soil, and water; (2) support regulatory and policy decisions; and (3) provide the scientific support needed to ensure effective implementation of environmental regulations and strategies.

The EPA's Superfund Innovative Technology Evaluation (SITE) Program evaluates technologies designed for characterization and remediation of contaminated Superfund and Resource Conservation and Recovery Act (RCRA) sites. The SITE Program was created to provide reliable cost and performance data in order to speed acceptance and use of innovative remediation, characterization, and monitoring technologies by the regulatory and user community.

Effective monitoring and measurement technologies are needed to assess the degree of contamination at a site, provide data that can be used to determine the risk to public health or the environment, and monitor the success or failure of a remediation process. One component of the EPA SITE Program, the Monitoring and Measurement Technology (MMT) Program, demonstrates and evaluates innovative technologies to meet these needs.

Candidate technologies can originate within the federal government or the private sector. Through the SITE Program, developers are given the opportunity to conduct a rigorous demonstration of their technologies under actual field conditions. By completing the demonstration and distributing the results, the agency establishes a baseline for acceptance and use of these technologies. The MMT Program is managed by the Office of Research and Development's Environmental Sciences Division in Las Vegas, NV.

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

Abstract

Milestone's Direct Mercury Analyzer (DMA-80) was demonstrated under the U.S. Environmental Protection Agency Superfund Innovative Technology Evaluation Program in May 2003 at the Oak Ridge National Laboratory (ORNL) in Oak Ridge, TN. The purpose of the demonstration was to collect reliable performance and cost data for the DMA-80 and four other field measurement devices for mercury in soil and sediment. The key objectives of the demonstration were: 1) determine sensitivity of each instrument with respect to a vendor-generated method detection limit (MDL) and practical quantitation limit (PQL); 2) determine analytical accuracy associated with vendor field measurements using field samples and standard reference materials (SRMs); 3) evaluate the precision of vendor field measurements; 4) measure time required to perform mercury measurements; and 5) estimate costs associated with mercury measurements for capital, labor, supplies, and investigation-derived wastes.

The demonstration involved analysis of standard reference materials (SRMs), field samples collected from four sites, and spiked field samples for mercury. The performance results for a given field measurement device were compared to those of an off-site laboratory using reference method, "Test Methods for Evaluating Solid Waste" (SW-846) Method 7471B.

The sensitivity, accuracy, and precision measurements were successfully completed. The DMA-80 performed well in all these categories. During the demonstration, Milestone required 22 hours and 10 minutes for the analysis of 173 samples. The measurement costs were estimated to be \$6,210 for Milestone's DMA-80 rental option or \$35.90 per sample; \$18.55 per sample excluding rental fees.

The DMA-80 exhibited good ease of use and durability, as well as no major health and safety concerns. However, the device portability is somewhat limited by its size. Despite these limitations, the demonstration findings collectively indicated that the DMA-80 is a reliable field mobile measurement device for mercury in soil.

Contents

Notice	ii
Verification Statement	iii
Foreword	vii
Abstract	viii
Contents	ix
Tables	xii
Figures	xiii
Abbreviations, Acronyms, and Symbols	xiv
Acknowledgments	xvi

<u>Chapter</u>		<u>Page</u>
1	Introduction	1
	1.1 Description of the SITE Program	1
	1.2 Scope of the Demonstration	2
	1.2.1 Phase I	2
	1.2.2 Phase II	2
	1.3 Mercury Chemistry and Analysis	3
	1.3.1 Mercury Chemistry	3
	1.3.2 Mercury Analysis	4
2	Technology Description	6
	2.1 Description of Atomic Absorption Spectroscopy	6
	2.2 Description of the DMA-80	6
	2.3 Developer Contact Information	8
3	Field Sample Collection Locations and Demonstration Site	9
	3.1 Carson River	10
	3.1.1 Site Description	10
	3.1.2 Sample Collection	10
	3.2 Y-12 National Security Complex	11
	3.2.1 Site Description	11
	3.2.2 Sample Collection	11
	3.3 Confidential Manufacturing Site	11
	3.3.1 Site Description	11

Contents (Continued)

<u>Chapter</u>		<u>Page</u>
	3.3.2 Sample Collection	12
3.4	Puget Sound	12
	3.4.1 Site Description	12
	3.4.2 Sample Collection	12
3.5	Demonstration Site	13
3.6	SAIC GeoMechanics Laboratory	14
4	Demonstration Approach	15
	4.1 Demonstration Objectives	15
	4.2 Demonstration Design	16
	4.2.1 Approach for Addressing Primary Objectives	16
	4.2.2 Approach for Addressing Secondary Objectives	20
	4.3 Sample Preparation and Management	21
	4.3.1 Sample Preparation	21
	4.3.2 Sample Management	24
	4.4 Reference Method Confirmatory Process	25
	4.4.1 Reference Method Selection	25
	4.4.2 Referee Laboratory Selection	25
	4.4.3 Summary of Analytical Methods	27
	4.5 Deviations from the Demonstration Plan	28
5	Assessment of Laboratory Quality Control Measurements	29
	5.1 Laboratory QA Summary	29
	5.2 Data Quality Indicators for Mercury Analysis	29
	5.3 Conclusions and Data Quality Limitations	30
	5.4 Audit Findings	32
6	Performance of the DMA-80	33
	6.1 Primary Objectives	34
	6.1.1 Sensitivity	34
	6.1.2 Accuracy	36
	6.1.3 Precision	44
	6.1.4 Time Required for Mercury Measurement	47
	6.1.5 Cost	48
	6.2 Secondary Objectives	49
	6.2.1 Ease of Use	49
	6.2.2 Health and Safety Concerns	52
	6.2.3 Portability of the Device	53
	6.2.4 Instrument Durability	54
	6.2.5 Availability of Vendor Instruments and Supplies	54
7	Economic Analysis	56
	7.1 Issues and Assumptions	56
	7.1.1 Capital Equipment Cost	56
	7.1.2 Cost of Supplies	57

Contents (Continued)

<u>Chapter</u>		<u>Page</u>
	7.1.3 Support Equipment Cost	57
	7.1.4 Labor Cost	57
	7.1.5 Investigation-Derived Waste Disposal Cost	58
	7.1.6 Costs Not Included	58
7.2	DMA-80 Costs	59
	7.2.1 Capital Equipment Cost	59
	7.2.2 Cost of Supplies	60
	7.2.3 Support Equipment Cost	61
	7.2.4 Labor Cost	61
	7.2.5 Investigation-Derived Waste Disposal Cost	61
	7.2.6 Summary of DMA-80 Costs	62
7.3	Typical Reference Method Costs	63
8	Summary of Demonstration Results	64
	8.1 Primary Objectives	64
	8.2 Secondary Objectives	65
9	Bibliography	68
Appendix A -	Milestone Comments	69
Appendix B -	Statistical Analysis	70

Tables

Table		Page
1-1	Physical and Chemical Properties of Mercury	4
1-2	Methods for Mercury Analysis in Solids or Aqueous Soil Extracts	5
3-1	Summary of Site Characteristics	10
4-1	Demonstration Objectives	15
4-2	Summary of Secondary Objective Observations Recorded During the Demonstration	20
4-3	Field Samples Collected from the Four Sites	22
4-4	Analytical Methods for Non-Critical Parameters	28
5-1	MS/MSD Summary	30
5-2	LCS Summary	30
5-3	Precision Summary	31
5-4	Low Check Standards	31
6-1	Distribution of Samples Prepared for Milestone and the Referee Laboratory	33
6-2	Milestone SRM Comparison	38
6-3	ALSI SRM Comparison	38
6-4	Accuracy Evaluation by Hypothesis Testing	39
6-5	Number of Samples Within Each %D Range	41
6-6	Concentration of Non-Target Analytes	41
6-7	Evaluation of Precision	45
6-8	Time Measurements for Milestone	48
7-1	Capital Cost Summary for the DMA-80	60
7-2	Carrier Gas Cost Summary	60
7-3	Mercury Trap Costs	61
7-4	Labor Costs	61
7-5	IDW Costs	62
7-6	Summary of Rental Costs for the DMA-80	62
7-7	DMA-80 Costs by Category	63
8-1	Distribution of Samples Prepared for Milestone and the Referee Laboratory	65
8-2	Summary of DMA-80 Results for the Primary Objectives	66
8-3	Summary of DMA-80 Results for the Secondary Objectives	67
B-1	Summary of Unified Hypothesis Test	72

Figures

Figure		Page
2-1	Schematic of DMA-80.	7
2-2	Photograph of the DMA-80 during the field demonstration.	8
3-1	Tent and field conditions during the demonstration at Oak Ridge, TN.	13
3-2	Demonstration site and Building 5507.	13
4-1	Test sample preparation at the SAIC GeoMechanics Laboratory.	23
6-1	Data plot for low concentration sample results	42
6-2	Data plot for high concentration sample results	43
6-3	Calibration result screen.	49
6-4	Calibration curve screen	50
6-5	System control display screen	51
6-6	Sample peak screen.	52
7-1	Capital equipment costs	59

Abbreviations, Acronyms, and Symbols

%	Percent
%D	Percent difference
°C	Degrees Celsius
µg/kg	Microgram per kilogram
AAS	Atomic absorption spectrometry
ALSI	Analytical Laboratory Services, Inc.
bgs	Below ground surface
cm	Centimeter
CFR	Code of Federal Regulations
CI	Confidence Interval
COC	Chain of custody
DMA-80	Direct Mercury Analyzer
DOE	Department of Energy
EPA	United States Environmental Protection Agency
g	Gram
H&S	Health and Safety
Hg	Mercury
HgCl ₂	Mercury (II) chloride
IDL	Instrument detection limit
IDW	Investigation-derived waste
ITVR	Innovative Technology Verification Report
kg	Kilogram
L	Liter
mL/min	Milliliter per minute
LCS	Laboratory control sample
LEFPC	Lower East Fork Poplar Creek
m	Meter
MDL	Method detection limit
mg	Milligram
mg/kg	Milligram per kilogram
mL	Milliliter
mm	Millimeter
MMT	Monitoring and Measurement Technology
MS/MSD	Matrix Spike/ Matrix Spike Duplicate
NERL	National Exposure Research Laboratory
ng	Nanogram
nm	Nanometer
ORD	Office of Research and Development

Abbreviations, Acronyms, and Symbols (Continued)

ORR	Oak Ridge Reservation
ORNL	Oak Ridge National Laboratory
OSWER	Office of Solid Waste and Emergency Response
ppb	Parts per billion
PPE	Personal protective equipment
ppm	Parts per million
PQL	Practical quantitation limit
QA	Quality assurance
QAPP	Quality Assurance Project Plan
QC	Quality control
RPD	Relative percent difference
RSD	Relative standard deviation
SAIC	Science Applications International Corporation
SITE	Superfund Innovative Technology Evaluation
SOP	Standard operating procedure
SRM	Standard reference material
SW-846	Test Methods for Evaluating Solid Waste; Physical/Chemical Methods
TOC	Total organic carbon
TOM	Task Order Manager
UL	Underwriters Laboratory
UEFPC	Upper East Fork of Poplar Creek
Y-12	Y-12 Oak Ridge Security Complex, Oak Ridge, TN

Acknowledgments

The U.S. Environmental Protection Agency (EPA) Superfund Innovative Technology Evaluation wishes to acknowledge the support of the following individuals in performing the demonstration and preparing this document: Elizabeth Phillips of the U.S. Department of Energy Oak Ridge National Laboratory (ORNL); Stephen Childs, Thomas Early, Roger Jenkins, and Monty Ross of the UT-Battelle ORNL; Dale Rector of the Tennessee Department of Environment and Conservation (TDEC) Department of Energy Oversight; Mikhail Mensh of Milestone, Inc.; Leroy Lewis of the Idaho National Engineering and Environmental Laboratory, retired; Ishwar Murarka of the EPA Science Advisory Board, member; Danny Reible of Louisiana State University; Mike Bolen, Joseph Evans, Julia Gartseff, Sara Hartwell, Cathleen Hubbard, Kevin Jago, Andrew Matuson, Allen Motley, John Nicklas, Maurice Owens, Nancy Patti, Fernando Padilla, Mark Pruitt, James Rawe, Herb Skovronek, and Joseph Tillman of Science Applications International Corporation (SAIC); Scott Jacobs and Ann Vega of the EPA National Risk Management Research Laboratory's Land Remediation and Pollution Control Division; and Brian Schumacher of the EPA National Exposure Research Laboratory.

This document was QA reviewed by George Brilis of the EPA National Exposure Research Laboratory.

Chapter 1 Introduction

The U.S. Environmental Protection Agency (EPA) under the Office of Research and Development (ORD), National Exposure Research Laboratory (NERL), conducted a demonstration to evaluate the performance of innovative field measurement devices for their ability to measure mercury concentrations in soils and sediments. This Innovative Technology Verification Report (ITVR) presents demonstration performance results and associated costs of Milestone's Direct Mercury Analyzer (DMA) -80. The vendor-prepared comments regarding the demonstration are presented in Appendix A.

The demonstration was conducted as part of the EPA Superfund Innovative Technology Evaluation (SITE) Monitoring and Measurement Technology (MMT) Program. Mercury contaminated soils and sediments, collected from four sites within the continental U.S., comprised the majority of samples analyzed during the evaluation. Some soil and sediment samples were spiked with mercury (II) chloride (HgCl_2) to provide concentrations not occurring in the field samples. Certified standard reference material (SRM) samples were also used to provide samples with certified mercury concentrations and to increase the matrix variety.

The demonstration was conducted at the Department of Energy (DOE) Oak Ridge National Laboratory (ORNL) in Oak Ridge, TN during the week of May 5, 2003. The purpose of the demonstration was to obtain reliable performance and cost data for field measurement devices in order to 1) provide potential users with a better understanding of the devices' performance and operating costs under well-defined field conditions and 2) provide the instrument vendors with documented results that can assist them in promoting acceptance and use of their devices. The results obtained using the five field mercury measurement devices were compared to the mercury

results obtained for identical sample sets (samples, spiked samples, and SRMs) analyzed at a referee laboratory. The referee laboratory, which was selected prior to the demonstration, used a well-established EPA reference method.

1.1 Description of the SITE Program

Performance verification of innovative environmental technologies is an integral part of the regulatory and research mission of the EPA. The SITE Program was established by EPA's Office of Solid Waste and Emergency Response (OSWER) and ORD under the Superfund Amendments and Reauthorization Act of 1986.

The overall goal of the SITE Program is to conduct performance verification studies and to promote the acceptance of innovative technologies that may be used to achieve long-term protection of human health and the environment. The program is designed to meet three main objectives: 1) identify and remove obstacles to the development and commercial use of innovative technologies; 2) demonstrate promising innovative technologies and gather reliable performance and cost information to support site characterization and cleanup activities; and 3) develop procedures and policies that encourage the use of innovative technologies at Superfund sites, as well as at other waste sites or commercial facilities.

The SITE Program includes the following elements:

- The MMT Program evaluates innovative technologies that sample, detect, monitor, or measure hazardous and toxic substances in soil, water, and sediment samples. These technologies are expected to provide better, faster, or more cost-effective methods for

producing real-time data during site characterization and remediation studies than conventional technologies.

- The Remediation Technology Program conducts demonstrations of innovative treatment technologies to provide reliable performance, cost, and applicability data for site cleanups.
- The Technology Transfer Program provides and disseminates technical information in the form of updates, brochures, and other publications that promote the SITE Program and participating technologies. The Technology Transfer Program also offers technical assistance, training, and workshops in the support of the technologies. A significant number of these activities are performed by EPA's Technology Innovation Office.

The Field Analysis of Mercury in Soils and Sediments demonstration was performed under the MMT Program. The MMT Program provides developers of innovative hazardous waste sampling, detection, monitoring, and measurement devices with an opportunity to demonstrate the performance of their devices under actual field conditions. The main objectives of the MMT Program are as follows:

- Test and verify the performance of innovative field sampling and analytical technologies that enhance sampling, monitoring, and site characterization capabilities.
- Identify performance attributes of innovative technologies that address field sampling, monitoring, and characterization problems in a cost-effective and efficient manner.
- Prepare protocols, guidelines, methods, and other technical publications that enhance acceptance of these technologies for routine use.

The MMT Program is administered by the Environmental Sciences Division of the NERL in Las Vegas, NV. The NERL is the EPA center for investigation of technical and management approaches for identifying and quantifying risks to human health and the environment. The NERL mission components include 1) developing and evaluating methods and technologies for sampling, monitoring, and characterizing water, air, soil, and sediment; 2) supporting regulatory and policy decisions; and 3) providing technical support to ensure the effective implementation of environmental regulations and strategies.

1.2 Scope of the Demonstration

The demonstration project consisted of two separate phases: Phase I involved obtaining information on prospective vendors having viable mercury detection instrumentation. Phase II consisted of field and planning activities leading up to and including the demonstration activities. The following subsections provide detail on both of these project phases.

1.2.1 Phase I

Phase I was initiated by making contact with knowledgeable sources on the subject of "mercury in soil" detection devices. Contacts included individuals within EPA, Science Applications International Corporation (SAIC), and industry where measurement of mercury in soil was known to be conducted. Industry contacts included laboratories and private developers of mercury detection instrumentation. In addition, the EPA Task Order Manager (TOM) provided contacts for "industry players" who had participated in previous MMT demonstrations. SAIC also investigated university and other research-type contacts for knowledgeable sources within the subject area.

These contacts led to additional knowledgeable sources on the subject, which in turn led to various Internet searches. The Internet searches were very successful in finding additional companies involved with mercury detection devices.

All in all, these research activities generated an original list of approximately 30 companies potentially involved in the measurement of mercury in soils. The list included both international and U.S. companies. Each of these companies was contacted by phone or email to acquire further information. The contacts resulted in 10 companies that appeared to have viable technologies.

Due to instrument design (i.e., the instrument's ability to measure mercury in soils and sediments), business strategies, and stage of technology development, only 5 of those 10 vendors participated in the field demonstration portion of phase II.

1.2.2 Phase II

Phase II of the demonstration project involved strategic planning, field-related activities for the demonstration, data analysis, data interpretation, and preparation of the ITVRs. Phase II included pre-demonstration and demonstration activities, as described in the following subsections.

1.2.2.1 Pre-Demonstration Activities

The pre-demonstration activities were completed in the fall 2002. There were six objectives for the pre-demonstration:

- Establish concentration ranges for testing vendors' analytical equipment during the demonstration.
- Collect soil and sediment field samples to be used in the demonstration.
- Evaluate sample homogenization procedures.
- Determine mercury concentrations in homogenized soils and sediments.
- Select a reference method and qualify potential referee laboratories for the demonstration.
- Provide soil and sediment samples to the vendors for self-evaluation of their instruments, as a precursor to the demonstration.

As an integral part of meeting these objectives, a pre-demonstration sampling event was conducted in September 2002 to collect field samples of soils and sediments containing different levels of mercury. The field samples were obtained from the following locations:

- Carson River Mercury site - near Dayton, NV
- Y-12 National Security Complex - Oak Ridge, TN
- A confidential manufacturing facility - eastern U.S.
- Puget Sound - Bellingham Bay, WA

Immediately after collecting field sample material from the sites noted above, the general mercury concentrations in the soils and sediments were confirmed by quick turnaround laboratory analysis of field-collected subsamples using method SW-7471B. The field sample materials were then shipped to a soil preparation laboratory for homogenization. Additional pre-demonstration activities are detailed in Chapter 4.

1.2.2.2 Demonstration Activities

Specific objectives for this SITE demonstration were developed and defined in a Field Demonstration and Quality Assurance Project Plan (QAPP) (EPA Report # EPA/600/R-03/053). The Field Demonstration QAPP is available through the EPA ORD web site

(<http://www.epa.gov/ORD/SITE>) or from the EPA Project Manager. The demonstration objectives were subdivided into two categories: primary and secondary. Primary objectives are goals of the demonstration study that need to be achieved for technology verification. The measurements used to achieve primary objectives are referred to as critical. These measurements typically produce quantitative results that can be verified using inferential and descriptive statistics.

Secondary objectives are additional goals of the demonstration study developed for acquiring other information of interest about the technology that is not directly related to verifying the primary objectives. The measurements required for achieving secondary objectives are considered to be noncritical. Therefore, the analysis of secondary objectives is typically more qualitative in nature and often uses observations and sometimes descriptive statistics.

The field portion of the demonstration involved evaluating the capabilities of five mercury-analyzing instruments to measure mercury concentrations in soil and sediment. During the demonstration, each instrument vendor received three types of samples 1) homogenized field samples referred to as "field samples", 2) certified SRMs, and 3) spiked field samples (spikes).

Spikes were prepared by adding known quantities of HgCl₂ to field samples. Together, the field samples, SRMs, and spikes are referred to as "demonstration samples" for the purpose of this ITVR. All demonstration samples were independently analyzed by a carefully selected referee laboratory. The experimental design for the demonstration is detailed in Chapter 4.

1.3 Mercury Chemistry and Analysis

1.3.1 Mercury Chemistry

Elemental mercury is the only metal that occurs as a liquid at ambient temperatures. Mercury naturally occurs, primarily within the ore, cinnabar, as mercury sulfide (HgS). Mercury easily forms amalgams with many other metals, including gold. As a result, mercury has historically been used to recover gold from ores.

Mercury is ionically stable; however, it is very volatile for a metal. Table 1-1 lists selected physical and chemical properties of elemental mercury.

Table 1-1. Physical and Chemical Properties of Mercury

Properties	Data
Appearance	Silver-white, mobile, liquid.
Hardness	Liquid
Abundance	0.5% in Earth's crust
Density @ 25 °C	13.53 g/mL
Vapor Pressure @ 25 °C	0.002 mm
Volatilizes @	356 °C
Solidifies @	-39 °C

Source: Merck Index, 1983

Historically, mercury releases to the environment included a number of industrial processes such as chloralkali manufacturing, copper and zinc smelting operations, paint application, waste oil combustion, geothermal energy plants, municipal waste incineration, ink manufacturing, chemical manufacturing, paper mills, leather tanning, pharmaceutical production, and textile manufacturing. In addition, industrial and domestic mercury-containing products, such as thermometers, electrical switches, and batteries, are disposed of as solid wastes in landfills (EPA, July 1995). Mercury is also an indigenous compound at many abandoned mining sites and is, of course, found as a natural ore.

At mercury-contaminated sites, mercury exists in mercuric form (Hg^{2+}), mercurous form (Hg_2^{2+}), elemental form (Hg^0), and alkylated form (e.g., methyl or ethyl mercury). Hg_2^{2+} and Hg^{2+} are the more stable forms under oxidizing conditions. Under mildly reducing conditions, both organically bound mercury and inorganic mercury may be degraded to elemental mercury, which can then be converted readily to methyl or ethyl mercury by biotic and abiotic processes. Methyl and ethyl mercury are the most toxic forms of mercury; the alkylated mercury compounds are volatile and soluble in water.

Mercury (II) forms relatively strong complexes with Cl^- and CO_3^{2-} . Mercury (II) also forms complexes with inorganic ligands such as fluoride (F^-), bromide (Br^-), iodide (I^-), sulfate (SO_4^{2-}), sulfide (S^{2-}), and phosphate (PO_4^{3-}) and forms strong complexes with organic ligands, such as sulfhydryl groups, amino acids, and humic and fulvic acids. The insoluble HgS is formed under mildly reducing conditions.

1.3.2 Mercury Analysis

There are several laboratory-based, EPA promulgated methods for the analysis of mercury in solid and liquid hazardous waste matrices. In addition, there are several performance-based methods for the determination of various mercury species. Table 1-2 summarizes the commonly used methods for measuring mercury in both solid and liquid matrices, as identified through a review of the EPA Test Method Index and SW-846. A discussion of the choice of reference method is presented in Chapter 4.

Table 1-2. Methods for Mercury Analysis in Solids or Aqueous Soil Extracts

Method	Analytical Technology	Type(s) of Mercury analyzed	Approximate Concentration Range	Comments
SW-7471B	CVAAS	<ul style="list-style-type: none"> • inorganic mercury • organo-mercury 	10-2,000 ppb	Manual cold vapor technique widely used for total mercury determinations
SW-7472	ASV	<ul style="list-style-type: none"> • inorganic mercury • organo-mercury 	0.1-10,000 ppb	Newer, less widely accepted method
SW-7473	TD, amalgamation, and AAS	<ul style="list-style-type: none"> • inorganic mercury • organo-mercury 	0.2 - 400 ppb	Allows for total decomposition analysis
SW-7474	AFS	<ul style="list-style-type: none"> • inorganic mercury • organo-mercury 	1 ppb - ppm	Allows for total decomposition analysis; less widely used/reference
EPA 1631	CVAFS	<ul style="list-style-type: none"> • inorganic mercury • organo-mercury 	0.5 - 100 ppt	Requires "trace" analysis procedures; written for aqueous matrices; Appendix A of method written for sediment/soil samples
EPA 245.7	CVAFS	<ul style="list-style-type: none"> • inorganic mercury • organo-mercury 	0.5 - 200 ppt	Requires "trace" analysis procedures; written for aqueous matrices; will require dilutions of high-concentration mercury samples
EPA 6200	FPXRF	<ul style="list-style-type: none"> • inorganic mercury 	>30 mg/kg	Considered a screening protocol

AAS = Atomic Absorption Spectrometry
 AAF = Atomic Fluorescence Spectrometry
 AFS = Atomic Fluorescence Spectrometry
 ASV = Anodic Stripping Voltammetry
 CVAAS = Cold Vapor Atomic Absorption Spectrometry
 CVAFS = Cold Vapor Atomic Fluorescence Spectrometry
 FPXRF = Field Portable X-ray Fluorescence
 EPA = U.S. Environmental Protection Agency
 mg/kg = milligram per kilogram
 ppb = parts per billion
 ppm = parts per million
 ppt = parts per trillion
 SW = solid waste
 TD = thermal decomposition

Chapter 2 Technology Description

This chapter provides a detailed description of 1) the thermal decomposition method of atomic absorption spectroscopy (AAS), which is the type of technology on which Milestone's instrument is based, and 2) a detailed description of the DMA-80.

2.1 Description of Atomic Absorption Spectroscopy

The principle of analysis used by the DMA-80 is AAS preceded by thermal decomposition, catalytic reduction, and amalgamation desorption. AAS is based on the direct relationship between the absorption of light of a specific wavelength by gas-phase atoms of an inorganic analyte, and the concentration of those atoms. Because samples analyzed by AAS are usually liquids or solids, the analyte atoms or ions must be vaporized in a flame or graphite furnace prior to the determination. The vaporized atoms absorb light of an analyte-specific wavelength, and make transitions to higher electronic energy levels. The analyte concentration is directly proportional to the amount of light absorbed. Concentration measurements are usually determined from a working curve after calibrating the instrument with standards of known concentration.

In reference to AAS as a general analytical application, thermal decomposition, followed by atomic absorption spectroscopy, is common; however, the mechanism of analyte recovery for the determination step may vary. Examples include cold vapor traps and amalgamation desorption.

When operating the DMA-80, a sample of known mass is placed in the drying and decomposition furnace and heated to 750 Celsius (°C). The liquid or solid sample is dried and organic materials are decomposed. AAS utilizes the absorption of light by an element, in this case mercury

vapor, as compared to a standard to quantify the mass of analyte present in a sample. The absorption of light is proportional to the concentration of analyte present. The wavelength of the light source is specific to the contaminant of interest. For mercury, the wavelength is 254 nm.

2.2 Description of the DMA-80

The Milestone DMA-80 is an integrated system that utilizes thermal decomposition, catalytic reduction, amalgamation desorption, and AAS to rapidly analyze solid and liquid samples.

Applications and Specifications – The Milestone DMA-80 enables analysts to rapidly determine total mercury concentrations in solid and liquid samples without sample pretreatment or digestion. Maximum sample sizes are 500 µL and 500 mg, respectively, for liquid and solid samples. According to Milestone, individual sample results are available in approximately 5 minutes and up to 40 samples can be processed, start to finish, in a 4-hour period. Per Milestone, results are reportedly independent of matrix, detection limits range from 0.5 to 600 ng mercury on a mass basis, and reproducibility (measurement error for two or more samples) is less than 1.5 percent. Results from this demonstration are reported in Chapter 6.

In areas where mercury contamination in the soil is an existing problem, the background signal may be significantly increased due to airborne dust containing mercury. As with other AAS technologies, memory effects between analyses may be encountered when analyzing a sample of high mercury content (e.g., 400 ng) prior to analyzing one of low content (e.g., 25 ng). Typically, to minimize memory effects, samples are analyzed in batches of low and high concentrations, analyzing those of low

concentration first. If this batching process cannot be accomplished, a blank analysis, with an extended decomposition time, may be required following the analysis of a highly-concentrated sample to limit memory effects. Co-absorbing gases, such as free chlorine and certain organics (as indicated in Methods 7470A and 7471B), should not interfere due to the release of decomposition products by the decomposition furnace, removal of some decomposition products by the decomposition catalyst, and the selective entrapment of mercury vapor on the amalgamator. As with other analytical devices, field conditions that may affect accuracy and precision include sample homogeneity, sample handling errors, unpredictable matrix effects, and sample and cell contamination (EPA, 1998).

Because no sample digestion or pre-treatment is required, no reagents are utilized. As a result, the only waste materials are residual sample material, excess sample, and decontamination solution. The DMA-80 volatilizes mercury into the oxygen stream flowing through the instrument, which ultimately exhausts to ambient air. The instrument exhaust may be attached to a fume hood with a filter, or a mercury trap may be assembled and attached in the field, based on instructions provided by the vendor. Figure 2-1 presents a schematic diagram of the thermal decomposition, catalytic reduction unit, and amalgamation desorption furnace for the DMA-80.

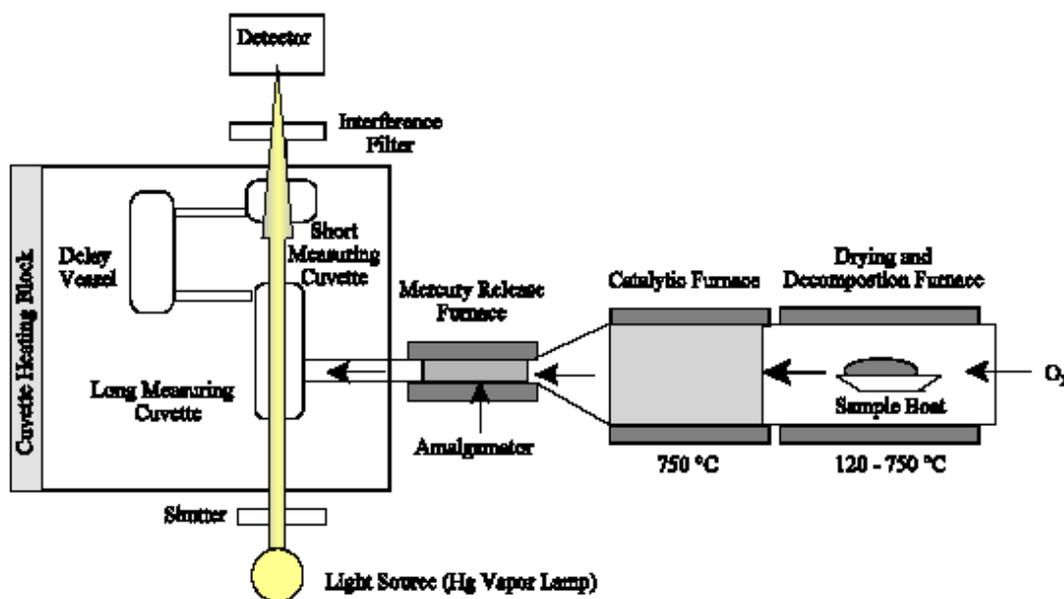


Figure 2-1. Schematic of DMA-80

The Milestone DMA-80 is approximately 56 kg and has a dimension of 80 cm by 42 cm by 30 cm (height). The terminal has a dimension of 33 cm by 27 cm by 26 cm (height) and weighs less than 2 kg. The instrument operates on either a 110V or 230V AC source at 50 to 60 Hz. The unit is equipped with a built-in, 40-position auto sampler for solids and liquids. An optional analytical

balance can be provided for automatic sample weight data transfer. (A steady table is needed for accurate weight determination.) Sample weight data transfer can be accomplished from other appropriate balances by utilizing a 9-pin connector. Other required equipment includes a micro spatula, tweezers, and digital pipets with 10-100 µl and 100-1,000 µl ranges.

Operation – Liquid or solid samples are introduced into the DMA-80 (Figure 2-2) individually or using the auto sampler. The sample is initially dried in an oxygen stream passing through a quartz tube located inside a controlled heating coil. A separate cylinder supplies oxygen as a carrier gas; the flow rate is approximately 200 mL/minute at 60 psig. The combustion gases are further decomposed on a catalytic column at 750 °C. Mercury vapor is collected on a gold amalgamation trap and subsequently desorbed for quantitation. Mercury content is determined using a single beam spectrophotometer with two sequential, flow-through measurement cells. The light source is a low-pressure mercury vapor lamp. The instrument detector is a silicon UV photo-detector at 253.65 nm, with a 254 nm interference filter having a 9-nm bandwidth.



Figure 2-2. Photograph of the DMA-80 during the field demonstration.

Each cell has its own calibration curve. The system provides automatic switch-over between the low and high working ranges. The low range is 0-35 ng mercury; the high range is 35-600 ng mercury. Calibration standards are not provided with the instrument; however, the electronic instructions provided by Milestone included embedded links to internet sites where standards can be purchased.

Results are displayed using a touch-screen, Pentium-based control terminal equipped with a keyboard and a mouse. The Windows-based system control software provides automatic data storage; edit functions to create, modify, and store commonly used methods; and options to select single or auto sample. Standard data include absorbance, mercury mass (nanogram), and total mercury concentration (parts per billion). Data can be printed to a standard printer or stored.

2.3 Developer Contact Information

Additional information about the DMA-80 can be obtained from the following source:

Milestone
160 B Shelton Road
Monroe, CT 06468
Telephone: (203) 261-6175
Fax: (203) 261-6592
Email: techsales@milestonesci.com
Internet: www.milestonesci.com

Chapter 3

Field Sample Collection Locations and Demonstration Site

As previously described in Chapter 1, the demonstration in part tested the ability of all five vendor instruments to measure mercury concentrations in demonstration samples. The demonstration samples consisted of field-collected samples, spiked field samples, and SRMs. The field-collected samples comprised the majority of demonstration samples. This chapter describes the four sites from which the field samples were collected, the demonstration site, and the sample homogenization laboratory. Spiked samples were prepared from these field samples.

Screening of potential mercury-contaminated field sample sites was conducted during Phase I of the project. Four sites were selected for acquiring mercury-contaminated samples that were diverse in appearance, consistency, and mercury concentration. A key criterion was the source of the contamination. These sites included:

- Carson River Mercury site - near Dayton, NV
- The Y-12 National Security Complex (Y-12) - Oak Ridge, TN
- A confidential manufacturing facility - eastern U.S.
- Puget Sound - Bellingham Bay, WA

Site Diversity – Collectively, the four sites provided sampling areas with both soil and sediment, having variable physical consistencies and variable ranges of mercury contamination. Two of the sites (Carson River and Oak Ridge) provided both soil and sediment samples. A third site (a manufacturing facility) provided just soil samples and a fourth site (Puget Sound) provided only sediment samples.

Access and Cooperation – Site representatives were instrumental in providing site access, and in some cases,

guidance on the best areas to collect samples from relatively high and low mercury concentrations. In addition, representatives from the host demonstration site (ORNL) provided a facility for conducting the demonstration.

At three of the sites, the soil and/or sediment sample was collected, homogenized by hand in the field, and subsampled for quick turnaround analysis. These subsamples were sent to analytical laboratories to determine the general range of mercury concentrations at each of the sites. (The Puget Sound site did not require confirmation of mercury contamination due to recently acquired mercury analytical data from another, ongoing research project.) The field-collected soil and sediment samples from all four sites were then shipped to SAIC's GeoMechanics Laboratory for a more thorough sample homogenization (see Section 4.3.1) and subsampled for redistribution to vendors during the pre-demonstration vendor self-evaluations.

All five of the technology vendors performed a self-evaluation on selected samples collected and homogenized during this pre-demonstration phase of the project. For the self-evaluation, the laboratory results and SRM values were supplied to the vendor, allowing the vendor to determine how well it performed the analysis on the field samples. The results were used to gain a preliminary understanding of the field samples collected and to prepare for the demonstration.

Table 3-1 summarizes key characteristics of samples collected at each of the four sites. Also included are the sample matrix, sample descriptions, and sample depth intervals. The analytical results presented in Table 3-1 are based on referee laboratory mercury results for the demonstration samples.

Table 3-1. Summary of Site Characteristics

Site Name	Sampling Area	Sample Matrix	Depth	Description	Hg Concentration Range
Carson River Mercury site	Carson River	Sediment	water/sediment interface	Sandy silt, with some organic debris present (plant stems and leaves)	10 ppb - 50 ppm
	Six Mile Canyon	Soil	3 - 8 cm bgs	Silt with sand to sandy silt	10 ppb - 1,000 ppm
Y-12 National Security Complex	Old Hg Recovery Bldg.	Soil	0 - 1 m bgs	Silty-clay to sandy-gravel	0.1 - 100 ppm
	Poplar Creek	Sediment	0 - 0.5 m bgs	Silt to coarse sandy gravel	0.1 - 100 ppm
Confidential manufacturing site	Former plant building	Soil	3.6 -9 m bgs	Silt to sandy silt	5 - 1,000 ppm
Puget Sound - Bellingham Bay	Sediment layer	Sediment	1.5 - 1.8 m thick	Clayey-sandy silt with various woody debris	10 - 400 ppm
	Underlying Native Material	Sediment	0.3 m thick	Medium-fine silty sands	0.16 - 10 ppm

bgs = below ground surface.

3.1 Carson River

3.1.1 Site Description

The Carson River Mercury site begins near Carson City, NV, and extends downstream to the Lahontan Valley and the Carson Desert. During the Comstock mining era of the late 1800s, mercury was imported to the area for processing gold and silver ore. Ore mined from the Comstock Lode was transported to mill sites, where it was crushed and mixed with mercury to amalgamate the precious metals. The Nevada mills were located in Virginia City, Silver City, Gold Hill, Dayton, Six Mile Canyon, Gold Canyon, and adjacent to the Carson River between New Empire and Dayton. During the mining era, an estimated 7,500 tons of mercury were discharged into the Carson River drainage, primarily in the form of mercury-contaminated tailings (EPA Region 9, 1994).

Mercury contamination is present at Carson River as either elemental mercury and/or inorganic mercury sulfides with less than 1%, if any, methylmercury. Mercury contamination exists in soils present at the former gold and silver mining mill sites; waterways adjacent to the mill sites; and sediment, fish, and wildlife over more than a 50-mile length of the Carson River. Mercury is also present in the sediments and adjacent flood plain of the Carson River, and in the sediments of Lahontan Reservoir, Carson Lake, Stillwater Wildlife Refuge, and Indian Lakes. In addition, tailings with elevated mercury levels are still present at, and around, the historic mill sites, particularly in Six Mile Canyon (EPA, 2002a).

3.1.2 Sample Collection

The Carson River Mercury site provided both soil and sediment samples across the range of contaminant concentrations desired for the demonstration. Sixteen near-surface soil samples were collected between 3-8 cm below ground surface (bgs). Two sediment samples were collected at the water-to-sediment interface. All 18 samples were collected on September 23-24, 2002 with a hand shovel. Samples were collected in Six Mile Canyon and along the Carson River.

The sampling sites were selected based upon historical data from the site. Specific sampling locations in the Six Mile Canyon were selected based upon local terrain and visible soil conditions (e.g., color and particle size). The specific sites were selected to obtain soil samples with as much variety in mercury concentration as possible. These sites included hills, run-off pathways, and dry river bed areas. Sampling locations along the Carson River were selected based upon historical mine locations, local terrain, and river flow.

When collecting the soil samples, approximately 3 cm of surface soil was scraped to the side. The sample was then collected with a shovel, screened through a 6.3-millimeter (mm) (0.25-inch) sieve to remove larger material, and collected in 4-liter (L) sealable bags identified with a permanent marker. The sediment samples were also collected with a shovel, screened through a 6.3-mm sieve to remove larger material, and collected in 4-L sealable bags identified with a permanent marker. Each of the 4-L sealable bags was placed into a second 4-L

sealable bag, and the sample label was placed onto the outside bag. The sediment samples were then placed into 10-L buckets, lidded, and identified with a sample label.

3.2 Y-12 National Security Complex

3.2.1 Site Description

The Y-12 site is located at the DOE ORNL in Oak Ridge, TN. The Y-12 site is an active manufacturing and developmental engineering facility that occupies approximately 800 acres on the northeast corner of the DOE Oak Ridge Reservation (ORR) adjacent to the city of Oak Ridge, TN. Built in 1943 by the U.S. Army Corps of Engineers as part of the World War II Manhattan Project, the original mission of the installation was development of electromagnetic separation of uranium isotopes and weapon components manufacturing, as part of the national effort to produce the atomic bomb. Between 1950 and 1963, large quantities of elemental mercury were used at Y-12 during lithium isotope separation pilot studies and subsequent production processes in support of thermonuclear weapons programs.

Soils at the Y-12 facility are contaminated with mercury in many areas. One of the areas of known high levels of mercury-contaminated soils is in the vicinity of a former mercury use facility (the "Old Mercury Recovery Building" – Building 8110). At this location, mercury-contaminated material and soil were processed in a Nicols-Herschoff roasting furnace to recover mercury. Releases of mercury from this process, and from a building sump used to secure the mercury-contaminated materials and the recovered mercury, have contaminated the surrounding soils (Rothchild, et al., 1984). Mercury contamination also occurred in the sediments of the East Fork of Poplar Creek (DOE, 1998). The Upper East Fork of Poplar Creek (UEFPC) drains the entire Y-12 complex. Releases of mercury via building drains connected to the storm sewer system, building basement dewatering sump discharges, and spills to soils, all contributed to contamination of UEFPC. Recent investigations showed that bank soils containing mercury along the UEFPC were eroding and contributing to mercury loading. Stabilization of the bank soils along this reach of the creek was recently completed.

3.2.2 Sample Collection

Two matrices were sampled at Y-12 in Oak Ridge, TN, creek sediment and soil. A total of 10 sediment samples was collected; one sediment sample was collected from the Lower East Fork of Poplar Creek (LEFPC) and nine sediment samples were collected from the UEFPC. A total

of six soil samples was collected from the Building 8110 area. The sampling procedures that were used are summarized below.

Creek Sediments – Creek sediments were collected on September 24-25, 2002 from the East Fork of Poplar Creek. Sediment samples were collected from various locations in a downstream to upstream sequence (i.e., the downstream LEFPC sample was collected first and the most upstream point of the UEFPC was sampled last).

The sediment samples from Poplar Creek were collected using a commercially available clam-shell sonar dredge attached to a rope. The dredge was slowly lowered to the creek bottom surface, where it was pushed by foot into the sediment. Several drops of the sampler (usually seven or more) were made to collect enough material for screening. On some occasions, a shovel was used to remove overlying "hardpan" gravel to expose finer sediments at depth. One creek sample consisted of creek bank sediments, which was collected using a stainless steel trowel.

The collected sediment was then poured onto a 6.3-mm sieve to remove oversize sample material. Sieved samples were then placed in 12-L sealable plastic buckets. The sediment samples in these buckets were homogenized with a plastic ladle and subsamples were collected in 20-milliliter (mL) vials for quick turnaround analyses.

Soil – Soil samples were collected from pre-selected boring locations September 25, 2002. All samples were collected in the immediate vicinity of the Building 8110 foundation using a commercially available bucket auger. Oversize material was hand picked from the excavated soil because the soil was too wet to be passed through a sieve. The soil was transferred to an aluminum pan, homogenized by hand, and subsampled to a 20-mL vial. The remaining soil was transferred to 4-L plastic containers.

3.3 Confidential Manufacturing Site

3.3.1 Site Description

A confidential manufacturing site, located in the eastern U.S., was selected for participation in this demonstration. The site contains elemental mercury, mercury amalgams, and mercury oxide in shallow sediments (less than 0.3 m deep) and deeper soils (3.65 to 9 m bgs). This site provided soil with concentrations from 5-1,000 mg/kg.

The site is the location of three former processes that resulted in mercury contamination. The first process

involved amalgamation of zinc with mercury. The second process involved the manufacturing of zinc oxide. The third process involved the reclamation of silver and gold from mercury-bearing materials in a retort furnace. Operations led to the dispersal of elemental mercury, mercury compounds such as chlorides and oxides, and zinc-mercury amalgams. Mercury values have been measured ranging from 0.05 to over 5,000 mg/kg, with average values of approximately 100 mg/kg.

3.3.2 Sample Collection

Eleven subsurface soil samples were collected on September 24, 2002. All samples were collected with a Geoprobe[®] unit using plastic sleeves. All samples were collected at the location of a former facility plant. Drilling locations were determined based on historical data provided by the site operator. The intention was to gather soil samples across a range of concentrations. Because the surface soils were from relatively clean fill, the sampling device was pushed to a depth of 3.65 m using a blank rod. Samples were then collected at pre-selected depths ranging from 3.65 to 9 m bgs. Individual cores were 1-m long. The plastic sleeve for each 1-m core was marked with a permanent marker; the depth interval and the bottom of each core was marked. The filled plastic tubes were transferred to a staging table where appropriate depth intervals were selected for mixing. Selected tubes were cut into 0.6-m intervals, which were emptied into a plastic container for premixing soils. When feasible, soils were initially screened to remove materials larger than 6.3-mm in diameter. In many cases, soils were too wet and clayey to allow screening; in these cases, the soil was broken into pieces by hand and, by using a wooden spatula, oversize materials were manually removed. These soils (screened or hand sorted) were then mixed until the soil appeared visually uniform in color and texture. The mixed soil was then placed into a 4-L sample container for each chosen sample interval. A subsample of the mixed soil was transferred into a 20-mL vial, and it was sent for quick turnaround mercury analysis. This process was repeated for each subsequent sample interval.

3.4 Puget Sound

3.4.1 Site Description

The Puget Sound site consists of contaminated offshore sediments. The particular area of the site used for collecting demonstration samples is identified as the Georgia Pacific, Inc. Log Pond. The Log Pond is located within the Whatcom Waterway in Bellingham Bay, WA, a

well-established heavy industrial land use area with a maritime shoreline designation. Log Pond sediments measure approximately 1.5 to 1.8-m thick, and contain various contaminants including mercury, phenols, polyaromatic hydrocarbons, polychlorinated biphenyls, and wood debris. Mercury was used as a preservative in the logging industry. The area was capped in late 2000 and early 2001 with an average of 7 feet of clean capping material, as part of a Model Toxics Control Act interim cleanup action. The total thickness ranges from approximately 0.15 m along the site perimeter to 3 m within the interior of the project area. The restoration project produced 2.7 acres of shallow sub-tidal and 2.9 acres of low intertidal habitat, all of which had previously exceeded the Sediment Management Standards cleanup criteria (Anchor Environmental, 2001).

Mercury concentrations have been measured ranging from 0.16 to 400 mg/kg (dry wt). The majority (98%) of the mercury detected in near-shore ground waters and sediments of the Log Pond is believed to be comprised of complexed divalent (Hg^{2+}) forms such as mercuric sulfide (Bothner, et al., 1980 and Anchor Environmental, 2000).

3.4.2 Sample Collection

Science Applications International Corporation (SAIC) is currently performing a SITE remedial technology evaluation in the Puget Sound (SAIC, 2002). As part of ongoing work at that site, SAIC collected additional sediment for use during this MMT project. Sediment samples collected on August 20-21, 2002 from the Log Pond in Puget Sound were obtained beneath approximately 3-6 m of water, using a vibra-coring system capable of capturing cores to 0.3 m below the proposed dredging prism. The vibra-corer consisted of a core barrel attached to a power head. Aluminum core tubes, equipped with a stainless steel "eggshell" core catcher to retain material, were inserted into the core barrel. The vibra-core was lowered into position on the bottom and advanced to the appropriate sampling depth. Once sampling was completed, the vibra-core was retrieved and the core liner removed from the core barrel. The core sample was examined at each end to verify that sufficient sediment was retained for the particular sample. The condition and quantity of material within the core was then inspected to determine acceptability.

The following criteria were used to verify whether an acceptable core sample was collected:

- Target penetration depth (i.e., into native material) was achieved.

- Sediment recovery of at least 65% of the penetration depth was achieved.
- Sample appeared undisturbed and intact without any evidence of obstruction/blocking within the core tube or catcher.

The percent sediment recovery was determined by dividing the length of material recovered by the depth of core penetration below the mud line. If the sample was deemed acceptable, overlying water was siphoned from the top of the core tube and each end of the tube capped and sealed with duct tape. Following core collection, representative samples were collected from each core section representing a different vertical horizon. Sediment was collected from the center of the core that had not been smeared by, or in contact with, the core tube. The volumes removed were placed in a decontaminated stainless steel bowl or pan and mixed until homogenous in texture and color (approximately 2 minutes).

After all sediment for a vertical horizon composite was collected and homogenized, representative aliquots were placed in the appropriate pre-cleaned sample containers. Samples of both the sediment and the underlying native material were collected in a similar manner. Distinct layers of sediment and native material were easily recognizable within each core.

3.5 Demonstration Site

The demonstration was conducted in a natural environment, outdoors, in Oak Ridge, TN. The area was a grass covered hill with some parking areas, all of which were surrounded by trees. Building 5507, in the center of the demonstration area, provided facilities for lunch, break, and sample storage for the project and personnel.

Most of the demonstration was performed during rainfall events ranging from steady to torrential. Severe puddling of rain occurred to the extent that boards needed to be placed under chairs to prevent them from sinking into the ground. Even when it was not raining, the relative humidity was high, ranging from 70.6 to 98.3 percent. Between two and four of the tent sides were used to keep rainfall from damaging the instruments. The temperature in the afternoons ranged from 65-70 degrees Fahrenheit, and the wind speed was less than 10 mph. The latitude is 36°N, the longitude 35°W, and the elevation 275 m. (Figure 3-1 is a photograph of the site during the demonstration and Figure 3-2 is a photograph of the location.)



Figure 3-1. Tent and field conditions during the demonstration at Oak Ridge, TN.



Figure 3-2. Demonstration site and Building 5507.

3.6 SAIC GeoMechanics Laboratory

Sample homogenization was completed at the SAIC GeoMechanics Laboratory in Las Vegas, NV. This facility is an industrial-type building with separate facilities for personnel offices and material handling. The primary function of the laboratory is for rock mechanics studies. The laboratory has rock mechanics equipment, including

sieves, rock crushers, and sample splitters. The personnel associated with this laboratory are experienced in the areas of sample preparation and sample homogenization. In addition to the sample homogenization equipment, the laboratory contains several benches, tables, and open space. Mercury air monitoring equipment was used during the sample preparation activities for personnel safety.

Chapter 4 Demonstration Approach

This chapter describes the demonstration approach that was used for evaluating the field mercury measurement technologies at ORNL in May 2003. It presents the objectives, design, sample preparation and management procedures, and the reference method confirmatory process used for the demonstration.

4.1 Demonstration Objectives

The primary goal of the SITE MMT Program is to develop reliable performance and cost data on innovative, field-ready measurement technologies. A SITE

demonstration must provide detailed and reliable performance and cost data in order that potential technology users have adequate information needed to make sound judgements regarding an innovative technology's applicability to a specific site, and to be able to compare the technology to conventional technologies.

Table 4-1 summarizes the project objectives for this demonstration. In accordance with *QAPP Requirements for Applied Research Projects (EPA, 1998)*, the technical project objectives for the demonstration were categorized as primary and secondary.

Table 4-1. Demonstration Objectives

Objective	Description	Method of Evaluation
Primary Objectives		
Primary Objective # 1	Determine sensitivity of each instrument with respect to vendor-generated MDL and PQL.	Independent laboratory confirmation of SRMs, field samples, and spiked field samples.
Primary Objective # 2	Determine potential analytical accuracy associated with vendor field measurements.	
Primary Objective # 3	Evaluate the precision of vendor field measurements.	
Primary Objective # 4	Measure time required to perform five functions related to mercury measurements: 1) mobilization and setup, 2) initial calibration, 3) daily calibration, 4) sample analysis, and 5) demobilization.	Documentation during demonstration; vendor-provided information.
Primary Objective # 5	Estimate costs associated with mercury measurements for the following four categories: 1) capital, 2) labor, 3) supplies, and 4) investigation-derived wastes.	
Secondary Objectives		
Secondary Objective # 1	Document ease of use, skills, and training required to operate the device properly.	Documentation of observations during demonstration; vendor-provided information.
Secondary Objective # 2	Document potential H&S concerns associated with operating the device.	
Secondary Objective # 3	Document portability of the device.	
Secondary Objective # 4	Evaluate durability of device based on materials of construction and engineering design.	
Secondary Objective # 5	Document the availability of the device and its spare parts.	Post-demonstration investigation.

Critical data support primary objectives and noncritical data support secondary objectives. With the exception of the cost information, primary objectives required the use of quantitative results to draw conclusions regarding technology performance. Secondary objectives pertained to information that was useful and did not necessarily require the use of quantitative results to draw conclusions regarding technology performance.

4.2 Demonstration Design

4.2.1 Approach for Addressing Primary Objectives

The purpose of this demonstration was to evaluate the performance of the vendor's instrumentation against a standard laboratory procedure. In addition, an overall average relative standard deviation (RSD) was calculated for all measurements made by the vendor and the referee laboratory. RSD comparisons used descriptive statistics, not inferential statistics, between the vendor and laboratory results. Other statistical comparisons (both inferential and descriptive) for sensitivity, precision, and accuracy were used, depending upon actual demonstration results.

The approach for addressing each of the primary objectives is discussed in the following subsections. A detailed explanation of the precise statistical determination used for evaluating primary objectives No. 1 through No. 3 is presented in Chapter 6.

4.2.1.1 Primary Objective #1: Sensitivity

Sensitivity is the ability of a method or instrument to discriminate between small differences in analyte concentration (EPA, 2002b). It can be discussed in terms of an instrument detection limit (IDL), a method detection limit (MDL), and as a practical quantitation limit (PQL). MDL is not a measure of sensitivity in the same respect as an IDL or PQL. It is a measure of precision at a predetermined, usually low, concentration. The IDL pertains to the ability of the instrument to determine with confidence the difference between a sample that contains the analyte of interest at a low concentration and a sample that does not contain that analyte. The IDL is generally considered to be the minimum true concentration of an analyte producing a non-zero signal that can be distinguished from the signals generated when no concentration of the analyte is present and with an adequate degree of certainty.

The IDL is not rigidly defined in terms of matrix, method, laboratory, or analyst variability, and it is not usually

associated with a statistical level of confidence. IDLs are, thus, usually lower than MDLs and rarely serve a purpose in terms of project objectives (EPA, 2002b). The PQL defines a specific concentration with an associated level of accuracy. The MDL defines a lower limit at which a method measurement can be distinguished from background noise. The PQL is a more meaningful estimate of sensitivity. The MDL and PQL were chosen as the two distinct parameters for evaluating sensitivity. The approach for addressing each of these indicator parameters is discussed separately in the following paragraphs.

MDL

MDL is the estimated measure of sensitivity as defined in 40 Code of Federal Regulations (CFR) Part 136. The purpose of the MDL measurement is to estimate the concentration at which an individual field instrument is able to detect a minimum concentration that is statistically different from instrument background or noise. Guidance for the definition of the MDL is provided in EPA G-5i (EPA, 2002b).

The determination of an MDL usually requires seven different measurements of a low concentration standard or sample. Following procedures established in 40 CFR Part 136 for water matrices, the demonstration MDL definition is as follows:

$$MDL = t_{(n-1, 0.99)} s$$

where: $t_{(n-1, 0.99)}$ = 99th percentile of the t-distribution with n-1 degrees of freedom
n = number of measurements
s = standard deviation of replicate measurements

PQL

The PQL is another important measure of sensitivity. The PQL is defined in EPA G-5i as the lowest level an instrument is capable of producing a result that has significance in terms of precision and bias. (Bias is the difference between the measured value and the true value.) It is generally considered the lowest standard on the instrument calibration curve. It is often 5-10 times higher than the MDL, depending upon the analyte, the instrument being used, and the method for analysis; however, it should not be rigidly defined in this manner.

During the demonstration, the PQL was to be defined by the vendor's reported calibration or based upon lower concentration samples or SRMs. The evaluation of vendor-reported results for the PQL included a determination of the percent difference (%D) between their calculated value and true value. The true value is considered the value reported by the referee laboratory for field samples or spiked field samples, or, in the case of SRMs, the certified value provided by the supplier. The equation used for the %D calculation is:

$$\%D = \frac{|C_{\text{true}} - C_{\text{calculated}}|}{C_{\text{true}}} \times 100$$

where: C_{true} = true concentration as determined by the referee laboratory or SRM reference value
 $C_{\text{calculated}}$ = calculated test sample concentration

The PQL and %D were reported for the vendor. The %D for the referee laboratory, at the same concentration, was also reported for purposes of comparison. No statistical comparison was made between these two values; only a descriptive comparison was made for purposes of this evaluation. (The %D requirement for the referee laboratory was defined as 10% or less. The reference method PQL was approximately 10 µg/kg.)

4.2.1.2 Primary Objective #2: Accuracy

Accuracy was calculated by comparing the measured value to a known or true value. For purposes of this demonstration, three separate standards were used to evaluate accuracy. These included: 1) SRMs, 2) field samples collected from four separate mercury-contaminated sites, and 3) spiked field samples. Four sites were to be used for the evaluation; however, the manufacturing site samples proved to be too high in concentration for the Milestone Field instrument (above 10 mg/kg) and therefore were not analyzed. Samples representing field samples and spiked field samples were prepared at the SAIC GeoMechanics Laboratory. In order to prevent cross contamination, SRMs were prepared in a separate location. Each of these standards is discussed separately in the following paragraphs.

SRMs

The primary standards used to determine accuracy for this demonstration were SRMs. SRMs provided very tight

statistical comparisons, although they did not provide all matrices of interest nor all ranges of concentrations. The SRMs were obtained from reputable suppliers, and had reported concentrations at associated 95% confidence intervals (CIs), and 95% prediction intervals. Prediction intervals were used for comparison because they represent a statistically infinite number of analyses, and therefore, would include all possible correct results 95% of the time. All SRMs were analyzed by the referee laboratory and selected SRMs were analyzed by the vendor, based upon instrument capabilities and concentrations of SRMs that could be obtained. Selected SRMs covered an appropriate range for each vendor. Replicate SRMs were also analyzed by the vendor and the laboratory.

The purpose for SRM analysis by the referee laboratory was to provide a check on laboratory accuracy. During the pre-demonstration, the referee laboratory was chosen, in part, based upon the analysis of SRMs. This was done to ensure a competent laboratory would be used for the demonstration. Because of the need to provide confidence in laboratory analysis during the demonstration, the referee laboratory analyzed SRMs as an ongoing check for laboratory bias.

Evaluation of vendor and laboratory analysis of SRMs was performed as follows. Accuracy was reported for individual sample concentrations of replicate measurements made at the same concentration.

Two-tailed 95% CIs were computed according to the following equation:

$$\bar{x} \pm t_{(n-1, 0.975)} \cdot s / \sqrt{n}$$

where: $t_{(n-1, 0.975)}$ = 97.5th percentile of the t-distribution with n-1 degrees of freedom
n = number of measurements
s = standard deviation of replicate measurements

The number of vendor-reported SRM results and referee laboratory-reported SRM results that were within the associated 95% prediction interval were evaluated. Prediction intervals were computed in a similar fashion to the CI, except that the Student's "t" value use "n" equal to infinity and, because they represented "n" approaching infinity, the square root of "n" is dropped from the equation.

A final measure of accuracy determined from SRMs is a frequency distribution that shows the percentage of vendor-

reported measurements that are within a specified window of the reference value. For example, a distribution within a 30% window of a reported concentration, within a 50% window, and outside a 50% window of a reported concentration. This distribution aspect could be reported as average concentrations of replicate results from the vendor for a particular concentration and matrix compared to the same sample from the laboratory. These are descriptive statistics and are used to better describe comparisons, but they are not intended as inferential tests.

Field Samples

The second accuracy standard used for this demonstration was actual field samples collected from four separate mercury-contaminated sites. (Only 3 of the 4 sites were used for the Milestone evaluation.) This accuracy determination consisted of a comparison of vendor-reported results for field samples to the referee laboratory results for the same field samples. The field samples were used to ensure that "real-world" samples were tested for each vendor. The field samples consisted of variable mercury concentrations within varying soil and sediment matrices. The referee laboratory results are considered the standard for comparison to each vendor.

Vendor sample results for a given field sample were compared to replicates analyzed by the laboratory for the same field sample. (A hypothesis test was used with $\alpha = 0.01$. The null hypothesis was that sample results were similar. Therefore, if the null hypothesis is rejected, then the sample sets are considered different.) Comparisons for a specific matrix or concentration were made in order to provide additional information on that specific matrix or concentration. Comparison of the vendor values to laboratory values were similar to the comparisons noted previously for SRMs, except that a more definitive or inferential statistical evaluation was used. $\alpha = 0.01$ was used to help mitigate inter-laboratory variability. Additionally, an aggregate analysis was used to mitigate statistical anomalies (see Section 6.1.2).

Spiked Field Samples

The third accuracy standard for this demonstration was spiked field samples. These spiked field samples were analyzed by the vendors and by the referee laboratory in replicate in order to provide additional measurement comparisons to a known value. Spikes were prepared to cover additional concentrations not available from SRMs or the samples collected in the field. They were grouped with the field sample comparison noted above.

4.2.1.3 Primary Objective #3: Precision

Precision can be defined as the degree of mutual agreement of independent measurements generated through repeated application of a process under specified conditions. Precision is usually thought of as repeatability of a specific measurement, and it is often reported as RSD. The RSD is computed from a specified number of replicates. The more replications of a measurement, the more confidence is associated with a reported RSD. Replication of a measurement may be as few as 3 separate measurements to 30 or more measurements of the same sample, dependent upon the degree of confidence desired in the specified result. The precision of an analytical instrument may vary depending upon the matrix being measured, the concentration of the analyte, and whether the measurement is made for an SRM or a field sample.

The experimental design for this demonstration included a mechanism to evaluate the precision of the vendors' technologies. Field samples from the four mercury-contaminated field sites were evaluated by each vendor's analytical instrument. (See previous note concerning Milestone.) During the demonstration, concentrations were predetermined only as low, medium, or high. Ranges of test samples (field samples, SRMs, and spikes) were selected to cover the appropriate analytical ranges of the vendor's instrumentation. It was known prior to the demonstration that not all vendors were capable of measuring similar concentrations (i.e., some instruments were better at measuring low concentrations and others were geared toward higher concentration samples or had other attributes such as cost or ease of use that defined specific attributes of their technology). Because of this fact, not all vendors analyzed the same samples.

During the demonstration, the vendor's instrumentation was tested with samples from the four different sites, having different matrices when possible (i.e., depending upon available concentrations) and having different concentrations (high, medium, and low) using a variety of samples. Sample concentrations for an individual instrument were chosen based upon vendor attributes in terms of expected low, medium, and high concentrations that the particular instrument was capable of measuring.

The referee laboratory measured replicates of all samples. The results were used for purposes of precision comparisons to the individual vendor. The RSD for the vendor and the laboratory were calculated individually, using the following equation:

$$\%RSD = \frac{s}{\bar{x}} \times 100$$

where: s = standard deviation of replicate results
 \bar{x} = mean value of replicate results

Using descriptive statistics, differences between vendor RSD and referee laboratory RSD were determined. This included RSD comparisons based upon concentration, SRMs, field samples, and different sites. In addition, an overall average RSD was calculated for all measurements made by the vendor and the laboratory. RSD comparisons were based upon descriptive statistical evaluations between the vendor and the laboratory, and results were compared accordingly.

4.2.1.4 Primary Objective #4: Time per Analysis

The amount of time required for performing the analysis was measured and reported for five categories:

- Mobilization and setup
- Initial calibration
- Daily calibration
- Sample analyses
- Demobilization

Mobilization and setup included the time to unpack and prepare the instrument for operation. Initial calibration included the time to perform the vendor recommended on-site calibrations. Daily calibration included the time to perform the vendor-recommended calibrations on subsequent field days. (Note that this could have been the same as the initial calibration, a reduced calibration, or none.) Sample analyses included the time to prepare, measure, and calculate the results for the demonstration samples, and the necessary quality control (QC) samples performed by the vendor.

The time per analysis was determined by dividing the total amount of time required to perform the analyses by the number of samples analyzed (173). In the numerator, sample analysis time included preparation, measurement, and calculation of results for demonstration samples and necessary QC samples performed by the vendor. In the denominator, the total number of analyses included only demonstration samples analyzed by the vendor, not QC analyses nor reanalyses of samples.

Downtime that was required or that occurred between sample analyses as a part of operation and handling was considered a part of the sample analysis time. Downtime

occurring due to instrument breakage or unexpected maintenance was not counted in the assessment, but it is noted in this final report as an additional time. Any downtime caused by instrument saturation or memory effect was addressed, based upon its frequency and impact on the analysis.

Unique time measurements are also addressed in this report (e.g., if soil samples were analyzed directly, and sediment samples required additional time to dry before the analyses started, then a statement was made noting that soil samples were analyzed in X amount of hours, and that sediment samples required drying time before analysis).

Recorded times were rounded to the nearest 15-minute interval. The number of vendor personnel used was noted and factored into the time calculations. No comparison on time per analysis is made between the vendor and the referee laboratory.

4.2.1.5 Primary Objective #5: Cost

The following four cost categories were considered to estimate costs associated with mercury measurements:

- Capital costs
- Labor costs
- Supply costs
- Investigation derived waste (IDW) disposal costs

Although both vendor and laboratory costs are presented, the calculated costs were not compared with the referee laboratory. A summary of how each cost category was estimated for the measurement device is provided below:

- The capital cost was estimated based on published price lists for purchasing, renting, or leasing each field measurement device. If the device was purchased, the capital cost estimate did not include salvage value for the device after work was completed.
- The labor cost was based on the number of people required to analyze samples during the demonstration. The labor rate was based on a standard hourly rate for a technician or other appropriate operator. During the demonstration, the skill level required was confirmed based on vendor input regarding the operation of the device to produce mercury concentration results and observations made in the field. The labor costs were based on: 1) the actual number of hours required to complete all analyses, quality assurance (QA), and reporting; and 2) the assumption that a technician who worked for a portion of a day was paid for an entire 8-hour day.

- The supply costs were based on any supplies required to analyze the field and SRM samples during the demonstration. Supplies consisted of items not included in the capital category, such as extraction solvent, glassware, pipettes, spatulas, agitators, and similar materials. The type and quantity of all supplies brought to the field and used during the demonstration were noted and documented.

Any maintenance and repair costs during the demonstration were documented or provided by the vendor. Equipment costs were estimated based on this information and standard cost analysis guidelines for the SITE Program.

- The IDW disposal costs included decontamination fluids and equipment, mercury-contaminated soil and sediment samples, and used sample residues. Contaminated personal protective equipment (PPE) normally used in the laboratory was placed into a separate container. The disposal costs for the IDW were included in the overall analytical costs for each vendor.

After all of the cost categories were estimated, the cost per analysis was calculated. This cost value was based on the number of analyses performed. As the number of samples analyzed increased, the initial capital costs and certain other costs were distributed across a greater number of samples. Therefore, the per unit cost decreased. For this reason, two costs were reported: 1) the initial capital costs and 2) the operating costs per analysis. No comparison to the referee laboratory's method cost was made; however,

a generic cost comparison is made. Additionally, when determining laboratory costs, the associated cost for laboratory audits and data validation should be considered.

4.2.2 Approach for Addressing Secondary Objectives

Secondary objectives were evaluated based on observations made during the demonstration. Because of the number of vendors involved, technology observers were required to make simultaneous observations of two vendors each during the demonstration. Four procedures were implemented to ensure that these subjective observations made by the observers were as consistent as possible.

First, forms were developed for each of the five secondary objectives. These forms assisted in standardizing the observations. Second, the observers met each day before the evaluations began, at significant break periods, and after each day of work to discuss and compare observations regarding each device. Third, an additional observer was assigned to independently evaluate only the secondary objectives; in order to ensure that a consistent approach was applied in evaluating these objectives. Finally, the SAIC TOM circulated among the evaluation staff during the demonstration to ensure that a consistent approach was being followed by all personnel. Table 4-2 summarizes the aspects observed during the demonstration for each secondary objective. The individual approaches to each of these objectives are detailed further in the following subsections.

Table 4-2. Summary of Secondary Objective Observations Recorded During the Demonstration

SECONDARY OBJECTIVE				
General Information	Secondary Objective # 1 Ease of Use	Secondary Objective # 2 H&S Concerns	Secondary Objective # 3 Instrument Portability	Secondary Objective # 4 Instrument Durability
- Vendor Name	- No. of Operators	- Instrument Certifications	- Instrument Weight	- Materials of Construction
- Observer Name	- Operator Names/Titles	- Electrical Hazards	- Instrument Dimensions	- Quality of Construction
- Instrument Type	- Operator Training	- Chemicals Used	- Power Sources	- Max. Operating Temp.
- Instrument Name	- Training References	- Radiological Sources	- Packaging	- Max. Operating Humidity
- Model No.	- Instrument Setup Time	- Hg Exposure Pathways	- Shipping & Handling	- Downtime
- Serial No.	- Instrument Calibration Time	- Hg Vapor Monitoring		- Maintenance Activities
	- Sample Preparation Time	- PPE Requirements		- Repairs Conducted
	- Sample Measurement Time	- Mechanical Hazard		
		- Waste Handling Issues		
H&S =	Health and Safety			
PPE =	Personal Protective Equipment			

4.2.2.1 Secondary Objective #1: Ease of Use

The skills and training required for proper device operation were noted; these included any degrees or specialized training required by the operators. This information was gathered by interviews (i.e., questioning) of the operators. The number of operators required was also noted. This objective was also evaluated by subjective observations regarding the ease of equipment use and major peripherals required to measure mercury concentrations in soils and sediments. The operating procedure was evaluated to determine if the instrument is easily useable and understandable.

4.2.2.2 Secondary Objective #2: Health and Safety Concerns

Health and safety (H&S) concerns associated with device operation were noted during the demonstration. Criteria included hazardous materials used, the frequency and likelihood of potential exposures, and any direct exposures observed during the demonstration. In addition, any potential for exposure to mercury during sample digestion and analysis was evaluated based upon equipment design. Other H&S concerns, such as basic electrical and mechanical hazards, were also noted. Equipment certifications, such as Underwriters Laboratory (UL), were documented.

4.2.2.3 Secondary Objective #3: Portability of the Device

The portability of the device was evaluated by observing transport, measuring setup and tear down time, determining the size and weight of the unit and peripherals, and assessing the ease with which the instrument was repackaged for movement to another location. The use of battery power or the need for an AC outlet was also noted.

4.2.2.4 Secondary Objective #4: Instrument Durability

The durability of each device and major peripherals was assessed by noting the quality of materials and construction. All device failures, routine maintenance, repairs, and downtime were documented during the demonstration. No specific tests were performed to evaluate durability; rather, subjective observations were made using a field form as guidance.

4.2.2.5 Secondary Objective #5: Availability of Vendor Instruments and Supplies

The availability of each device was evaluated by determining whether additional units and spare parts are readily available from the vendor or retail stores. The

vendor's office (or a web page) and/or a retail store was contacted to identify and determine the availability of supplies of the tested measurement device and spare parts. This portion of the evaluation was performed after the field demonstration, in conjunction with the cost estimate.

4.3 Sample Preparation and Management

4.3.1 Sample Preparation

4.3.1.1 Field Samples

Field samples were collected during the pre-demonstration portion of the project, with the ultimate goal of producing a set of consistent test soils and sediments to be distributed among all participating vendors and the referee laboratory for analysis during the demonstration. Samples were collected from the following four sites:

- Carson River Mercury site (near Dayton, NV)
- Y-12 National Security Complex (Oak Ridge, TN)
- Manufacturing facility (eastern U.S.)
- Puget Sound (Bellingham, WA)

The field samples collected during the pre-demonstration sampling events comprised a variety of matrices, ranging from material having a high clay content to material composed mostly of gravelly, coarse sand. The field samples also differed with respect to moisture content; several were collected as wet sediments. Table 4-3 shows the number of distinct field samples that were collected from each of the four field sites.

Prior to the start of the demonstration, the field samples selected for analysis during the demonstration were prepared at the SAIC GeoMechanics Laboratory in Las Vegas, NV. The specific sample homogenization procedure used by this laboratory largely depended on the moisture content and physical consistency of the sample. Two specific sample homogenization procedures were developed and tested by SAIC at the GeoMechanics Laboratory during the pre-demonstration portion of the project. The methods included a non-slurry sample procedure and a slurry sample procedure.

A standard operating procedure (SOP) was developed detailing both methods. The procedure was found to be satisfactory based upon the results of replicate samples during the pre-demonstration. This SOP is included as Appendix A of the *Field Demonstration Quality Assurance Project Plan* (SAIC, August 2003 EPA/600/R-03/053). Figure 4-1 summarizes the homogenization steps of the

SOP, beginning with sample mixing. This procedure was used for preparing both pre-demonstration and demonstration samples. Prior to the mixing process (i.e., Step 1 in Figure 4-1), all field samples being processed were visually inspected to ensure that oversized materials were removed and that there were no clumps that would hinder homogenization. Non-slurry samples were air-dried in accordance with the SOP so that they could be passed

multiple times through a riffle splitter. Due to the high moisture content of many of the samples, they were not easily air-dried and could not be passed through a riffle splitter while wet. Samples with very high moisture contents, termed “slurries,” were not air-dried, and bypassed the riffle splitting step. The homogenization steps for each type of matrix are briefly summarized as follows.

Table 4-3. Field Samples Collected from the Four Sites

Field Site	No. of Samples / Matrices Collected	Areas For Collecting Sample Material	Volume Required
Carson River	12 Soil 6 Sediment	Tailings Piles (Six Mile Canyon) River Bank Sediments	4 L each for soil 12 L each for sediment
Y-12	10 Sediment 6 Soil	Poplar Creek Sediments Old Mercury Recovery Bldg. Soils	12 L each for sediment 4 L each for soil
Manufacturing Site	12 Soil	Subsurface Soils	4 L each
Puget Sound	4 Sediment	High-Level Mercury (below cap) Low-Level Mercury (native material)	12 L each

Preparing Slurry Matrices

For slurries (i.e., wet sediments), the mixing steps were sufficiently thorough that the sample containers could be filled directly from the mixing vessel. There were two separate mixing steps of the slurry-type samples. Each slurry was initially mixed mechanically within the sample container (i.e., bucket) in which the sample was shipped to the SAIC GeoMechanics Laboratory. A subsample of this premixed sample was transferred to a second mixing vessel. A mechanical drill equipped with a paint mixing attachment was used to mix the subsample. As shown in Figure 4-1, slurry samples bypassed the sample riffle splitting step. To ensure all sample bottles contained the same material, the entire set of containers to be filled was submerged into the slurry as a group. The filled vials were allowed to settle for a minimum of two days, and the standing water was removed using a Pasteur pipette. The removal of the standing water from the slurry samples was the only change to the homogenization procedure between the pre-demonstration and the demonstration.

Preparing "Non-Slurry" Matrices

Soils and sediments having no excess moisture were initially mixed (Step 1) and then homogenized in the sample riffle splitter (Step 2). Prior to these steps, the material was air-dried and subsampled to reduce the volume of material to a size that was easier to handle.

As shown in Figure 4-1 (Step 1), the non-slurry subsample was manually stirred with a spoon or similar equipment until the material was visually uniform. Immediately following manual mixing, the subsample was mixed and split six times for more complete homogenization (Step 2). After the 6th and final split, the sample material was leveled to form a flattened, elongated rectangle and cut into transverse sections to fill the containers (Steps 3 and 4). After homogenization, 20-ml sample vials were filled and prepared for shipment (Step 5).

For the demonstration, the vendor analyzed 173 samples, which included replicates of up to 7 samples per sample lot. The majority of the samples distributed had concentrations within the range of the vendor’s technology. Some samples had expected concentrations at or below the estimated level of detection for each of the vendor instruments. These samples were designed to evaluate the reported MDL and PQL and also to assess the prevalence of false positives. Field samples distributed to the vendor included sediments and soils collected from all four sites and prepared by both the slurry and dry homogenization procedures. The field samples were segregated into broad sample sets: low, medium, and high mercury concentrations. This gave the vendor the same general understanding of the sample to be analyzed as they would typically have for field application of their instrument.

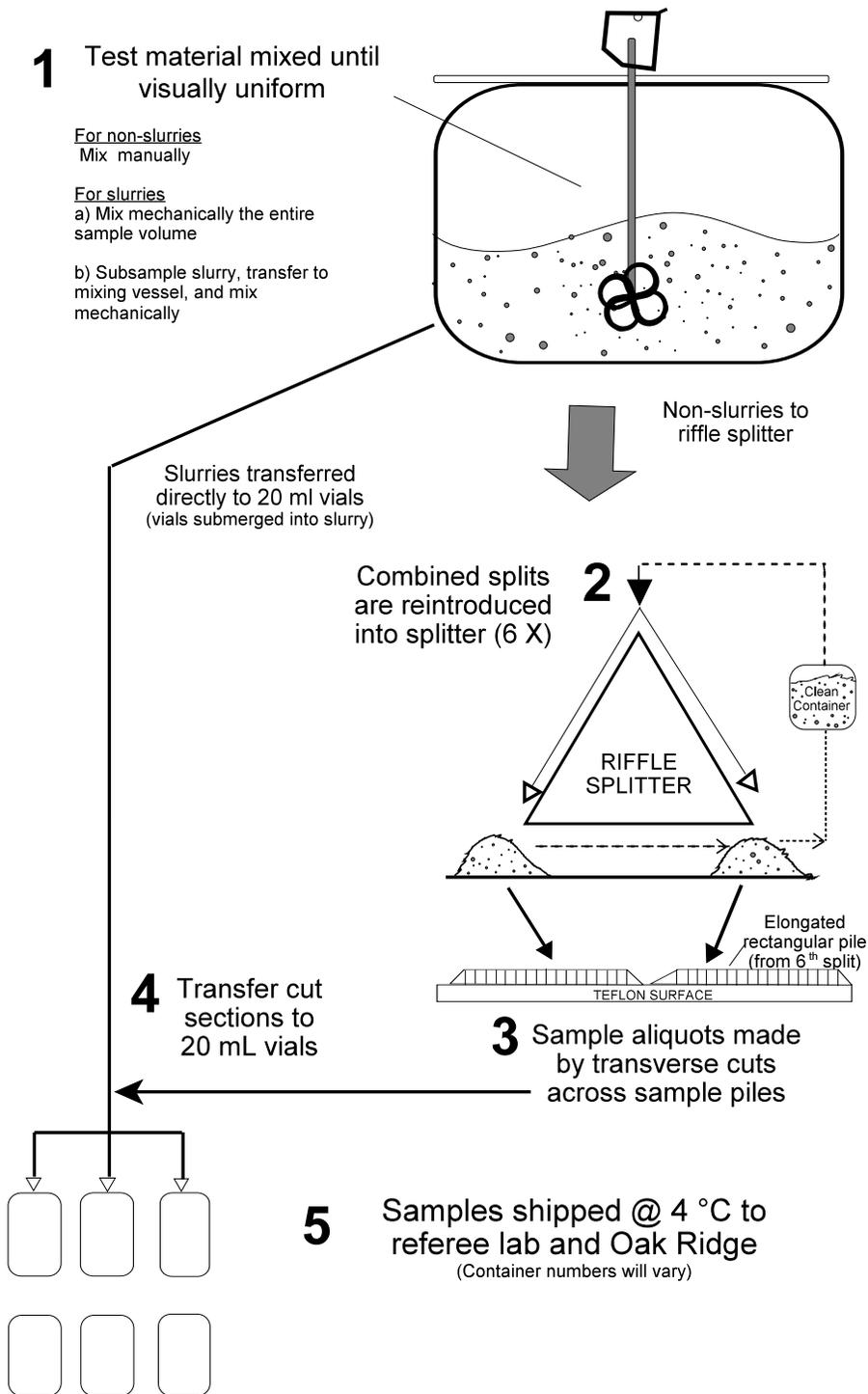


Figure 4-1. Test sample preparation at the SAIC GeoMechanics Laboratory.

In addition, selected field samples were spiked with mercury (II) chloride to generate samples with additional concentrations and test the ability of the vendor's instrumentation to measure the additional species of mercury. Specific information regarding the vendor's sample distribution is included in Chapter 6.

4.3.1.2 Standard Reference Materials

Certified SRMs were analyzed by both the vendors and the referee laboratory. These samples were homogenized matrices which had a known concentration of mercury. Concentrations were certified values, as provided by the supplier, based on independent confirmation via multiple analyses of multiple lots and/or multiple analyses by different laboratories (i.e., round robin testing). These analytical results were then used to determine "true" values, as well as a statistically derived intervals (a 95% prediction interval) that provided a range within which the true value were expected to fall.

The SRMs selected were designed to encompass the same contaminant ranges indicated previously: low-, medium-, and high-level mercury concentrations. In addition, SRMs of varying matrices were included in the demonstration to challenge the vendor technology as well as the referee laboratory. The referee laboratory analyzed all SRMs. SRM samples were intermingled with site field samples and labeled in the same manner as field samples.

4.3.1.3 Spiked Field Samples

Spiked field samples were prepared by the SAIC GeoMechanics Laboratory using mercury (II) chloride. Spikes were prepared using field samples from the selected sites. Additional information was gained by preparing spikes at concentrations not previously obtainable. The SAIC GeoMechanics Laboratory's ability to prepare spikes was tested prior to the demonstration and evaluated in order to determine expected variability and accuracy of the spiked sample. The spiking procedure was evaluated by preparing several different spikes using two different spiking procedures (dry and wet). Based upon replicate analyses results, it was determined that the wet, or slurry, procedure was the only effective method of obtaining a homogeneous spiked sample.

4.3.2 Sample Management

4.3.2.1 Sample Volumes, Containers, and Preservation

A subset from the pre-demonstration field samples was selected for use in the demonstration based on the sample's mercury concentration range and sample type

(i.e., sediment versus soil). The SAIC GeoMechanics Laboratory prepared individual batches of field sample material to fill sample containers for each vendor. Once all containers from a field sample were filled, each container was labeled and cooled to 4 °C. Because mercury analyses were to be performed both by the vendors in the field and by the referee laboratory, adequate sample size was taken into account. Minimum sample size requirements for the vendors varied from 0.1 g or less to 8-10 g. Only the referee laboratory analyzed separate sample aliquots for parameters other than mercury. These additional parameters included arsenic, barium, cadmium, chromium, lead, selenium, silver, copper, zinc, oil and grease, and total organic carbon (TOC). Since the mercury method (SW-846 7471B) being used by the referee laboratory requires 1 g for analysis, the sample size sent to all participants was a 20 mL vial (approximately 10 g), which ensured a sufficient volume and mass for analysis by all vendors.

4.3.2.2 Sample Labeling

The sample labeling used for the 20 mL vials consisted of an internal code developed by SAIC. This "blind" code was used throughout the entire demonstration. The only individuals who knew the key coding of the homogenized samples to the specific field samples were the SAIC TOM, the SAIC GeoMechanics Laboratory Manager, and the SAIC QA Manager.

4.3.2.3 Sample Record Keeping, Archiving, and Custody

Samples were shipped to the laboratory and the demonstration site the week prior to the demonstration. A third set of vials was archived at the SAIC GeoMechanics Laboratory as reserve samples.

The sample shipment to Oak Ridge was retained at all times in the custody of SAIC at their Oak Ridge office and until arrival of the demonstration field crew. Samples were shipped under chain-of-custody (CoC) and with custody seals on both the coolers and the inner plastic bags. Once the demonstration crew arrived, the coolers were retrieved from the SAIC office. The custody seals on the plastic bags inside the cooler were broken by the vendor upon transfer.

Upon arrival at the ORNL site, the vendor set up the instrumentation at the direction and oversight of SAIC. At the start of sample testing, the vendor was provided with a sample set representing field samples collected from a particular field site intermingled with SRM and spiked samples. Due to variability of vendor instrument

measurement ranges for mercury detection, not all vendors received samples from the same field material. All samples were stored in an ice cooler prior to demonstration startup and were stored in an on-site sample refrigerator during the demonstration. Each sample set was identified and distributed as a set with respect to the site from which it was collected. This was done because, in any field application, the location and general type of the samples would be known.

The vendor was responsible for analyzing all samples provided, performing any dilutions or reanalyses as needed, calibrating the instrument if applicable, performing any necessary maintenance, and reporting all results. Any samples that were not analyzed during the day were returned to the vendor for analysis at the beginning of the next day. Once analysis of the samples from the first location were completed by the vendor, SAIC provided a set of samples from the second location. Samples were provided at the time that they were requested by the vendor. Once again, the transfer of samples was documented using a COC form.

This process was repeated for samples from each location. SAIC maintained custody of all remaining sample sets until they were transferred to the vendor. SAIC maintained custody of samples that already had been analyzed and followed the waste handling procedures in Section 4.2.2 of the Field Demonstration QAPP to dispose of these wastes.

4.4 Reference Method Confirmatory Process

The referee laboratory analyzed all samples that were analyzed by the vendor technologies in the field. The following subsections provide information on the selection of the reference method, selection of the referee laboratory, and details regarding the performance of the reference method in accordance with EPA protocols. Other parameters that were analyzed by the referee laboratory are also discussed briefly.

4.4.1 Reference Method Selection

The selection of SW-846 Method 7471B as the reference method was based on several factors, predicated on information obtained from the technology vendors, as well as the expected contaminant types and soil/sediment mercury concentrations expected in the test matrices. There are several laboratory-based, promulgated methods for the analysis of total mercury. In addition, there are several performance-based methods for the determination

of various mercury species. Based on the vendor technologies, it was determined that a reference method for total mercury would be needed (Table 1-2 summarizes the methods evaluated, as identified through a review of the EPA Test Method Index and SW-846).

In selecting which of the potential methods would be suitable as a reference method, consideration was given to the following questions:

- Was the method widely used and accepted? Was the method an EPA-recommended, or similar regulatory method? The selected reference method should be sufficiently used so that it could be cited as an acceptable method for monitoring and/or permit compliance among regulatory authorities.
- Did the selected reference method provide QA/QC criteria that demonstrate acceptable performance characteristics over time?
- Was the method suitable for the species of mercury that were expected to be encountered? The reference method must be capable of determining, as total mercury, all forms of the chemical contaminant known or likely to be present in the matrices.
- Would the method achieve the necessary detection limits to evaluate the sensitivity of each vendor technology adequately?
- Was the method suitable for the concentration range that was expected in the test matrices?

Based on the above considerations, it was determined that SW-846 Method 7471B (analysis of mercury in solid samples by cold-vapor AAS) would be the best reference method. SW-846 Method 7474, (an atomic fluorescence spectrometry method using method 3052 for microwave digestion of the solid) had also been considered a likely technical candidate; however, because this method was not as widely used or referenced, Method 7471B was considered the better choice.

4.4.2 Referee Laboratory Selection

During the planning of the pre-demonstration phase of this project, nine laboratories were sent a statement of work (SOW) for the analysis of mercury to be performed as part of the pre-demonstration. Seven of the nine laboratories responded to the SOW with appropriate bids. Three of the seven laboratories were selected as candidate laboratories based upon technical merit, experience, and pricing. These laboratories received and analyzed blind samples and SRMs during pre-demonstration activities. The referee

laboratory to be used for the demonstration was selected from these three candidate laboratories. Final selection of the referee laboratory was based upon: 1) the laboratory's interest in continuing in the demonstration, 2) the laboratory-reported SRM results, 3) the laboratory MDL for the reference method selected, 4) the precision of the laboratory calibration curve, 5) the laboratory's ability to support the demonstration (scheduling conflicts, backup instrumentation, etc.), and 6) cost.

One of the three candidate laboratories was eliminated from selection based on a technical consideration. It was determined that one of the laboratories would not be able to meet demonstration quantitation limit requirements. (Its lower calibration standard was approximately 50 µg/kg and the vendor comparison requirements were well below this value.) Two candidates thus remained, including the eventual demonstration laboratory, Analytical Laboratory Services, Inc. (ALSI):

Analytical Laboratory Services, Inc.
Ray Martrano, Laboratory Manager
34 Dogwood Lane
Middletown, PA 17057
(717) 944-5541

In order to make a final decision on selecting a referee laboratory, a preliminary audit was performed by the SAIC QA Manager at the remaining two candidate laboratories. Results of the SRM samples were compared for the two laboratories. Each laboratory analyzed each sample (there were two SRMs) in triplicate. Both laboratories were within the 95% prediction interval for each SRM. In addition, the average result from the two SRMs was compared to the 95% CI for the SRM.

Calibration curves from each laboratory were reviewed carefully. This included calibration curves generated from previously performed analyses and those generated for other laboratory clients. There were two QC requirements regarding calibration curves; the correlation coefficient had to be 0.995 or greater and the lowest point on the calibration curve had to be within 10% of the predicted value. Both laboratories were able to achieve these two requirements for all curves reviewed and for a lower standard of 10 µg/kg, which was the lower standard required for the demonstration, based upon information received from each of the vendors. In addition, an analysis of seven standards was reviewed for MDLs. Both laboratories were able to achieve an MDL that was below 1 µg/kg.

It should be noted that vendor sensitivity claims impacted how low this lower quantitation standard should be. These claims were somewhat vague, and the actual quantitation limit each vendor could achieve was uncertain prior to the demonstration (i.e., some vendors claimed a sensitivity as low as 1 µg/kg, but it was uncertain at the time if this limit was actually a PQL or a detection limit). Therefore, it was determined that, if necessary, the laboratory actually should be able to achieve even a lower PQL than 10 µg/kg.

For both laboratories, SOPs based upon SW-846 Method 7471B, were reviewed. Each SOP followed this reference method. In addition, interferences were discussed because there was some concern that organic interferences may have been present in the samples previously analyzed by the laboratories. Because these same matrices were expected to be part of the demonstration, there was some concern associated with how these interferences would be eliminated. This is discussed at the end of this subsection.

Sample throughput was somewhat important because the selected laboratory was to receive all demonstration samples at the same time (i.e., the samples were to be analyzed at the same time in order to eliminate any question of variability associated with loss of contaminant due to holding time). This meant that the laboratory would receive approximately 400 samples for analysis over the period of a few days. It was also desirable for the laboratory to produce a data report within a 21-day turnaround time for purposes of the demonstration. Both laboratories indicated that this was achievable. Instrumentation was reviewed and examined at both laboratories. Each laboratory used a Leeman mercury analyzer for analysis. One of the two laboratories had backup instrumentation in case of problems. Each laboratory indicated that its Leeman mercury analyzer was relatively new and had not been a problem in the past.

Previous SITE program experience was another factor considered as part of these pre-audits. This is because the SITE program generally requires a very high level of QC, such that most laboratories are not familiar with the QC required unless they have previously participated in the program. A second aspect of the SITE program is that it generally requires analysis of relatively "dirty" samples and many laboratories are not use to analyzing such "dirty" samples. Both laboratories have been longtime participants in this program.

Other QC-related issues examined during the audits included: 1) analyses on other SRM samples not previously examined, 2) laboratory control charts, and 3) precision

and accuracy results. Each of these issues was closely examined. Also, because of the desire to increase the representativeness of the samples for the demonstration, each laboratory was asked if sample aliquot sizes could be increased to 1 g (the method requirement noted 0.2 g). Based upon previous results, both laboratories routinely increased sample size to 0.5 g, and each laboratory indicated that increasing the sample size would not be a problem. Besides these QC issues, other less tangible QA elements were examined. This included analyst experience, management involvement in the demonstration, and internal laboratory QA management. These elements were also factored into the final decision.

Selection Summary

There were very few factors that separated the quality of these two laboratories. Both were exemplary in performing mercury analyses. There were, however, some minor differences based upon this evaluation that were noted by the auditor. These were as follows:

- ALSI had backup instrumentation available. Even though neither laboratory reported any problems with its primary instrument (the Leeman mercury analyzer), ALSI did have a backup instrument in case there were problems with the primary instrument, or in the event that the laboratory needed to perform other mercury analyses during the demonstration time.
- As noted, the low standard requirement for the calibration curve was one of the QC requirements specified for this demonstration in order to ensure that a lower quantitation could be achieved. This low standard was 10 µg/kg for both laboratories. ALSI, however, was able to show experience in being able to calibrate much lower than this, using a second calibration curve. In the event that the vendor was able to analyze at concentrations as low as 1 µg/kg with precise and accurate determinations, ALSI was able to perform analyses at lower concentrations as part of the demonstration. ALSI used a second, lower calibration curve for any analyses required below 0.05 mg/kg. Very few vendors were able to analyze samples at concentrations at this low a level.
- Management practices and analyst experience were similar at both laboratories. ALSI had participated in a few more SITE demonstrations than the other laboratory, but this difference was not significant because both laboratories had proven themselves capable of handling the additional QC requirements for the SITE program. In addition, both laboratories had

internal QA management procedures to provide the confidence needed to achieve SITE requirements.

- Interferences for the samples previously analyzed were discussed and data were reviewed. ALSI analyzed two separate runs for each sample. This included analyses with and without stannous chloride. (Stannous chloride is the reagent used to release mercury into the vapor phase for analysis. Sometimes organics can cause interferences in the vapor phase. Therefore, an analysis with no stannous chloride would provide information on organic interferences.) The other laboratory did not routinely perform this analysis. Some samples were thought to contain organic interferences, based on previous sample results. The pre-demonstration results reviewed indicated that no organic interferences were present. Therefore, while this was thought to be a possible discriminator between the two laboratories in terms of analytical method performance, it became moot for the samples included in this demonstration.

The factors above were considered in the final evaluation. Because there were only minor differences in the technical factors, cost of analysis was used as the discriminating factor. (If there had been significant differences in laboratory quality, cost would not have been a factor.) ALSI was significantly lower in cost than the other laboratory. Therefore, ALSI was chosen as the referee laboratory for the demonstration.

4.4.3 Summary of Analytical Methods

4.4.3.1 Summary of Reference Method

The critical measurement for this study was the analysis of mercury in soil and sediment samples. Samples analyzed by the laboratory included field samples, spiked field samples, and SRM samples. Detailed laboratory procedures for subsampling, extraction, and analysis were provided in the SOPs included as Appendix B of the Field Demonstration QAPP. These are briefly summarized below.

Samples were analyzed for mercury using Method 7471B, a cold-vapor atomic absorption method, based on the absorption of radiation at the 253.7-nm wavelength by mercury vapor. The mercury is reduced to the elemental state and stripped/volatilized from solution in a closed system. The mercury vapor passes through a cell positioned in the light path of an AA spectrophotometer. Absorbance (peak height) is measured as a function of mercury concentration. Potassium permanganate is added to eliminate possible interference from sulfide. As per the

method, concentrations as high as 20 mg/kg of sulfide, as sodium sulfide, do not interfere with the recovery of added inorganic mercury in reagent water. Copper has also been reported to interfere; however, the method states that copper concentrations as high as 10 mg/kg had no effect on recovery of mercury from spiked samples. Samples high in chlorides require additional permanganate (as much as 25 mL) because, during the oxidation step, chlorides are converted to free chlorine, which also absorbs radiation at 254 nm. Free chlorine is removed by using an excess (25 mL) of hydroxylamine sulfate reagent. Certain volatile organic materials that absorb at this wavelength may also cause interference. A preliminary analysis without reagents can determine if this type of interference is present.

Prior to analysis, the contents of the sample container are stirred, and the sample mixed prior to removing an aliquot for the mercury analysis. An aliquot of soil/sediment (1 g) is placed in the bottom of a biological oxygen demand bottle, with reagent water and aqua regia added. The mixture is heated in a water bath at 95 °C for 2 minutes. The solution is cooled and reagent water and potassium permanganate solution are added to the sample bottle. The bottle contents are thoroughly mixed, and the bottle is placed in the water bath for 30 minutes at 95 °C. After cooling, sodium chloride-hydroxylamine sulfate is added to reduce the excess permanganate. Stannous chloride is then added and the bottle attached to the analyzer; the sample is aerated and the absorbance recorded. An analysis without non-stannous chloride is also included as an interference check when organic contamination is suspected. In the event of positive results of the non-stannous chloride analysis, the laboratory was to report those results to SAIC so that a determination of organic interferences could be made.

4.4.3.2 Summary of Methods for Non-Critical Measurements.

A selected set of non-critical parameters was also measured during the demonstration. These parameters were measured to provide a better insight into the chemical

constituency of the field samples, including the presence of potential interferences. The results of the tests for potential interferences were reviewed to determine if a trend was apparent in the event that inaccuracy or low precision was observed. Table 4-4 presents the analytical method reference and method type for these non-critical parameters.

Table 4-4. Analytical Methods for Non-Critical Parameters

Parameter	Method Reference	Method Type
Arsenic, barium, cadmium, chromium, lead, selenium, silver, copper, zinc	SW-846 3050/6010	Acid digestion, ICP
Oil and Grease	EPA 1664	n-Hexane extraction, Gravimetric analysis
TOC	SW-846 9060	Carbonaceous analyzer
Total Solids	EPA 2540G	Gravimetric

4.5 Deviations from the Demonstration Plan

During the demonstration Milestone determined that their field instrument was not appropriate for analysis of samples with greater than 5 mg/kg mercury. Previously, during pre-demonstration trials Milestone was able to analyze higher concentration samples, therefore, samples supplied during the demonstration had concentrations greater than 5 mg/kg. SAIC prepared several samples with these higher concentrations, including all samples collected from the Manufacturing Site. During the final day of the demonstration, Milestone was slightly behind in performing sample analysis and therefore a decision was made that because the only samples left to analyze were Manufacturing Site samples and because these samples were all known to be above 5 mg/kg, it was decided by the EPA TOM and Milestone personnel that these samples need not be analyzed. This resulted in the analysis of fewer samples (173), than all other vendors.

Chapter 5

Assessment of Laboratory Quality Control Measurements

5.1 Laboratory QA Summary

QA may be defined as a system of activities, the purpose of which is to provide assurance that defined standards of quality are met with a stated level of confidence. A QA program is a means of integrating the quality planning, quality assessment, QC, and quality improvement efforts to meet user requirements. The objective of the QA program is to reduce measurement errors to agreed-upon limits, and to produce results of acceptable and known quality. The QAPP specified the necessary guidelines to ensure that the measurement system for laboratory analysis was in control, and provided detailed information on the analytical approach to ensure that data of high quality could be obtained to achieve project objectives. The laboratory analyses were critical to project success, as the laboratory results were used as a standard for comparison to the field method results. The field methods are of unknown quality, and therefore, for comparison purposes the laboratory analysis needed to be a known quantity. The following sections provide information on the use of data quality indicators, and a detailed summary of the QC analyses associated with project objectives.

5.2 Data Quality Indicators for Mercury Analysis

To assess the quality of the data generated by the referee laboratory, two important data quality indicators of primary concern are precision and accuracy. Precision can be defined as the degree of mutual agreement of independent measurements generated through repeated application of the process under specified conditions. Accuracy is the degree of agreement of a measured value with the true or expected value. Both accuracy and precision were measured by the analysis of matrix spike/matrix spike

duplicates (MS/MSDs). The precision of the spiked duplicates is evaluated by expressing, as a percentage, the difference between results of the sample and sample duplicate results. The relative percent difference (RPD) is calculated as:

$$RPD = \frac{(\text{Maximum Value} - \text{Minimum Value})}{(\text{Maximum Value} + \text{Minimum Value})/2} \times 100$$

To determine and evaluate accuracy, known quantities of the target analytes were spiked into selected field samples. All spikes were post-digestion spikes because of the high sample concentrations encountered during the demonstration. Pre-digestion spikes, on high-concentration samples would either have been diluted or would have required additional studies to determine the effect of spiking more analyte and subsequent recovery values. To determine matrix spike recovery, and hence measure accuracy, the following equation was applied:

$$\%R = \frac{C_{ss} - C_{us}}{C_{sa}} \times 100$$

where,

C_{ss}	=	Analyte concentration in spiked sample
C_{us}	=	Analyte concentration in unspiked sample
C_{sa}	=	Analyte concentration added to sample

Laboratory control samples (LCSs) were used as an additional measure of accuracy in the event of significant

matrix interference. To determine the percent recovery of LCS analyses, the equation below was used:

$$\%R = \frac{\text{Measured Concentration}}{\text{Theoretical Concentration}} \times 100$$

While several precautions were taken to generate data of known quality through control of the measurement system, the data must also be representative of true conditions and comparable to separate sample aliquots. Representativeness refers to the degree with which analytical results accurately and precisely reflect actual conditions present at the locations chosen for sample collection. Representativeness was evaluated as part of the pre-demonstration and combined with the precision measurement in relation to sample aliquots. Sample aliquoting by the SAIC GeoMechanics Laboratory tested the ability of the procedure to produce homogeneous, representative, and comparable samples. All samples were carefully homogenized in order to ensure comparability between the laboratory and the vendor. Therefore, the RSD measurement objective of 25% or less for replicate sample lot analysis was intended to assess not only precision but representativeness and comparability.

Sensitivity was another critical factor assessed for the laboratory method of analysis. This was measured as a practical quantitation limit and was determined by the low standard on the calibration curve. Two separate calibration curves were run by the laboratory when necessary. The higher calibration curve was used for the majority of the samples and had a lower calibration limit of 25 µg/kg. The lower calibration curve was used when samples were below this lower calibration standard. The lower calibration curve had a lower limit standard of 5 µg/kg. The lower limit standard of the calibration curve was run with each sample batch as a check standard and was required to be within 10% of the true value (QAPP QC requirement). This additional check on analytical sensitivity was performed to ensure that this lower limit standard was truly representative of the instrument and method practical quantitation limit.

5.3 Conclusions and Data Quality Limitations

Critical sample data and associated QC analyses were reviewed to determine whether the data collected were of adequate quality to provide proper evaluation of the project's technical objectives. The results of this review are summarized below.

Accuracy objectives for mercury analysis by Method 7471B were assessed by the evaluation of 23 spiked duplicate pairs, analyzed in accordance with standard procedures in the same manner as the samples. Recovery values for the critical compounds were well within objectives specified in the QAPP, except for two spiked samples summarized in Table 5-1. The results of these samples, however, were only slightly outside specified limits, and given the number of total samples (46 or 23 pairs), this is an insignificant number of results that did not fall within specifications. The MS/MSD results therefore, are supportive of the overall accuracy objectives.

Table 5-1. MS/MSD Summary

Parameter	Value
QC Limits	80%- 120%
Recovery Range	85.2% - 126%
Number of Duplicate Pairs	23
Average Percent Recovery	108%
No. of Spikes Outside QC Specifications	2

An additional measure of accuracy was LCSs. These were analyzed with every sample batch (1 in 20 samples) and results are presented in Table 5-2. All results were within specifications, thereby supporting the conclusion that QC assessment met project accuracy objectives.

Table 5-2. LCS Summary

Parameter	Value
QC Limits	90%- 110%
Recovery Range	90% - 100%
Number of LCSs	24
Average Percent Recovery	95.5%
No. of LCSs Outside QC Specifications	0

Precision was assessed through the analysis of 23 duplicate spike pairs for mercury. Precision specifications were established prior to the demonstration as a RPD less

than 20%. All but two sample pairs were within specifications, as noted in Table 5-3. The results of these samples, however, were only slightly outside specified limits, and given the number of total samples (23 pairs), this is an insignificant number of results that did not fall within specifications. Therefore, laboratory analyses met precision specifications.

Table 5-3. Precision Summary

Parameter	Value
QC Limits	RPD < 20%
MS/MSD RPD Range	0.0% to 25%
Number of Duplicate Pairs	23
Average MS/MSD RPD	5.7%
No. of Pairs Outside QC Specifications	2

Sensitivity results were within specified project objectives. The sensitivity objective was evaluated as the PQL, as assessed by the low standard on the calibration curve. For the majority of samples, a calibration curve of 25-500 µg/kg was used. This is because the majority of samples fell within this calibration range (samples often required dilution). There were, however, some samples below this range and a second curve was used. The calibration range for this lower curve was 5-50 µg/kg. In order to ensure that the lower concentration on the calibration curve was a true PQL, the laboratory ran a low check standard (lowest concentration on the calibration curve) with every batch of samples. This standard was required to be within 10% of the specified value. The results of this low check standard are summarized in Table 5-4.

Table 5-4. Low Check Standards

Parameter	Value
QC Limits	Recovery 90% - 110%
Recovery Range	88.6% - 111%
Number of Check Standards Analyzed	23
Average Recovery	96%

There were a few occasions where this standard did not meet specifications. The results of these samples, however, were only slightly outside specified limits, and given the number of total samples (23), this is an insignificant number of results that did not fall within specifications. In addition, the laboratory reanalyzed the standard when specifications were not achieved, and the second determination always fell within the required limits. Therefore laboratory objectives for sensitivity were achieved according to QAPP specifications.

As noted previously, comparability and representativeness were assessed through the analysis of replicate samples. Results of these replicates are presented in the discussion on primary project objectives for precision. These results show that data were within project and QA objectives.

Completeness objectives were achieved for the project. All samples were analyzed and data were provided for 100% of the samples received by the laboratory. No sample bottles were lost or broken.

Other measures of data quality included method blanks, calibration checks, evaluation of linearity of the calibration curve, holding time specifications, and an independent standard verification included with each sample batch. These results were reviewed for every sample batch run by ALSI, and were within specifications. In addition, 10% of the reported results were checked against the raw data. Raw data were reviewed to ensure that sample results were within the calibration range of the instrument, as defined by the calibration curve. A 6-point calibration curve was generated at the start of each sample batch of 20. A few data points were found to be incorrectly reported. Recalculations were performed for these data, and any additional data points that were suspected outliers were checked to ensure correct results were reported. Very few calculation or dilution errors were found. All errors were corrected so that the appropriate data were reported.

Another measure of compliance were the non-stannous chloride runs performed by the laboratory for every sample analyzed. This was done to check for organic interference. There were no samples that were found to have any organic interference by this method. Therefore, these results met expected QC specifications and data were not qualified in any fashion.

Total solids data were also reviewed to ensure that calculations were performed appropriately and dry weights reported when required. All of these QC checks met

QAPP specifications. In summary, all data quality indicators and QC specifications were reviewed and found to be well within project specifications. Therefore, the data are considered suitable for purposes of this evaluation.

5.4 Audit Findings

The SAIC SITE QA Manager conducted audits of both field activities and of the subcontracted laboratory as part of the QA measures for this project. The results of these technical system reviews are discussed below.

The field audit resulted in no findings or non-conformances. The audit performed at the subcontract laboratory was conducted during the time of project sample analysis. One non-conformance was identified and corrective action was initiated. It was discovered that the laboratory PQL was not meeting specifications due to a reporting error. The analyst was generating the calibration curves as specified above; however, the lower limit on the calibration curve was not being reported. This was immediately rectified and no other findings or non-conformances were identified.

Chapter 6 Performance of the DMA-80

Milestone, Inc. analyzed 173 samples from May 5-8, 2003 in Oak Ridge, TN. Results for these samples were reported by Milestone, and a statistical evaluation was performed. Additionally, the observations performed by SAIC during the demonstration were reviewed, and the remaining primary and secondary objectives were completed. The results of the primary and secondary objectives, identified in Chapter 1, are discussed in Sections 6.1 and 6.2, respectively.

The DMA-80 was used during the pre-demonstration in October, 2002 and during the demonstration by Milestone. Some of the pre-demonstration samples had concentrations above the 5 ppm upper concentration limit identified by Milestone. Results for those samples were reported, and although no statistical evaluation was performed, the results were similar to those reported by the analytical laboratory. To analyze samples above the 5 ppm concentration, a soil to solid dilution using silica gel is prepared. Samples were prepared for Milestone for the

demonstration that were above the 5 ppm concentration. After arriving at the demonstration, it was determined by Milestone that the soil to solid dilution was not appropriate for field analyses; therefore, Milestone elected to not analyze the samples from the manufacturing site (with concentrations between approximately 5-1,000 ppm). Subsequently, precision and accuracy were only determined for sample concentrations below 5 ppm. Use of Milestone's field instrument for higher sample concentrations would likely add additional unknown variance and therefore, it is not recommended for concentrations above 5 ppm based upon the results of this field study.

The distribution of the samples prepared for Milestone Inc. and the referee laboratory, is presented in Table 6-1. Milestone, Inc. received samples at 31 different concentrations for a total of 173 samples. These 173 samples consisted of 20 concentrations in replicates of 7, and 11 concentrations in replicates of 3.

Table 6-1. Distribution of Samples Prepared for Milestone and the Referee Laboratory

Site	Concentration Range	Soil	Sample Type		SRM
			Sediment	Spiked Soil	
Carson River (Subtotal = 75)	Low (1-500 ppb)	7	10	7	7
	Mid (0.5-50 ppm)	9	0	14	21
	High (50->1,000 ppm)	0	0	0	0
Puget Sound (Subtotal = 57)	Low (1 ppb - 10 ppm)	26	0	14	17
	High (10-500 ppm)	0	0	0	0
Oak Ridge (Subtotal = 41)	Low (0.1-10 ppm)	17	3	7	14
	High (10-800 ppm)	0	0	0	0
Subtotal (Total = 173)		72	13	42	70

6.1 Primary Objectives

6.1.1 Sensitivity

Sensitivity objectives are explained in Chapter 4. The two primary sensitivity evaluations performed for this demonstration were the MDL and PQL. Determinations of these two measurements are explained in the paragraphs below, along with a comparison to the referee laboratory. These determinations set the standard for the evaluation of accuracy and precision for the Milestone field instrument. Any sample analyzed by Milestone and subsequently reported below their level of detection was not used as part of any additional evaluations. This was done because of the expectation that values below the lower limit of instrument sensitivity would not reflect the true instrument accuracy and precision. In addition, samples that were reported as greater than 5 mg/kg were not used in the evaluation of primary objectives. Therefore, there were fewer than the 173 samples previously noted in Section 6.0 used in the accuracy and precision evaluation.

The sensitivity measurements of MDL and PQL are both dependent upon the matrix and method. Hence, the MDL and PQL will vary, depending upon whether the matrix is a soil, waste, or water. Only soils and sediments were tested during this demonstration and therefore, MDL calculations for this evaluation reflect soil and sediment matrices. PQL determinations are not independent calculations, but are dependent upon results provided by the vendor for the samples tested.

Comparison of the MDL and PQL to laboratory sensitivity required that a standard evaluation be performed for all instruments tested during this demonstration. PQL, as previously noted, is defined in EPA G-5i as the lowest level of method and instrument performance with a specified accuracy and precision. This is often defined by the lowest point on the calibration curve. Our approach was to let the vendor provide the lower limit of quantitation as determined by their particular standard operating procedure, and then test this limit by comparing results of samples analyzed at this low concentration to the referee laboratory results, or comparing the results to a standard reference material, if available. Comparison of these data are, therefore, presented for the lowest concentration sample results, as provided by the vendor. If the vendor provided "non-detect" results, then no formal evaluation of that sample was presented. In addition, the sample(s) was not used in the evaluation of precision and accuracy.

Method Detection Limit – The standard procedure for determining MDLs is to analyze a low standard or

reference material seven times, calculate the standard deviation and multiply the standard deviation by the "t" value for seven measurements at the 99th percentile ($\alpha = 0.01$). (This value is 3.143 as determined from a standard statistics table.) This procedure for determination of an MDL is defined in 40 CFR Part 136, and while determinations for MDLs may be defined differently for other instruments, this method was previously noted in the demonstration QAPP and is intended to provide a comparison to other similar MDL evaluations. The purpose is to provide a lower level of detection with a statistical confidence at which the instrument will detect the presence of a substance above its noise level. There is no associated accuracy or precision provided or implied.

Several blind standards and field samples were provided to Milestone at their estimated lower limit of sensitivity. The Milestone lower limit of sensitivity was previously estimated at 0.008 mg/kg. Because there are several different SRMs and field samples at concentrations close to the MDL, evaluation of the MDL was performed using more than a single concentration. Samples chosen for calculation were based upon: 1) concentration and how close it was to the estimated MDL, 2) number of analyses performed for the same sample (e.g., more than 4), and 3) if non-detects were reported by Milestone for a sample used to calculate the MDL. Then the next highest concentration sample was selected based upon the premise that a non-detect result reported for one of several samples indicates the selected sample is on the "edge" of the instruments detection capability.

Milestone ran two separate, blind low standards (SRMs), each seven times. This included one standard at 0.082 mg/kg and one at 0.62 mg/kg. For testing the method sensitivity claim, another blind SRM was provided to Milestone which was 0.017 mg/kg. This standard was analyzed only three times; therefore, it was not used in the MDL calculation. The two standards used had standard deviations of 0.29908 for the 0.082 mg/kg SRM, and 0.021769 for the 0.62 mg/kg SRM. Multiplying each of these standard deviations by the "t" statistic noted previously, one can calculate MDLs of 0.94 and 0.068 mg/kg, respectively. There was, however, one of the seven results for the 0.082 mg/kg SRM that appears to be outside the expected range of the other six determinations (e.g. almost an order of magnitude above all other results). If this result is not used, then the recalculated MDL (using 6 values and a "t" statistic of 3.365) is 0.049 mg/kg. Even when an outlier value was excluded from the calculation of the MDL for the 0.082 mg/kg value, the recalculated MDL is still well above the Milestone claim of 0.008 mg/kg. It

would therefore appear that the Milestone claim for sensitivity of 0.008 mg/kg is not applicable to soil and sediment materials.

The objective in estimating an MDL is to run a "low" instrument standard so that a more accurate MDL evaluation can be determined. Discounting the 0.94 mg/kg value calculated for the low standard of 0.082 mg/kg (calculated with an apparent outlier value), the MDL is estimated between 0.049 and 0.068 mg/kg. The equivalent MDL for the referee laboratory based upon analysis of a low standard analyzed seven times is 0.0026 mg/kg. The calculated result is only intended as a statistical estimation, and not a true test of instrument sensitivity.

Practical Quantitation Limit – This value is usually calculated by determining a low standard on the instrument calibration curve and it is estimated as the lowest standard at which the instrument will accurately and precisely determine a given concentration within specified QC limits. The PQL is often around 5-10 times the MDL. This PQL estimation, however, is method- and matrix-dependent. In order to determine the PQL, several low standards were provided to Milestone and subsequent %Ds were calculated.

Using the MDL calculations presented above, this would translate into a low limit PQL of 0.24 mg/kg. The instrument manufacturer, however, suggests a low calibration of 0.010 mg/kg standard in a 1 g sample (over an order of magnitude below the PQL calculated above) which translates to a 10 ng standard "on column." It would appear, based on the information gained during this demonstration, that this low standard calibration is likely well below instrument capabilities in determining an accurate and precise calculation for a PQL. In fact, a low standard of 0.017 mg/kg was tested during the demonstration, as noted above. This was run three times by Milestone, and the average value calculated was 0.0089 mg/kg, with a standard deviation of 0.00108 mg/kg. A 95% CI for this value is, therefore, 0.0077 to 0.010 mg/kg which is outside the range of the reference value determination of 0.017 mg/kg. The %D for the average value reported by Milestone, compared to the reference value of 0.017 mg/kg is 48%. Therefore, it appears that the instrument PQL is above the 0.010 mg/kg value suggested by the manufacturer, and above the 0.017 mg/kg value tested during the demonstration. The PQL may be close to the average MDL determined above, 0.058 mg/kg.

The next lowest standard tested during the demonstration was the 0.082 mg/kg SRM. This standard was run seven

different times during the demonstration. Seven different blind samples were analyzed by Milestone during several different instrument analytical batches. The average value calculated by Milestone for this standard was 0.206 mg/kg, with a standard deviation of 0.29908 mg/kg. The 95% CIs for this standard is therefore -0.016 to 0.428 mg/kg, and encompasses zero. The %D for this calculated value compared to the reference value is 151%. The reference value falls within the CI because the standard deviation is extremely wide (note that the relative standard deviation or coefficient of variation is 145%). Therefore, it would appear that because of this wide CIs range, this may be outside instrument capabilities for a precise evaluation of this low standard. There was, however, one of these seven results that appears to be outside the expected range of the other 6 determinations (see note above). If this result is not used, then the average value is 0.0935 mg/kg, the standard deviation is 0.0134 mg/kg, and a relative standard deviation of 15.7%. The 95% CI is 0.0828 to 0.1042 mg/kg. This CI does not quite include the reference value of 0.082 mg/kg, but is a much narrower range. It is close to the given reference value for the SRM, and overlaps the 95% SRM CI. The %D between the calculated average and reference value for this determination is 13.8%.

The laboratory results for this same standard (0.082 mg/kg SRM) estimated an average value of 0.0729 mg/kg, a standard deviation of 0.005 mg/kg, and a relative standard deviation of 6.7%. The %D between the referee laboratory average and the reported standard is 11%. This is given for purposes of comparison to the Milestone result.

The next lowest SRM value was 0.62 mg/kg. This was run as seven different blind samples by Milestone. The average value was 0.627 mg/kg, and the standard deviation was 0.0218 mg/kg, with a relative standard deviation of 3.47%. The 95% CIs for this standard is 0.611 to 0.643 mg/kg. The SRM value falls within this CIs. As previously noted, this is a very narrow CIs, suggesting not only an accurate, but also a very precise evaluation. The %D between the calculated average and reference value is 1.1%. The laboratory reported an average value of 0.533 mg/kg and a standard deviation of 0.033 and relative standard deviation of 6.2%. The %D between the calculated average and reference value for the referee laboratory is 14%.

It could be suggested that the instrument PQL is above 0.017 mg/kg, perhaps close to 0.082 mg/kg, and below 0.62 mg/kg. Given the information associated with the

MDL determination, it would appear that the PQL is likely not below the average MDL, 0.058 mg/kg.

Sensitivity Summary

The MDL is estimated between 0.049 and 0.068 mg/kg. The equivalent MDL for the referee laboratory based upon analysis of a low standard analyzed seven times is 0.0026 mg/kg. The MDL determination, however, is only a statistical calculation that has been used in the past by EPA and is currently not considered a "true" MDL by SW-846 methodology. SW-846 is suggesting that performance based methods be used and that PQLs be determined using low standard calculations. The low standard calculations suggest that a PQL for the Milestone field instrument is somewhere around 0.082 mg/kg. The referee laboratory PQL confirmed during the demonstration is 0.005 mg/kg. The %D for Milestone field instrumentation at concentrations of 0.082 and 0.62 mg/kg is very comparable to the reference laboratory method suggesting a PQL close to the lower SRM, 0.082 mg/kg for soil and sediment materials.

6.1.2 Accuracy

Accuracy is the instrument measurement compared to a standard or true value. For this demonstration, three separate standards were used for determining accuracy. The primary standard is SRMs. The SRMs are traceable to national systems. These were obtained from reputable suppliers with reported concentrations and an associated 95% CI and 95% prediction interval. The CI from the reference material is used as a measure of comparison with the CI calculated from replicate analyses for the same sample analyzed by the laboratory or vendor. Results are considered comparable if CIs of the SRM overlap with the CIs computed from the replicate analyses by the vendor. While this is not a definitive measure of comparison, it provides some assurance that the two values are equivalent.

Prediction intervals are intended as a measure of comparison for a single laboratory or vendor result with the SRM. When computing a prediction interval, the equation assumes an infinite number of analyses, and it is used to compare individual sample results. A 95% prediction interval would, therefore, predict the correct result from a single analysis 95% of the time for an infinite number of samples, if the result is comparable to that of the SRM. It should be noted that the corollary to this statement is that 5% of the time a result will be outside the prediction interval if determined for an infinite number of samples. If several samples are analyzed, the percentage of results within the

prediction interval will be slightly above or below 95%. The more samples analyzed, the more likely the percentage of correct results will be close to 95% if the result for the method being tested is comparable to the SRM.

Most SRMs were analyzed in replicates of seven by both the vendor and by the referee laboratory. SRMs for reanalyses were analyzed as replicates of three. In some cases there were apparent outlier results, as noted from the Milestone data. When this occurred, calculations were performed both with and without the outlier data for purposes of comparison. These were statistical outliers, generally an order of magnitude above the other results. There was no other reason to consider these samples as outliers other than statistical anomalies, and therefore, these results were not completely discounted from the analysis.

The second accuracy determination used a comparison of vendor results of field samples and SRMs to the referee laboratory results for these same samples. Field samples were used to ensure that "real-world" samples were tested by the vendor. The referee laboratory result is considered as the standard for comparison to the vendor result. This comparison is in the form of a hypothesis test with $\alpha = 0.01$. (Detailed equations along with additional information about this statistical comparison is included in Appendix B.)

It should be noted that there is evidence of a laboratory bias. This bias was determined by comparing average laboratory values to SRM reference values, and is discussed below. The laboratory bias is low in comparison to the reference value. A bias correction was not made when comparing individual samples (replicate analyses) between the laboratory and vendor; however, setting $\alpha = 0.01$ helps mitigate for this possible bias by widening the range of acceptable results between the two data sets.

In addition, there appears to be a Milestone bias, and this bias is high in comparison to the average reference value. This will be discussed in more detail later, however, in general the laboratory and Milestone data were within expected ranges, except for two SRMs which were subsequently reanalyzed by the laboratory after the demonstration, as explained below.

An aggregate analysis or unified hypothesis test was also performed for all 28 sample lots. (A detailed discussion of this statistical comparison is included in Appendix B.) This analysis provides additional statistical evidence in relation to the accuracy evaluation. A bias term is included in this calculation in order to account for any data bias.

The third measure of accuracy is obtained by the analysis of spiked field samples. These were analyzed by the vendors and the laboratory in replicate in order to provide additional measurement comparisons and are treated the same as field samples. Spikes were prepared to cover additional concentrations not available from SRMs or field samples. There is no comparison to the spiked concentration, only a comparison between the vendor and the laboratory reported value.

The purpose for SRM analysis by the referee laboratory is to provide a check on laboratory accuracy. During the pre-demonstration, the referee laboratory was chosen, in part, based upon the analysis of SRMs. This was done in order to ensure that a competent laboratory would be used for the demonstration. The pre-demonstration laboratory qualification showed that the laboratory was within prediction intervals for all SRMs analyzed. Because of the need to provide confidence in laboratory analysis during the demonstration, the referee laboratory also analyzed SRMs as an on-going check of laboratory bias.

The pre-demonstration laboratory evaluation was conducted to help ensure that laboratory SRM data would fall within expected ranges. It was considered possible, however, that during the demonstration the laboratory may fail to fall within the expected concentration ranges for a particular SRM. This did occur, and laboratory corrective action included a check of the laboratory calibration and calibration criteria for those particular samples. (See Table 6-2 for results.) These QC checks were found to be well within compliance, therefore, the laboratory was asked to recalibrate and rerun the two SRMs, as noted in the table. (SRM values were not provided to the laboratory upon reanalysis, nor was the laboratory told why the samples needed to be reanalyzed.) In particular, the SRM labeled as sample lot 40 had a reference value of 1.12 mg/kg.

The laboratory analysis of 7 different blind samples on 7 different runs recorded an average analysis of 0.12 with a relative standard deviation of 30%. This result would suggest that this was not a statistical anomaly, but something else, such as some type of sample interference or simply a mis-labeled sample lot number. Therefore, it was decided that this sample and one additional SRM (sample lot 44) should be reanalyzed. Results of the reanalysis for the two SRMs showed that the laboratory

was well within expected CIs and prediction intervals for both sample lots.

Evaluation of vendor and laboratory analysis of SRMs is performed in the following manner. Accuracy was determined by comparing the 95% CI of the sample analyzed by the vendor and laboratory to the 95% CI for the SRM. (95% CIs around the true value are provided by the SRM supplier.) This information is provided in Tables 6-2 and 6-3, with notations when the CIs overlap, suggesting comparable results. In addition, the number of SRM results for the vendor's analytical instrumentation and the referee laboratory that are within the associated 95% prediction interval are reported. This is a more definitive evaluation of laboratory and vendor accuracy. The percentage of total results within the prediction interval for the vendor and laboratory are reported in Tables 6-2 and 6-3, respectively.

The single most important number from these tables is the percentage of samples within the 95% prediction interval. As noted for the Milestone data, this percentage is 93% with $n = 54$. (This result is computed after removing statistical outliers.) This suggests that the Milestone data are within expected accuracy accounting for statistical variation. For 8 of the 10 determinations (where 7 samples were analyzed), Milestone average results are above the reference value. This would also suggest that possibly Milestone data are biased high. Six of the eight sample lots overlap with the 95% CIs calculated from the Milestone data, compared to values provided by the supplier of the SRM. This number is also suggestive of a reasonable comparison to the SRM value, accounting for statistical variation.

The percentage of samples within the 95% prediction interval for the laboratory data (after taking into account sample reanalysis for apparent anomalies) is 96%. This result also suggests that the ALSI data are within expected accuracy accounting for statistical variation. For 7 of the 10 determinations, ALSI average results are below the reference value. This would also suggest that the ALSI data are potentially biased low. Seven of the ten sample lots overlap with the 95% CIs calculated from the ALSI data, compared to values provided by the supplier of the SRM. This number is also suggestive of a reasonable comparison to the SRM value accounting for statistical variation.

Table 6-2. Milestone SRM Comparison

Sample Lot No.	SRM Value/ 95% CI	Milestone Avg./ 95% CI	CI Overlap (yes/no)	No. of Samples Analyzed	95% Prediction Interval	Milestone No. w/in Prediction Interval
37	0.158 / 0.132 - 0.184	0.482 / 0.0502 - 0.914	yes	7	0 - 0.357	5
37	0.158 / 0.132 - 0.184 ^a	0.307 / 0.230 - 0.384	no	6	0 - 0.357	5
44	4.7 / 4.3 - 5.1	4.40 / 4.23 - 4.57	yes	7	3.0 - 6.4	7
36	0.082 / 0.073 - 0.091	0.206 / 0 - 0.483	yes	7	0.0579 - 0.106 ^b	6
36	0.082 / 0.073 - 0.091 ^a	0.0935 / 0.0781 - 0.109	yes	6	0.0579 - 0.106 ^b	6
40	1.12 / 1.08 - 1.17	1.30 / 1.18 - 1.42	no	7	0.49 - 1.76	7
38	0.62 / 0.54 - 0.70	0.627 / 0.607 - 0.647	yes	7	0.200 - 1.04 ^b	7
39	1.09 / 0.94 - 1.24	1.09 / 1.03 - 1.15	yes	7	0.303 - 1.88 ^b	7
41	2.42 / 2.16 - 2.46	2.83 / 1.64 - 4.02	yes	7	1.30 - 3.32	6
43	3.80 / 3.50 - 4.11	4.17 / 3.51 - 4.83	yes	7	2.41 - 5.20	7
Total Samples				56		50
Total Samples w/o outliers				54		50
% of samples w/in prediction interval						93%

a Calculated results w/out suspected outlier value.

b Prediction interval is estimated based upon n=30. A 95% CI was provided by the SRM supplier but no prediction interval was given.

Table 6-3. ALSI SRM Comparison

Sample Lot No.	SRM Value/ 95% (CI)	ALSI Avg./ 95% CI	CI Overlap (yes/no)	No. of Samples Analyzed	95% Prediction Interval	ALSI No. w/in Prediction Interval
37	0.158 / 0.132 - 0.184	0.139 / 0.0928 - 0.185	yes	7	0 - 0.357	7
44	0.158 / 0.132 - 0.184 ^a	2.33 / 1.05 - 3.61	no	7	0 - 0.357	2
44 ^a	4.7 / 4.3 - 5.1	4.09 / 3.60 - 4.58	yes	3	3.0 - 6.4	3
36	0.082 / 0.073 - 0.091	0.073 / 0.0684 - 0.0776	yes	7	0.0579 - 0.106 ^b	7
40	0.082 / 0.073 - 0.091 ^a	0.12 / 0.087 - 0.15	no	7	0.0579 - 0.106 ^b	0
40 ^a	1.12 / 1.08 - 1.17	1.02 / 0.464 - 1.58	yes	3	0.49 - 1.76	3
38	0.62 / 0.54 - 0.70	0.533 / 0.502 - 0.564	yes	7	0.200 - 1.04 ^b	7
39	1.09 / 0.94 - 1.24	1.24 / 0.634 - 1.84	yes	7	0.303 - 1.88 ^b	6
41	2.42 / 2.16 - 2.46	1.79 / 1.28 - 2.29	yes	7	1.30 - 3.32	6
43	3.80 / 3.50 - 4.11	2.76 / 2.51 - 3.01	no	7	2.41 - 5.20	7
Total Samples				56		42
% of samples w/in prediction interval						75%
Reanalysis						
Total Samples				48		46
% of samples w/in prediction interval						96%

a Reanalysis of SRM samples was performed by laboratory based upon QAPP corrective action procedures.

b Prediction interval is estimated based upon n=30. A 95% CI was provided by the SRM supplier but no prediction interval was given.

Hypothesis Testing

Sample results from field and spiked field samples for the vendor compared to similar tests by the referee laboratory are used as another accuracy check. Spiked samples were used to cover concentrations not found in the field samples, and they are considered the same as the field samples for purposes of comparison. Because of the limited data available for determining the accuracy of the spiked value, these were not considered the same as

reference standards. Therefore, these samples were evaluated in the same fashion as field samples, but they were not compared to individual spiked concentrations.

Using a hypothesis test with alpha = 0.01, vendor results for all samples were compared to laboratory results to determine if sample populations are the same or significantly different. This was performed for each sample lot separately. Alpha was set at 0.01 to help mitigate for inter-laboratory bias as mentioned earlier. This mitigation attempt, however, has limited value because Milestone

data are likely biased high, and ALSI data are likely biased low. As a result of this bias, some sample lots would be expected to result in comparisons that would test as significantly different. Because this test does not separate precision from bias, if Milestone's or ALSI's computed standard deviation was large due to a highly variable result (indication of poor precision), the two CIs could overlap, and therefore, the fact that there was no significant

difference between the two results may be due to high sample variability. Overall precision, however, as noted from the precision evaluation (section 6.1.3) is within expected ranges for both Milestone and ALSI data. Accordingly, associated RSDs have also been reported in Table 6-4 along with results of the hypothesis testing for each sample lot.

Table 6-4. Accuracy Evaluation by Hypothesis Testing

Sample Lot No./ Site	Avg. Conc. mg/kg	RSD or CV	Number of Measurements	Significantly Different at Alpha = 0.01	Relative Percent Difference (Milestone to ALSI)
03/ Oak Ridge				no	53.5%
Milestone	0.45	23.2%	3		
ALSI	0.26	3.8%	3		
09/ Oak Ridge				no	40.5%
Milestone	0.70	39.8%	7		
ALSI	0.47	34.2%	7		
14/ Oak Ridge				no	10.5%
Milestone	5.28	12.6%	3		
ALSI	4.75	27.5%	6		
37/ Oak Ridge SRM				no	111%
Milestone	0.48	96.9%	7		
ALSI	0.14	36.4%	7		
44/ Oak Ridge SRM				yes	61.4%
Milestone	4.40	4.2%	7		
ALSI	2.33	59.4%	7		
44/ Oak Ridge SRM ^a				no	7.3%
Milestone	4.40	4.2%	7		
ALSI	4.09	4.8%	3		
02/ Puget Sound				no	38.7%
Milestone	0.089	41.3%	7		
ALSI	0.060	23.6%	7		
05/ Puget Sound				no	27.5%
Milestone	0.28	15.8%	3		
ALSI	0.21	33.3%	3		
08/ Puget Sound				no	45.2%
Milestone	0.87	31.6%	3		
ALSI	0.55	13.4%	7		
10/ Puget Sound				no	64.7%
Milestone	0.71	25.3%	3		
ALSI	0.36	20.5%	3		
11/ Puget Sound				yes	49.9%
Milestone	1.35	19.2%	7		
ALSI	0.81	32.6%	7		
12/ Puget Sound				yes	29.9%
Milestone	1.46	2.6%	3		
ALSI	1.08	2.8%	3		
35/ Puget Sound SRM				no	1.5%
Milestone	0.0089	12.2%	3		
ALSI	0.0087	6.3%	7		
36/ Puget Sound SRM				no	95.4%
Milestone	0.21	144%	7		
ALSI	0.073	6.7%	7		
40/ Puget Sound SRM				yes	166%
Milestone	1.30	10.1%	7		
ALSI	0.12	30.0%	7		

Table 6-4. Continued

Sample Lot No./ Site	Avg. Conc. mg/kg	RSD or CV	Number of Measurements	Significantly Different at Alpha = 0.01	Relative Percent Difference (Milestone to ALSI)
^a 40/ Puget Sound SRM					
Milestone	1.30	10.1%	7	no	24.1%
ALSI	1.02	22.0%	3		
57/ Puget Sound				yes	33.6%
Milestone	1.03	20.1%	7		
ALSI	0.73	16.2%	7		
01/ Carson River				no	1.3%
Milestone	0.24	41.6%	7		
ALSI	0.24	37.8%	7		
04/ Carson River				no	19.7%
Milestone	0.13	16.1%	7		
ALSI	0.11	9.1%	7		
06/ Carson River				no	55.3%
Milestone	0.46	15.5%	3		
ALSI	0.26	15.7%	7		
15/ Carson River				no	-12.0%
Milestone	3.75	11.1%	3		
ALSI	4.23	24.5%	7		
16/ Carson River				no	20.6%
Milestone	8.78	34.2%	3		
ALSI	7.14	13.7%	3		
18/ Carson River				no	-15.4%
Milestone	8.66	7.8%	3		
ALSI	10.1	8.0%	7		
38/ Carson River SRM				yes	17.2%
Milestone	0.63	3.5%	7		
ALSI	0.53	6.2%	7		
39/ Carson River SRM				no	-12.9%
Milestone	1.09	6.1%	7		
ALSI	1.24	52.9%	7		
41/ Carson River SRM				no	45.0%
Milestone	2.83	45.5%	7		
ALSI	1.79	30.5%	7		
43/ Carson River SRM				yes	40.7%
Milestone	4.17	17.1%	7		
ALSI	2.76	9.6%	7		
56/ Carson River				no	19.9%
Milestone	0.28	27.8%	7		
ALSI	0.23	12.6%	7		
58/ Carson River				no	13.0%
Milestone	0.86	12.3%	7		
ALSI	0.76	8.6%	7		
59/ Carson River				no	16.6%
Milestone	2.02	13.8%	7		
ALSI	1.71	7.9%	7		

^a Reanalysis performed due to SRM results outside expected accuracy specifications.

CV Coefficient of variance

Of the 28 sample lots, 7 results are significantly different. This may be a higher number of results outside expected accuracy evaluations for Milestone and ALSI if these two sets of data were equal with alpha = 0.01. (Normally this would mean that 1 in 100 results would be outside expected ranges for both sets of data due to statistical variation.) There were, however, two SRMs analyzed by the laboratory that appeared in question. Per QAPP

specifications, reanalysis for the noted SRMs (sample lots 40 and 44) were performed. Upon reanalysis with three replicates, both of the sample lots were within expected accuracy specifications. Hypothesis testing at alpha = 0.01 resulted in no significant difference between ALSI and Milestone for the reanalyzed SRMs. Therefore, the number of sample lots that are significantly different dropped to 5. Most of the relative percent differences are

positive (all but three) which indicates that the Milestone result is generally higher than the laboratory result. This is indicative of the previously noted low bias associated with the laboratory data.

The second set of three analyses (sample lot 40) produced similar RSD results as had been achieved with 7 replicates, indicating a very precise determination. When accounting for differences between the ALSI and Milestone data, it should be noted that there may be inherent biases in both sets of data. ALSI may be biased low and Milestone may be biased high, therefore comparisons of the two data sets would likely result in an additional number of data sets that were significantly different over and above the 1 in 100 difference as noted previously.

In determining the number of results significantly above or below the value reported by the referee laboratory, 16 of 30 Milestone average results were found to have relative percent differences less than 30%. Only 2 of 30 Milestone average results have relative percent differences greater than 100% for this same group of samples (see Table 6-5). The differences are accentuated by the low bias for the laboratory results and the high bias for Milestone results.

There appears to be more significant differences in the Puget Sound sample set than any of the other sample lots, which may be due to an interference for these particular samples (see Table 6-6).

Table 6-5. Number of Sample Lots Within Each %D Range

	<30%	>30%, <50%	>50%, <100%	>100%	Total
Positive %D	13	7	5	2	27
Negative %D	3	0	0	0	3
Total	16	7	5	2	30

Table 6-6. Concentration (in mg/kg) of Non-Target Analytes

Lot #	Site	TOC	O&G	Ag	As	Ba	Cd	Cr	Cu	Pb	Se	Sn	Zn	Hg
1	Carson River	870	190	<0.5	9	210	<0.5	19	13	3	<2	<5	60	0.19
2	Puget Sound	3500	290	<0.5	3	23	<0.5	16	10	1	<2	<5	24	0.04
3	Oak Ridge	2300	530	1.8	4	150	<0.5	46	20	15	<2	<5	55	0.31
4	Carson River	2400	200	<0.5	8	240	<0.5	17	32	12	<2	<5	66	0.10
5	Puget Sound	3500	210	<0.5	3	28	<0.5	18	11	3	<2	<5	28	0.16
6	Carson River	7200	200	<0.5	4	32	<0.5	16	9	1	<2	<5	24	0.23
8	Puget Sound	8100	200	<0.5	3	27	1.0	17	23	99	2	<5	37	0.37
9	Oak Ridge	3300	150	1.9	5	160	0.5	70	49	24	<2	<5	100	0.66
10	Puget Sound	4200	130	<0.5	3	24	<0.5	18	8	1	<2	<5	24	0.62
11	Puget Sound	3800	130	<0.5	4	20	<0.5	18	8	1	<2	<5	24	0.63
12	Puget Sound	3500	290	<0.5	3	23	0.8	16	7	2	<2	<5	23	1.1
14	Oak Ridge	7800	180	0.32	2	41	0.4	16	9	11	<2	<4	74	78
15	Carson River	2700	70	3.2	22	100	<0.5	13	18	18	<2	<5	49	3.3
16	Carson River	2100	80	<0.5	4	150	<0.5	18	39	14	<2	<5	81	7.3
18	Carson River	1900	70	26	17	46	2.0	6	62	200	<2	<5	390	9.3
35	SRM Canmet S0-3	NR	NR	NR	NR	300	NR	26	17	14	NR	NR	52	0.02
36	SRM Canmet S0-2	NR	NR	NR	NR	970	NR	16	7	21	NR	NR	120	0.08
37	SRM CRM-016	NR	NR	0.7	7.8	79	0.47	14	16	14	1	NR	70	0.16
38	SRM NWRI TH-2	NR	NR	5.8	8.7	570	5.2	120	120	190	0.83	NR	900	0.62
39	SRM NWRI WQB-1	NR	NR	1	23	600	2	89	80	84	1	3.9	275	1.09
40	SRM CRM 020	NR	NR	38	400	25	15	14	730	5100	6.6	NR	3000	1.1
41	SRM CRM 026	NR	NR	0.57	5.4	210	12	27	19	26	1.9	NR	140	2.4
43	SRM CRM 027	NR	NR	6	12	170	12	27	9.9	52	14	NR	51	3.8
44	SRM CRM 021	NR	NR	6.5	25	590	1.2	11	4800	6500	NR	300	550	4.7
56	Spiked Lot 1	870	190	<0.5	9	210	<0.5	19	13	3	<2	<5	60	0.20
57	Spiked PS- X1,X4	3500	290	<0.5	3	23	<0.5	16	10	1	<2	<5	24	0.61
58	Spiked CR-SO-14	870	190	<0.5	9	210	<0.5	19	13	3	<2	<5	60	0.74
59	Spiked CR-SO-14	870	190	<0.5	9	210	<0.5	19	13	3	<2	<5	60	1.6

CRM = Canadian Reference Material

NR = Not Reported by Standard Supplier

Upon examination of non-target analyte data (Table 6-6) collected for these samples, no obvious interference was noted. For example, a high organic content may cause interference, but these samples do not necessarily have a higher organic content than other samples tested. In addition, the method 7471B mercury analysis requires that a non-stannous chloride analysis be conducted with each sample analyzed, in order to test for organic interferences. Upon examination of the laboratory data for the sample sets mentioned above, there was no apparent interference noted in the non-stannous chloride analysis. Other interferences caused by additional elements were also not found to be significant. Of course, there could be interferences that were not tested, and therefore, while it may be an interference particular to this sample lot, the exact cause remains unknown.

In addition to the statistical summary presented above, data plots (Figures 6-1 and 6-2) are included in order to present a visual interpretation of the accuracy. Two separate plots have been included for the Milestone data. These two plots are divided based upon sample concentration in order to provide a more detailed

presentation. Concentrations of samples analyzed by Milestone ranged approximately from 0.01 to 10 mg/kg. The previous statistical summary eliminated some of these data based upon whether concentrations were interpreted to be in the analytical range of the Milestone field instrument. This graphical presentation presents all data points. It shows Milestone data compared to ALSI data plotted against concentration. Sample groups are shown by connecting lines. Breaks between groups indicate a different set of samples at a different concentration. Sample groups were arranged from lowest to highest concentration.

As can be seen by this presentation, samples analyzed by Milestone appear to match well with the ALSI results. There are some outlier data points, however, most of these data points when averaged with other data points from the same sample group were within approximate averaged concentrations. This is only a visual interpretation and does not provide statistical significance. It does, however, provide a visual interpretation that supports the previous statistical results for accuracy, as presented above.

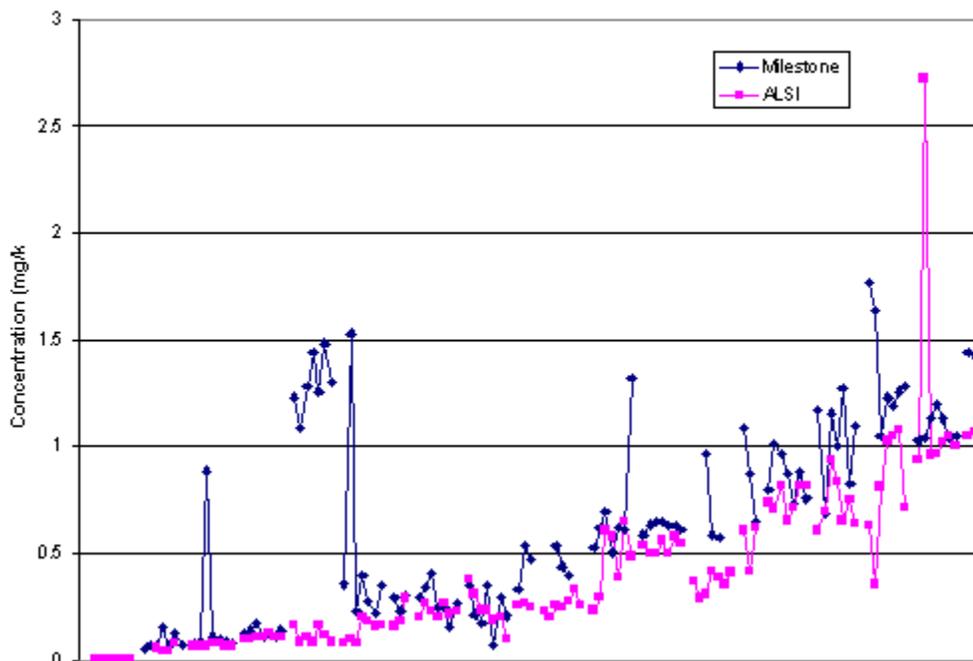


Figure 6-1. Data plot for low concentration sample results.

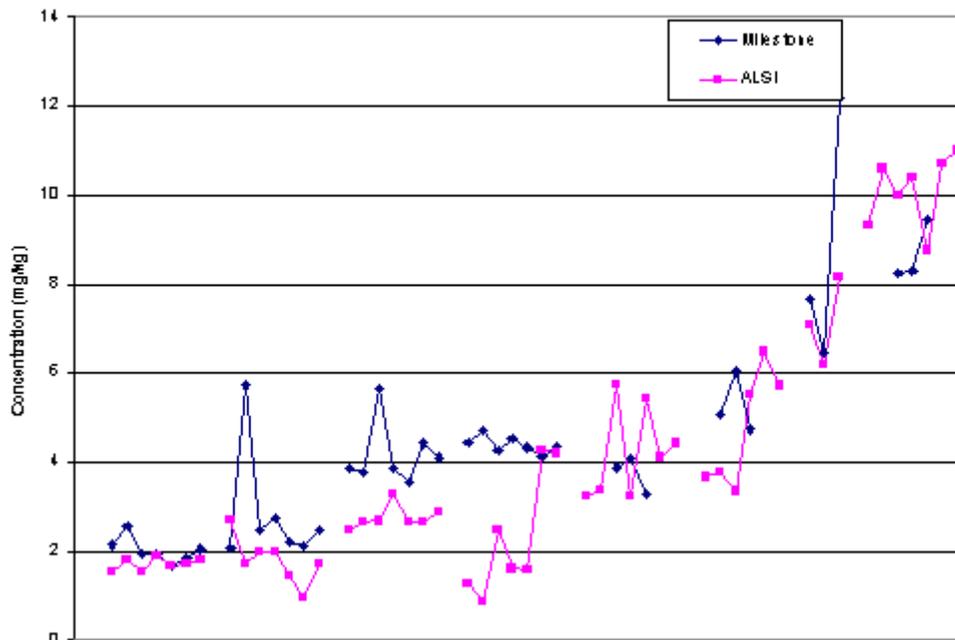


Figure 6-2. Data plot for high concentration sample results.

Unified Hypothesis Test

SAIC performed a unified hypothesis test analysis to assess the comparability of analytical results provided by Milestone and those provided by ALSI. (See appendix B for a detailed description of this test.) Milestone and ALSI both supplied multiple assays on replicates derived from a total of 28 different sample lots, either field materials or reference materials. (Only samples above the previously estimated Milestone PQL were used.) The Milestone and ALSI data from these assays formed the basis of this assessment. The results of this unified hypothesis test show that the two data sets are similar and therefore Milestone data compared well to the referee laboratory. This confirms previous statistical determinations from above showing very few differences between the two data sets.

Milestone analytical results for sample lot 37 were large relative to the concentration provided with the sample reference material, and the Milestone data for sample lot 37 made a substantial contribution to the chi-square statistic. Accordingly, excluding sample lot 37 (n = 27)

from the data set resulted in a chi-square statistic of 42.1, which does not exceed the upper 99th percentile of the chi-square distribution with 26 degrees of freedom with value 45.6. So, excluding sample lot 37 data, results from this analysis suggest that the two data sets are the same for the ALSI laboratory and the Milestone field instrument. The null hypothesis tested was that, on average, Milestone and ALSI produce the same results within a given sample lot. Additional information about this statistical evaluation is included in Appendix B.

Accuracy Summary

In summary, Milestone data compared to SRM values were within expected accuracy determinations. ALSI data compared to SRM values were also within expected accuracy determinations, after reanalysis of two sample lots. These two comparisons are the best evidence suggesting that the Milestone field instrument and the ALSI analysis provide accurate data. The additional comparison of these two data sets (hypothesis test for each sample lot), do not provide evidence contrary to the results of this comparison, but do not necessarily support this conclusion.

The number of Milestone average values less than 30% different from the referee laboratory results or SRM reference values was 16 of 30 different sample lots. However, when making the comparison between Milestone and ALSI data, and taking into account the possible bias associated with both sets of data, the hypothesis test and the %D comparison may be within reasonable expectations for considering these two separate analyses to be equivalent.

The unified hypothesis test provides additional evidence that there is no statistical difference between data sets provided by ALSI and Milestone. Overall, the accuracy evaluations suggest that the Milestone field instrument provides results that are comparable within expected accuracy specifications, and should be considered equivalent.

6.1.3 Precision

Precision is usually thought of as repeatability of a specific measurement, and it is often reported as RSD. The RSD is computed from a specified number of replicates. The more replications of a measurement, the higher confidence associated with a reported RSD. Replication of a measurement may be as few as 3 separate measurements to 30 or more measurements of the same sample, depending upon the degree of confidence desired in the specified result. Most samples were analyzed seven times by both Milestone and the referee laboratory. In some cases, samples may have been analyzed as few as three times. This was often the situation when it was believed that the chosen sample, or SRM, was likely to be below the vendor quantitation limit. The precision goal for the referee laboratory, based upon pre-demonstration results is an RSD of 25% or less. A descriptive evaluation for differences between Milestone RSDs and the referee laboratory RSDs was determined. In Table 6-7, the RSD for each separate sample lot is shown for Milestone compared to the referee laboratory. The average RSD was then computed for all measurements made by Milestone, and this value was compared to the average RSD for the laboratory.

In addition, the precision of an analytical instrument may vary, depending upon the matrix being measured, the concentration of the analyte, and whether the measurement is made for an SRM or a field sample. To evaluate precision for clearly different matrices, an overall average RSD for the SRMs is calculated and compared to the average RSD for the field samples. This comparison

is also included in Table 6-7 and shown for both Milestone and the referee laboratory.

The purpose of this evaluation is to determine the field instrument's capability to precisely measure analyte concentrations under real-life conditions. Instrument repeatability was measured using samples from each of three different sites. Within each site, there may be two separate matrices, soil and sediment. Not all sites have both soil and sediment matrices, nor are there necessarily high, medium, and low concentrations for each sample site. Therefore, spiked samples were included to cover additional ranges. (Originally there were 4 different sites chosen for each vendor; however, Milestone's capability to measure high concentration samples was limited under field conditions. Therefore, because the samples from the manufacturing site were believed to be above Milestone's upper quantitation limit, these samples were not analyzed.)

Table 6-7 shows results from Oak Ridge, Puget Sound, and Carson River. It was thought that because these three different field sites represented different matrices, measures of precision may vary from site to site. The average RSD for each site is shown in Table 6-7 and compared between Milestone and the referee laboratory. SRM RSDs are not included in this comparison because SRMs, while grouped with different sites for purposes of ensuring that the samples remained blind during the demonstration, were not actually samples from that site, and were, therefore, compared separately.

The RSDs of various concentrations are compared by noting the RSD of the individual sample lots. The ranges of test samples (field, SRMs, and spikes) were selected to cover the appropriate analytical ranges of Milestone's instrumentation. Average referee laboratory values for sample concentrations are included in the table, along with SRM values, when appropriate. These are discussed in detail in the Section 6.1.2 and are included here for purposes of precision comparison. Sample concentrations were separated into approximate ranges: low, medium, and high, as noted in Table 6-7 and Table 6-1. Milestone's field instrument, however, is an atomic absorption instrument and therefore less subject to concentration variations. This means that variations in precision due to varying concentrations is less likely. Because Milestone performed no sample dilution, there are no additional operations that would likely affect precision measurements.

Table 6-7. Evaluation of Precision

Sample Lot No. Milestone and Lab	Avg. Conc. or Reference SRM Value	RSD	Number of Samples	w/in 25% RSD Goal?
OAK RIDGE				
Lot no. 03 Milestone	0.26 (medium)	23.2%	3	yes
ALSI		3.8%	3	yes
Lot no. 09 Milestone	0.47 (medium)	39.8%	7	no
ALSI		34.2%	7	no
Lot no. 14 Milestone	4.75 (high)	12.6%	3	yes
ALSI		27.5%	6	no
Lot no.37/ SRM Milestone	0.16 (medium)	96.9%	7	no
*Milestone w/out outlier		23.8%	6	yes
ALSI		36.4%	7	no
Lot no. 44/ SRM Milestone	4.70 (high)	4.2%	7	yes
ALSI		59.4%	7	no
Oak Ridge Avg. RSD Milestone		25.0%		yes
ALSI		28.9%		no
PUGET SOUND				
Lot no. 02 Milestone	0.060 (low)	41.3%	7	no
ALSI		23.6%	7	yes
Lot no. 05 Milestone	0.21 (medium)	15.8%	3	yes
ALSI		33.3%	3	no
Lot no. 08 Milestone	0.36 (medium)	31.6%	3	no
ALSI		13.4%	7	yes
Lot no. 10 Milestone	0.55 (medium)	25.0%	3	yes
ALSI		20.5%	3	yes
Lot no. 11 Milestone	0.81 (medium)	19.2%	7	yes
ALSI		32.6%	7	no
Lot no. 12 Milestone	1.08 (high)	2.6%	3	yes
ALSI		2.8%	3	yes
Lot no. 35/ SRM Milestone	0.02 (low)	12.2%	3	yes
ALSI		6.3%	7	yes
Lot no. 36/ SRM Milestone	0.08 (low)	144%	7	no
*Milestone w/out outlier		15.7%	6	yes
ALSI		6.7%	7	yes
Lot no. 40/ SRM Milestone	1.12 (high)	10.1%	7	yes
ALSI		30.0%	7	no
Lot no. 57 Milestone	0.73 (medium)	20.1%	7	yes
ALSI		16.2%	7	yes
Puget Sound/ Avg. RSD Milestone		22.3%		
ALSI		20.4%		

Table 6-7. Continued

Sample Lot No. Milestone and Lab	Avg. Conc. or Reference SRM Value	RSD	Number of Samples	w/in 25% RSD Goal?
CARSON RIVER				
Lot no. 01	0.24 (medium)			
Milestone		41.6%	7	no
ALSI		37.8%	7	no
Lot no. 04	0.11 (medium)			
Milestone		16.1%	7	yes
ALSI		9.1%	7	yes
Lot no. 06	0.26 (medium)			
Milestone		15.5%	3	yes
ALSI		15.7%	7	yes
Lot no. 15	4.23 (high)			
Milestone		11.1%	3	yes
ALSI		24.5%	7	yes
Lot no. 16	7.14 (high)			
Milestone		34.2%	3	no
ALSI		13.7%	3	yes
Lot no. 18	10.1 (high)			
Milestone		7.8%	3	yes
ALSI		8.0%	7	yes
Lot no. 38/ SRM	0.62 (medium)			
Milestone		3.5%	7	yes
ALSI		6.2%	7	yes
Lot no. 39/ SRM	1.09 (high)			
Milestone		6.1%	7	yes
ALSI		52.9%	7	no
Lot no. 41/ SRM	2.42 (high)			
Milestone		45.5%	7	no
ALSI		30.5%	7	no
Lot no. 43/ SRM	3.85 (high)			
Milestone		17.1%	7	yes
ALSI		9.6%	7	yes
Lot no. 56	0.23 (medium)			
Milestone		27.8%	7	no
ALSI		12.6%	7	yes
Lot no. 58	0.76 (medium)			
Milestone		12.3%	7	yes
ALSI		8.6%	7	yes
Lot no. 59	1.71 (high)			
Milestone		13.8%	7	yes
ALSI		7.9%	7	yes
Carson River/ Avg. RSD				
Milestone		20.0%		yes
ALSI		15.3%		yes
SUMMARY STATISTICS				
Overall Avg. RSD				
Milestone		19.4%		yes
ALSI		23.7%		yes
Field Samples/ Avg. RSD				
Milestone		22.0%		yes
ALSI		19.6%		yes
SRMs/ Avg. RSD				
Milestone		15.3%		yes
ALSI		26.5%		no

Samples below the MDL, as determined in the section discussing sensitivity, were not included in Table 6-7. There appears to be no correlation between concentration (low, medium, or high) and RSD; therefore, no other formal evaluations of this comparison were performed.

The referee laboratory analyzed replicates of all samples analyzed by Milestone. This was used for purposes of precision comparison to Milestone. RSD for the vendor and the laboratory were calculated individually and shown in Table 6-7.

Milestone precision is very comparable to the referee laboratory (Table 6-7). The single most important measure of precision provided in Table 6-7, overall average RSD, is 23.7% for the referee laboratory compared to the Milestone average RSD of 19.4%. Both of these RSDs are within the predicted 25% RSD objective for precision, expected from both analytical and sampling variance.

In addition, field sample precision compared to SRM precision shows that there may be no significant difference between these two sample lots; field sample RSD 19.6% for ALSI and 22.0% for Milestone; SRM RSD 26.5% for ALSI and 15.3% for Milestone. Differences in these overall RSD numbers suggest differences in the two methods and/or instruments but not differences attributable to field samples or SRMs. This would suggest that not only was there no difference in analysis of these samples, but that the preparation procedure for the field samples (see Section 4.3.1 for description of field sample homogenization) was very thorough and complete. For purposes of this analysis, spiked samples are considered the same as field samples because these were similar field matrices and the resulting variance was expected to be equal to field samples. The replicate sample RSDs also confirm the pre-demonstration results, showing that sample homogenization procedures met their originally stated objectives, and that SRM and field sample variation were not significantly different.

There also appears to be no significant site variation between Oak Ridge, Puget Sound, and the Carson River site samples. (See Table 6-7 showing average RSDs for each of these sample lots. These average RSDs are computed using only the results of the field samples and not the SRMs.) In addition, there appears to be no difference in precision for different concentrations, as noted in the discussion above.

Precision Summary

The precision of the Milestone field instrument is very comparable to laboratory precision, and within expected

precision variation for soil and sediment matrices. The Milestone field instrument can therefore obtain very precise measurements, equivalent to laboratory variation covering the entire range of the instrument (PQL as determined in Section 6.1.1 and an upper limit set by Milestone of 5 mg/kg) as determined during this demonstration.

6.1.4 Time Required for Mercury Measurement

During the demonstration, the time required for mercury measurement activities was measured. Specific activities that were timed included: instrument setup, sample analysis, and instrument disassembly. One field technician performed all operations during the demonstration, with the exception of unloading the DMA-80, which measures 80 by 42 by 30 (H) cm and weighs 56 kg.

Setup and disassemble times were measured once. Analytical time was measured each day, beginning when the first blank was started, and continuing until the last blank was completed at the end of the day. Any downtime was noted and then subtracted from the total daily operational time. Finally, the total of the operational time from all four days was divided by the total number of analyses performed. For this calculation, analyses of blanks and calibration standards, and reanalyses of samples were not included in the total number of samples.

Setup time for the DMA-80 consisted of removing the instrument from the shipping container, placement on a level working surface, establishment of all electrical and gas tubing connections, and instrument warmup. The time required to remove the DMA-80 from the shipping container could not be measured, because the device was shipped to the site in the back of a vehicle without any packaging. However, based on information provided by the vendor, it is estimated that two people could remove the device from the corrugated cardboard shipping container in less than 5 minutes. Setup time for other peripheral devices, such as the computer/monitor and analytical balance, was also included in the instrument setup time. These two devices were packaged, along with other supplies, in a large corrugated cardboard box. The balance came in two pieces: the base and the top cover. The balance was set up and leveled in 10 minutes. Setup of the computer/monitor took less than 5 minutes.

During the demonstration, the DMA-80 was moved to a table on the first and last days of field activities. The vendor required the assistance of one person to perform this task. It is conceivable that one person could perform this operation, but not all individuals would be able to move

the large, heavy instrument without assistance. On the second and third days of the demonstration, the vendor operated out of the back of an SUV, and required no assistance in setting up the DMA-80. It is estimated that this activity took 5 minutes on average.

After all devices were set in place, and electrical and gas flow connections were made. The DMA-80 was connected to a power source and to the computer/monitor. The balance was also connected to the power source and the computer/monitor. Gas connections were made from the oxygen cylinder, through a pressure regulator (this part was already completed), and then to the DMA-80. A mercury trap (pre-assembled) was inserted in the vent line, which was then attached to the DMA-80. Overall, the electrical and gas flow connections required approximately 10 minutes. However, if the mercury trap had to be assembled and the gas flow regulator installed, as would be the case for most operations, the total setup time is estimated at 20-30 minutes for the first usage. After that the trap can be used for 3 months without reassembly.

After setup was complete, the instrument required approximately 20 minutes to come to operating temperature. It is worth noting that setup of the balance was performed during this time period.

Overall, the time required to remove the DMA-80 from its shipping container, set up the device, allow the instrument to reach operating temperature, and set up peripheral devices during instrument warmup is estimated at approximately 30-40 minutes.

Individual sample analysis times were not measured for the duration of the demonstration. Analysis time was estimated by recording start and stop times each day and accounting for any instrument downtime due to operator breaks or device failure and maintenance activities. Therefore, the total time for analyses included blanks, calibration standards, and any sample reanalyses; however, the total number of analyses performed includes only demonstration samples (samples, spikes, and SRMs), not vendor blanks, calibration standards, or reanalyses. Table 6-8 presents the time measurements recorded for each of the four days of operation of the DMA-80.

Table 6-8. Time Measurements for Milestone

Day	Day 1	Day 2	Day 3	Day 4	Total
Run Time (minutes)	260	590	410	70	1330

Instrument disassembly was measured from the time that sample or blank analyses ended until the instrument was disassembled and placed in the original shipping container. During the demonstration, the balance was disassembled and packaged while the final samples were being analyzed (an advantage of an auto-sampler). This complete process took about 5 minutes.

The DMA-80 was not re-packaged because it was not brought to the site in a shipping container. Disassembly of the DMA-80 involved turning off power, disconnecting the power source and interface cables to the computer/monitor, removal of the auto-sampler tray, and disconnecting the oxygen supply. This process required 15 minutes to complete. Packaging would require that the DMA-80 be placed in a custom shipping container with reinforced corners and buffer spaces. The auto-sampler tray, cables, gas tubing, weigh boats, and any other supplies would need to be packaged also. Finally, the oxygen cylinder would need to be disassembled by closing the main valve, bleeding off any pressure in the line, and removing the plastic tubing and pressure regulator. It is estimated that this complete process would take approximately 30 minutes, not including the time to return the oxygen cylinder to the supplier.

Analysis Time Summary

In total, Milestone analyzed 173 samples during the demonstration. The turnaround time on individual sample analyses was 5 minutes. However, using the total analytical time reported in Table 6-2 (1330 minutes), 7.7 minutes per analysis is a better approximation of real world operating conditions. It should be noted that the number of analyses does not include blanks, standards, and reanalyzed samples. These numbers will vary from site to site depending on project goals (e.g., are “greater than” results acceptable, or must all samples be quantified) and sample demands (e.g., high concentration samples or very heterogeneous samples). If project goals require all samples to be quantified, the number of reanalyses and blanks required could be higher and, therefore, the time per analysis could be greater. On the other hand, if sample results can be reported as “greater than” values (as was generally done during the demonstration), then 8 minutes per analysis is a reasonable average time.

6.1.5 Cost

Background information, assumptions used in the cost analysis, demonstration results, and a cost estimate are provided in Chapter 7.

6.2 Secondary Objectives

This section discusses the performance results for the DMA-80 in terms of secondary objectives described in Section 4.1. These secondary objectives were addressed based on observations of the DMA-80 and information provided by Milestone.

6.2.1 Ease of Use

Documents the ease of use, as well as the skills and training required to properly operate the device.

Based on observations made during the demonstration, the DMA-80 is easy to operate, requiring one field technician with a basic knowledge of chemistry acquired on the job or in a university and training on the DMA-80.

The vendor provided an SOP, entitled "Getting Started: Calibration and Analysis Procedures," for use with the

DMA-80 (see Appendix B). This procedure was evaluated during the demonstration. The procedure was generally easy to understand. SOP Section 1.0, Calibration, could not be evaluated because the vendor performed equipment calibration in the office, prior to shipping the instrument to the field. The vendor performed daily calibration checks as recommended in Section 2.0 of the SOP. Calibration involved weighing a small amount of a standard in a weigh boat, placing it in the auto-sampler, and processing the standard sample through the DMA-80. Figure 6-3 presents an example of typical calibration results. Figure 6-4 shows example 3-point calibration curves for the two cuvettes.

The instruction on calibration checks and analyses was clear. Combined with instrument training, this SOP would provide a user with adequate direction on basic use of the DMA-80. Included were instructions on calibrating the instrument and running blanks before processing samples. In addition, Milestone provides a 1-day training course (at the purchaser's cost) and telephone support at no cost to anyone who purchases the DMA-80. Neither of these was evaluated during the demonstration.

Sample	Sample ID	Weight [g]	Height	Hg [ng]	Result [µg/kg]	Calibration Factor	Remarks
1	Cleaning	0.0001	0.0102	0.00	0.00	1	
1	Blank	0.1000	0.0038	0.00	0.00	1	
1	Blank	0.1000	0.0023	0.00	0.00	1	
1	Blank	0.1000	0.0019	0.00	0.00	1	
1	Std 50 ng	0.1000	0.1257	5.00	50.00	1	
1	Std 50 ng	0.1000	0.1265	5.00	50.00	1	
1	Std 50 ng	0.1000	0.1293	5.00	50.00	1	
1	Std 100 ng	0.1000	0.2514	10.00	100.00	1	
1	Std 100 ng	0.1000	0.2615	10.00	100.00	1	
1	Cleaning	0.0001	0.0078	0.00	0.00	1	
1	Std 100 ng	0.1000	0.2518	10.00	100.00	1	
1	Std 200 ng	0.1000	0.4949	20.00	200.00	1	
1	Std 200 ng	0.1000	0.5067	20.00	200.00	1	
1	Std 200 ng	0.1000	0.4971	20.00	200.00	1	

Figure 6-3. Calibration result screen.

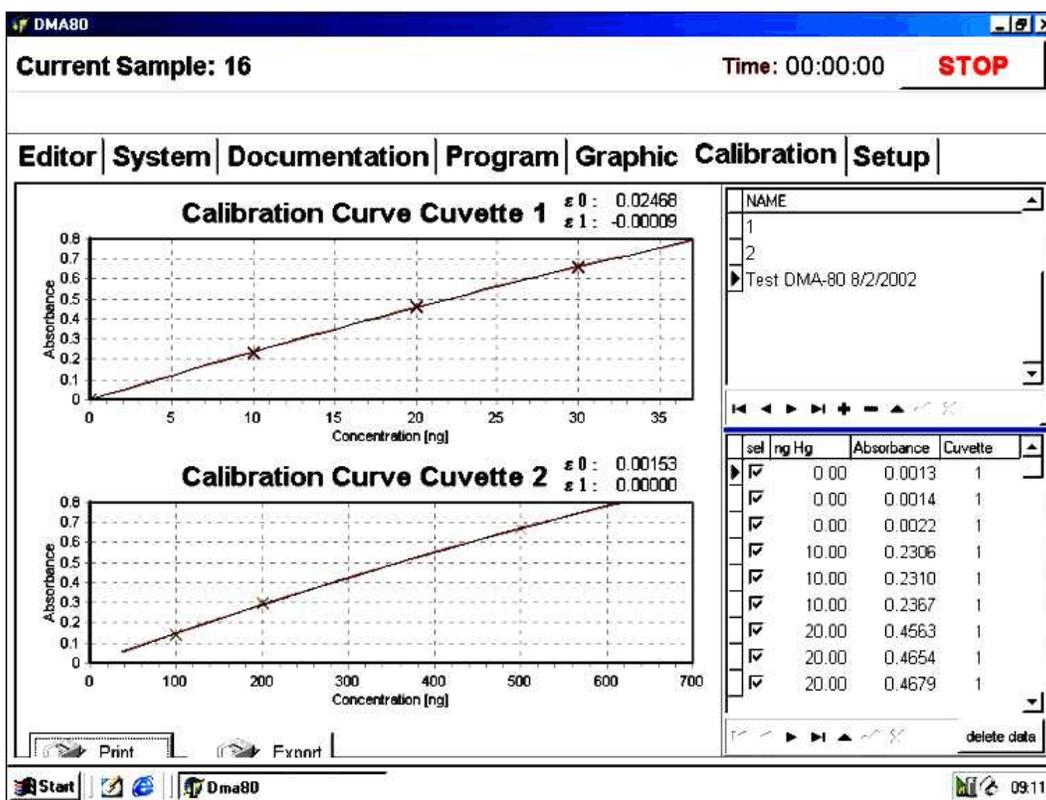


Figure 6-4. Calibration curve screen.

Items not covered in the SOP were trouble shooting and maintenance. For example, during the demonstration, the auto sampler jammed and the pneumatic sample insertion arm required realignment twice. Neither of these maintenance items was discussed in the SOP. In addition, there are two crucial operational elements that are not addressed in the SOP. The first is the selection of sample size such that the results will be within the calibration range. Selection of sample size requires an estimate of the expected mercury concentration. This problem is not unique to the DMA-80; any AAS instrument requires an estimate of sample concentration in order to get sample results within a specified calibration range. Second, no information was provided on how to handle samples that were outside of the calibration range. Procedures implemented during the demonstration included running a blank sample after a sample was above the calibration range (to purge the system of mercury) and reducing sample size on subsequent reanalyses (if quantitative results are reported). These procedures were not described in the SOP; however, the software prompted the analyst to run a clean-out blank. It is not known whether

these procedures are covered in the vendor training course.

Milestone chose to operate the DMA-80 with one chemist during the demonstration. The chemist held a B.S. degree in chemistry. Milestone claimed that a laboratory or field technician with a high school diploma and basic computer knowledge could operate the equipment after a 1-day training course. Field observations supported this claim. Most operations required either use of a keyboard or mouse with a Microsoft Windows-based system, or alternatively, the use of a touch screen with icons. The prompts and icons were clear and easy to understand.

The input screen includes a table which is pre-numbered to correspond to the auto sampler slots. The user enters the sample number in the first column and the sample identification or description in a second column. The operator then performs the sample weighing step and enters the sample weight or the sample weight can be automatically transferred from appropriate balances. A status column shows the sample analysis status.

The operator was able to perform sample preparation and analysis on a continuous basis. Sample preparation took approximately one minute per sample. Sample preparation consisted of mixing samples in the original container using a clean stainless steel spatula. A clean weigh boat was placed on the balance (not part of the system, but can be provided), the balance was zeroed, the weigh boat was removed from the balance, and a small amount of sample was placed in the weigh boat. The weigh boat and sample were placed on the balance again. The net weight was displayed on the digital balance and within the input screen for the DMA-80. When the weight stabilized, the operator input the weight by touching a screen icon for the scale. This operation was easy to understand and could be performed by a trained technician.

Sample analysis took 7.7 minutes per sample, on average. Because sample analysis was automated, sample preparation of additional samples continued during sample analysis of previous samples. Typically, three to four samples were prepared during the time it took to perform an analysis, allowing time for observation of equipment performance.

Sample analysis consisted of placing the pre-weighed sample boat in the proper slot on the auto sampler. The slot number corresponded to the number in the input screen. The samples were then automatically advanced as samples were processed. The auto sampler picked up the sample boat and inserted it into the furnace opening. When sample analysis was completed, the sample boat was automatically removed from the furnace and placed back on the auto sampler. The auto sampler then advanced the next sample for analysis. Because this process was automated, it was extremely easy to use. The only potential difficulty was ensuring that the sample boat was placed in the appropriate auto sampler slot so that the results matched with the proper sample number. As with sample preparation, sample analysis was easy to understand and could be performed by a trained technician.

As samples were analyzed, vendor-proprietary software screens allowed the user to track the approximate location of the sample mercury in a graphic display of the analyzer furnace, amalgamator, and photo cell (Figure 6-5).

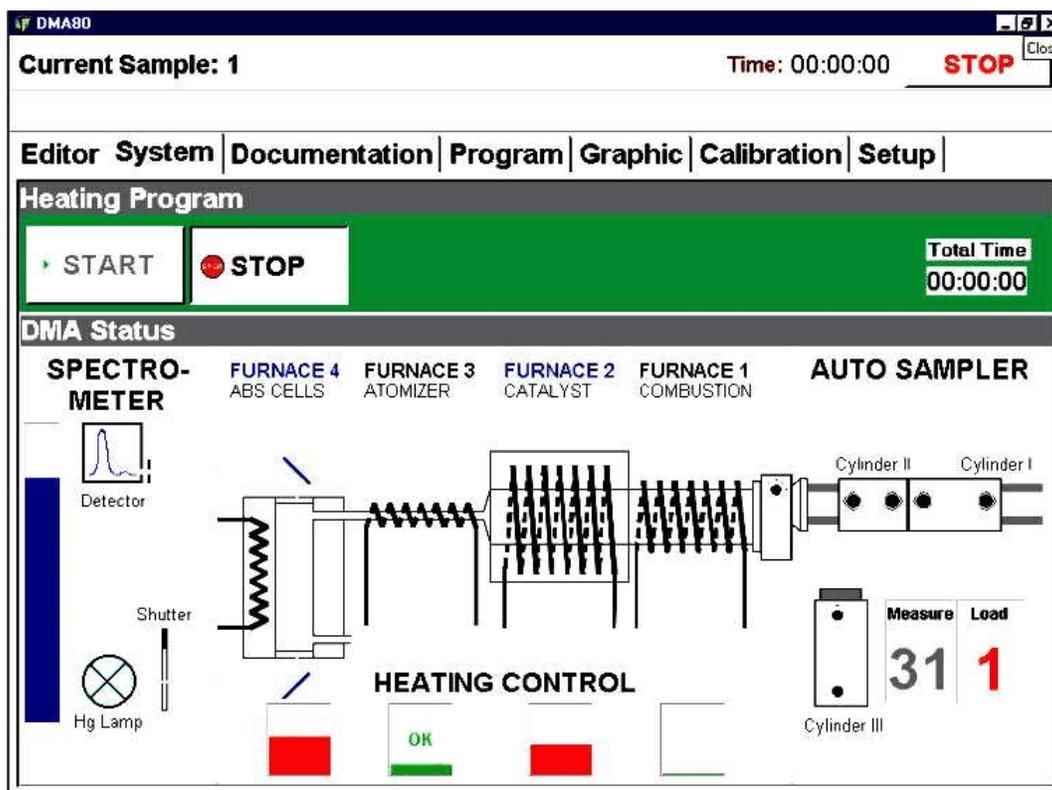


Figure 6-5. System control display screen.

During the demonstration, samples with concentrations outside of the equipment calibration/operation range were encountered. These samples would result in a peak outside of the calibration range. The vendor software flagged these samples (red "X" instead of a green "√" for "in range" samples) and prompted the user to run a blank

to demonstrate that excess mercury had been purged from the system. The software messages were clear and easy to follow. Another screen presented a graph of the adsorption peak. Figure 6-6 shows a representative peak for a sample that was "in range."

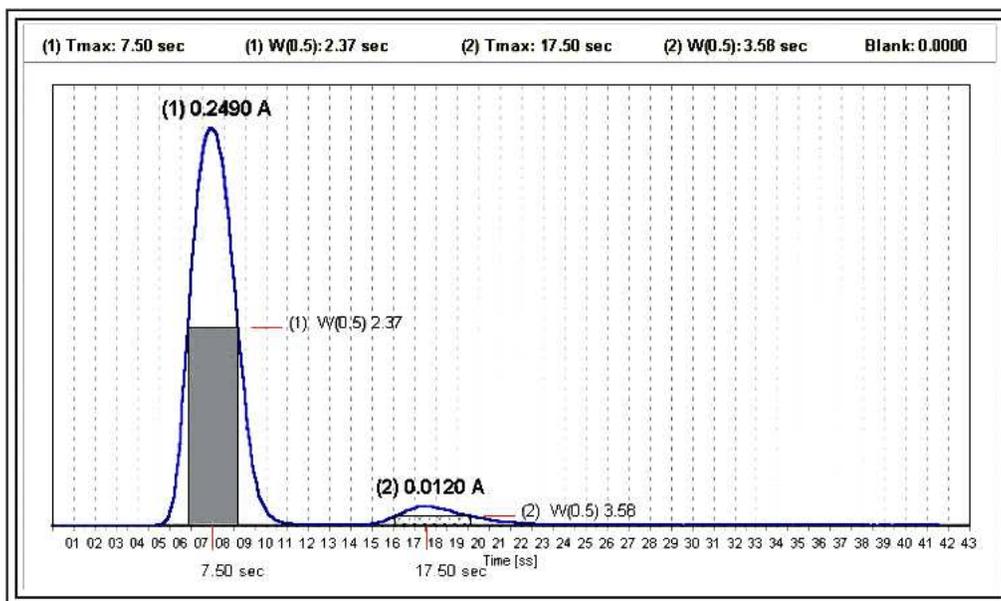


Figure 6-6. Sample peak screen.

The digital balance was the major peripheral item. The vendor will supply a balance with cables for direct input into the system monitor/software, or the user can supply his/her own balance. Though the balance is not part of the required vendor equipment, a balance is a necessary peripheral. Therefore, the balance was evaluated during the demonstration. The reader should note that other brands and models of balances may be used and these may not perform in the same manner as the balance used during the demonstration. The interface of the balance with the monitor/software was seamless. Overall, the balance was easy to use in conjunction with the DMA-80. According to the vendor, the sample weight is currently stored in a second database, requiring re-entry of the data into the main sample database with the corresponding potential for data entry errors. The vendor claims that a new edition of the software will eliminate the need to re-enter sample weights into the sample data base.

6.2.2 Health and Safety Concerns

Documents potential health and safety concerns associated with operating the device.

No significant health and safety concerns were noted during the demonstration. The only potential health and safety concerns identified were the generation of mercury vapors and the use of oxygen as the carrier gas. The vendor recommends and can provide a mercury filter; oxygen can be safely handled using standard laboratory procedures.

Health and safety concerns, including chemical hazards, radiation sources, electrical shock, explosion, and mechanical hazards were evaluated.

No chemicals were used in the preparation or processing of samples, except for analytical standards. During this demonstration, the analytical standards were soil SRMs for mercury. These were handled with gloves and the operator wore safety glasses with side shields at all times. Such standard laboratory precautions mitigate the potential for dermal exposure. Similar procedures were also used for soil samples which contained mercury. Because the DMA-80 is designed to thermally convert mercury compounds to mercury vapors as part of the analytical process, and no fume hood was present to exhaust mercury vapors after analysis, inhalation of mercury was a concern. The vendor installed a mercury trap, containing potassium permanganate, in the exhaust line from the DMA-80. Measurements were taken with a Jerome 431-x gold film mercury vapor analyzer, manufactured by Arizona Instruments Corporation. The instrument has a range of 0.000 to 0.999 mg/m³. In all cases, readings were 0.000 mg/m³ in the breathing zone of the operator.

In evaluating electrical shock potential, two factors were evaluated: 1) obvious areas where electrical wires are exposed and 2) safety certifications. No exposed wires were noted during the demonstration. All connections between equipment were made using standard electrical power cords, modem interface lines, and 9-pin cords. Power cords were grounded with ground fault interrupters, and a surge protector was utilized. The DMA-80 was not UL certified, but did have CE certification; no other safety certifications were marked on the transformer.

No obvious explosion hazards were noted. The use of oxygen as a carrier gas does present the possibility of explosion in the presence of ignition sources; however, implementation of good laboratory safety practices can mitigate any such hazard. The cylinder needs to be secured both when in use and when not in use. When not in use, the cylinder should be disconnected from the DMA-80 and the cap replaced to prevent damage to the cylinder valve. The cylinder was clearly marked as oxygen and the appropriate hazard label was present.

No serious mechanical hazards were noted during the demonstration. All equipment edges were smooth, minimizing any chance of cuts or scrapes. The hinged lid on the DMA-80 presents the possibility of a pinch hazard, as would any hinged device; however, the lid is not overly heavy, does not need to be routinely opened, and is designed to remain securely in place when the lid is open.

6.2.3 Portability of the Device

Documents the portability of the device.

The DMA-80 was not easily portable (by hand) due to its size and weight. It was easy to set up and can be taken anywhere that a small van or SUV can go. The instrument is better characterized as mobile rather than field portable.

The DMA-80 measured 80 cm (L) by 43 cm (W) by 30 cm (H). The weight was estimated at 45 kg. Also included as a standard feature with the DMA-80 were a controller with monitor and a keyboard; both were light weight and easily portable. The controller measured approximately 38 cm (L) by 23 cm (W) by 22 cm inches high.

The one negative aspect of the DMA-80, with respect to portability, was its size and weight. This equipment required the assistance of one SAIC person to unload from the transport vehicle to the table used during the demonstration. It should be noted that the DMA-80 can be used out of the back of the vehicle (SUV), and, in fact, was used this way on the second and third days of the demonstration. This device may be better characterized as a "mobile" instrument rather than "field portable". The device is not hand held, and can not be easily moved by hand from one location to another. That said, the DMA-80 can certainly be transported to any place that a small van or SUV can go and would be practical for most field applications.

The balance required a flat, stable surface. Because the width of the DMA-80 prevented placement of the balance in the SUV, a table was required. The vendor utilized a marble slab on the table to provide extra stability and simplify leveling the balance. A marble slab is not a part of the standard equipment supplied with the DMA-80. A flat surface area is also required for staging samples while filling weigh boats.

The DMA-80 is not equipped with a battery. Operation of the instrument requires a standard electrical source of 110 volts. The vendor asserts that the DMA-80 can be powered by a generator, although this was not evaluated during the demonstration.

For the demonstration, the vendor was supplied with a folding table, two chairs, and a tent to provide shelter from inclement weather. In addition, one 1-gallon container

each was provided for waste soil and decontamination water utilized to clean weigh boats. A 2-gallon zip-lock bag was furnished for disposal of used gloves, wipes, and other wastes which were contaminated during the demonstration. Finally, a large trash bag was supplied for disposal of non-contaminated wastes.

6.2.4 Instrument Durability

Evaluates the durability of the device based on its materials of construction and engineering design.

The DMA-80 was well designed and constructed for durability.

The outside of the DMA-80 is constructed of sturdy stainless steel. Parts were securely connected with screws and lock washers. The top of the device could be opened to access inner components. The lid was secured with stainless steel hinges. No environmental (e.g., corrosion) or mechanical (e.g., shear stress or impact) tests were performed; however, the outer shell of the instrument appeared to be well-designed and constructed, indicating that the device would likely be durable under field conditions.

No evaluation could be made regarding the long-term durability of the furnace, analytical cell, or circuitry. Visual inspection did not indicate that any problems were likely. The vendor offers a standard 1-year warranty and will provide an extended warranty and maintenance plan at the owner's cost.

Minor problems were identified during the demonstration with two moving parts on the auto sampler component of the system. During the first day of the demonstration, the vendor adjusted the alignment of the pneumatic arm used to take sample boats from the auto sampler and insert them into the furnace. The vendor explained that this alignment is frequently required after the instrument has been shipped. This alignment took less than 10 minutes to accomplish. During the second day of the demonstration, the 23rd sample was dropped by the auto sampler. The vendor indicated that this was due to minor alignment problems with the pneumatic arm. This resulted in approximately 5 minutes of downtime. Later that same day, the auto sample manifold jammed, causing the loss of one sample in the queue. (It should be noted that there was additional sample available, and a replacement weigh boat was prepared while the instrument ran other samples, resulting in no net downtime due to sample loss.) The auto

sampler jam resulted in approximately 5 to 7 minutes of downtime. Pneumatic pressure was released by the operator closing the oxygen tank and bleeding oxygen from the system. The auto sample manifold was then disengaged (it had been bent slightly by the pneumatic pressure when it jammed). The pneumatic arm was re-aligned by loosening set screws, aligning the arm, and resetting the set screws. The auto sampler manifold was reinserted, oxygen pressure reestablished, and the system operation tested with a blank. A new sample was running approximately 7 minutes after the jam originally occurred.

Finally, most of the demonstration was performed during rainfall events ranging from steady to torrential. The DMA-80 was located either under a tent (Days 1 and 4) or in the back of the SUV (Days 2 and 3). Even when it was not raining, the relative humidity was high, ranging from 70.6 to 98.3 percent. The high humidity and rainfall had no apparent impact on the reliability of the instrument operation.

6.2.5 Availability of Vendor Instruments and Supplies

Documents the availability of the device and spare parts.

The DMA-80 is readily available for lease, or purchase. DMA-80 rental is available on a limited basis. Spare parts and consumable supplies can be added to the original DMA-80 order or can be received within 24 to 48 hours of order placement. Supplies and standards not provided by Milestone are readily available from laboratory supply firms.

EPA representatives contacted Milestone regarding the availability of the DMA-80 and supplies. Milestone asserted that 95 percent of its current business is purchase or long-term lease arrangement. According to Milestone, such systems are available within 3 to 4 weeks of order placement, but can be expedited with a minimum 2-week turnaround. The DMA-80 also is available for rental on a limited basis (special requests). There is only one unit in the rental pool, so lead time is subject to availability.

The instrument comes standard with 40 weigh boats and a complete set of consumable items (catalyst,

amalgamator, and o-rings) installed in the instrument so that the instrument is fully operable upon receipt. Spare consumable items are available as part of a consumables kit or can be ordered individually. These and any other parts are available within 24-48 hours.

Other supplies and standards, not provided by Milestone, can be purchased from a laboratory supply firm. Typical delivery times, per Milestone, for most supplies will range from 1 day (with express delivery) to less than one week. Cost for capital equipment and supplies are discussed in Chapter 7.

Chapter 7

Economic Analysis

The purpose of the economic analysis was to estimate the total cost of mercury measurement at a hypothetical site. The cost per analysis was estimated; however, because the cost per analysis would decrease as the number of samples analyzed increased, the total capital cost was also estimated and reported. Because unit analytical costs are dependent upon the total number of analyses, no attempt was made to compare the cost of field analyses with the DMA-80 to the costs associated with the referee laboratory. "Typical" unit cost results, gathered from analytical laboratories, were reported to provide a context in which to review DMA-80 costs. No attempt was made to make a direct comparison between these costs for different methods because of differences in sample throughput, overhead factors, total equipment utilization factors, and other issues that make a head-to-head comparison impractical.

This Chapter describes the issues and assumptions involved in the economic analysis, presents the costs associated with field use of the DMA-80, and presents a cost summary for a "typical" laboratory performing sample analyses using the reference method.

7.1 Issues and Assumptions

Several factors can affect mercury measurement costs. Wherever possible in this Chapter, these factors are identified in such a way that decision-makers can independently complete a project-specific economic analysis. Milestone offers three options for potential DMA-80 users: 1) purchase of the instrument, 2) monthly rental (with a 3-month minimum), and 3) equipment leasing with an option to purchase at the end of 24 months. Because site and user requirements vary significantly, all

three of these options are discussed to provide each user with the information to make a case-by-case decision.

A more detailed cost analysis was performed on the equipment rental option for three months or less because this case represents the most frequently encountered field scenario. The results of that cost analysis are provided in section 7.2.

7.1.1 Capital Equipment Cost

The DMA-80 (the analytical instrument) comes complete with a 40-position auto sampler; Pentium Controller with keyboard, mouse, and touch screen monitor; Windows™ based software; and a set of stainless steel weigh boats, whether the instrument is purchased, rented, or leased. A portable computer may be substituted for the controller at the user's request. An optional digital balance, with sensitivity to 0.1 mg, is available for purchase from Milestone (no rental or leasing), but not included in the base cost of any of these three options. Alternatively, the user may provide his/her own balance. The vendor claims that virtually any balance adaptable to a 9-pin sub-D socket (female) interface cable will communicate with the DMA-80 (Milestone, 2003). This claim was not evaluated during the demonstration. A printer can also be purchased as an option from Milestone; no lease agreement or rental is available for the printer. Per the vendor, any Windows compatible printer can be used.

The cost quoted by Milestone does not include packaging or freight costs to ship the instrument to the user location. A 1-month, non-refundable deposit is required for rental and lease agreements. The deposit is not applied to payments. A user manual is provided at no cost. An 8-hour training session is available for an additional fee.

7.1.2 Cost of Supplies

The cost of supplies was estimated based on the supplies required to analyze demonstration samples, and based on discussions with Milestone. Requirements vary, depending upon whether solid or liquid samples are being analyzed. For purposes of this cost estimate, only supplies required to analyze solid samples are factored into the cost estimate. Supplies required for liquid samples are noted, and approximate prices provided, but those costs are not incorporated into the overall cost estimate because liquid samples were not analyzed during the demonstration. Supplies consisted of consumable items (e.g., standards and compressed oxygen) and non-consumables that could not be returned because they were contaminated or the remainder of a set. Non-consumable supplies consisted of a set of 3 micro-spatulas (for solid samples).

Consumable supplies consisted of:

- Adjustable micro-pipettes (for liquid samples)
- Housing and tubing for the mercury trap
- Calibration standards
- Compressed oxygen (welding grade)
- Potassium permanganate for mercury trap
- Glass wool for the mercury trap
- Silica gel for dilution of high-concentration samples

The purchase prices and supply sources were obtained from Milestone. Because the user cannot return unused or remaining portions of supplies, no salvage value was included in the cost of supplies. Personal protective equipment (PPE) supplies were assumed to be part of the overall site investigation or remediation costs; therefore, no PPE costs were included as supplies. During the demonstration, high-concentration samples generally were not quantified; they were usually reported as “greater than” values. In cases where the user wants to quantify high-concentration samples, a dilution material is needed. The vendor recommends silica gel. Even though silica gel was not used during the demonstration, it could have been used for high-concentration samples. (Milestone made the decision to not use silica gel during the demonstration and therefore this additional variation was not evaluated as part of overall instrument accuracy and precision.) Such samples are likely to be encountered at most other sites. Therefore, the cost was estimated based on the assumption that 25 percent of samples may have to be diluted.

7.1.3 Support Equipment Cost

During the demonstration, the DMA-80, controller, and balance were operated using AC power. The costs

associated with providing the power supply and electrical energy were not included in the economic analysis; the demonstration site provided AC power at no cost. None of the items mentioned above can operate on DC power, although a portable generator can be used to power the equipment.

Because of the large number of samples expected to be analyzed during the demonstration, EPA provided support equipment, including tables and chairs, for the field technician's comfort. In addition, the EPA provided a tent to ensure that there were no delays in the project due to inclement weather. These costs may not be incurred in all cases; however, such equipment is frequently needed in field situations, so these costs were included in the overall cost analysis.

7.1.4 Labor Cost

The labor cost was estimated based on the time required for DMA-80 setup, sample preparation, sample analysis, summary data preparation, and instrument packaging at the end of the day. Setup time covered the time required to take the instrument out of its packaging, setup all components, and ready the device for operation. However, the DMA-80 was brought to the site in a vehicle and was not in an original shipping container. Therefore, this time was estimated rather than measured. Sample preparation involved mixing samples with a micro-spatula. Other than the first couple of samples, sample preparation was easily completed while previous samples were being analyzed. Sample analysis was the time required to analyze all samples and submit a data summary. The data summary was strictly a tabulation of results in whatever form the vendor chose to provide. In this case, the vendor transcribed results from the electronic database to the field chain of custody forms (no printer was available in the field). The time required to perform all tasks was rounded to the nearest 5 minutes; however, for the economic analysis, times were rounded to the nearest hour and it was assumed that a field technician who had worked for a fraction of a day would be paid for an entire 8-hour day. Based on this assumption, a daily rate for a field technician was used in the analysis.

During the demonstration, EPA representatives evaluated the skill level required for the field technician to analyze and report results for mercury samples. Based on these field observations, a field technician with basic chemistry skills acquired on the job or in a university setting, and a 1-day training course specific to the DMA-80, was considered qualified to operate the instrument. For the economic analysis, an hourly rate of \$15 was used for a

field technician. A multiplication factor of 2.5 was applied to labor costs to account for overhead costs. Based on this hourly rate and multiplication factor, and an 8-hour day, a daily rate of \$300 was used for the economic analysis. Monthly labor rates are based on the assumption of an average of 21 work days per month. This assumes 365 days per year, and non work days totaling 113 days per year (104 weekend days and 9 holidays; vacation days are discounted assuming vacations will be scheduled around short-term work or staff will be rotated during long projects). Therefore, 252 total annual work days are assumed.

7.1.5 Investigation-Derived Waste Disposal Cost

Milestone was instructed to segregate its waste into four categories during the demonstration: 1) general trash; 2) lightly contaminated PPE and wipes; 3) contaminated soil (both analyzed and unanalyzed) and other highly contaminated wastes; and 4) wash water used for cleaning micro spatulas and weigh boats. General trash was not included as IDW and is not discussed in this document.

Lightly contaminated wastes consisted primarily of used surgical gloves and wipes. The surgical gloves were discarded for one of three reasons: 1) they posed a risk of cross contamination (noticeably soiled), 2) they posed a potential health and safety risk (holes or tears), or 3) the operator needed to perform other tasks (e.g., using cell phone to provide customer support). The rate of waste generation was in excess of what would be expected in a typical application of this instrument. In addition, the EPA evaluators occasionally contributed used gloves to this waste accumulation point. Wipes were used primarily to clean weigh boats and micro spatulas between samples. In cases where cross contamination is not a major concern (e.g., field screening or all samples are in the same concentration range), lesser amounts of waste would likely be generated.

Contaminated soils consisted primarily of soil placed in the weigh boat and then removed because the weight was above the target weight. Soil mass that was analyzed was also placed in this waste container as a precaution. It is expected that such soils would be free of mercury after being heated to high temperatures in the analytical instrument. In some cases, these sample residuals may not need to be handled as hazardous waste.

Finally, the vendor generated small amounts of waste water by cleaning weigh boats and micro spatulas. Weigh boats are considered clean after the completion of

analyses due to high temperatures in the analytical instrument. Therefore, weigh boats were not washed after analyses were completed; however, during the demonstration, the vendor was required to wash weigh boats for samples that were not analyzed (e.g., part of the sample spilled in placing the weigh boat on the auto sampler. The boats were rinsed with water and dried with a clean wipe to prevent potential cross contamination of low concentration samples.

The waste water, contaminated soil, excess sample material, and lightly contaminated gloves and wipes were considered hazardous wastes for purposes of this cost analysis.

7.1.6 Costs Not Included

Items for which costs were not included in the economic analysis are discussed in the following subsections, along with the rationale for exclusion of each.

Oversight of Sample Analysis Activities. A typical user of the DMA-80 would not be required to pay for customer oversight of sample analysis. EPA representatives observed and documented all activities associated with sample analysis during the demonstration. Costs for this oversight were not included in the economic analysis because they were project specific. For the same reason, costs for EPA oversight of the referee laboratory were also not included in the analysis.

Travel and Per Diem for Field Technician. Field technicians may be available locally. Because the availability of field technicians is primarily a function of the location of the project site, travel and per diem costs for field technicians were not included in the economic analysis.

Sample Collection and Management. Costs for sample collection and management activities, including sample homogenization and labeling, are site specific and, therefore, not included in the economic analysis. Furthermore, these activities were not dependent upon the selected reference method or field analytical tool. Likewise, sample shipping, COC activities, preservation of samples, and distribution of samples were specific requirements of this project that applied to all vendor technologies and may vary from site to site. None of these costs was included in the economic analysis.

Items Costing Less than \$10. The costs of inexpensive items, such as paper towels, were not included in the economic analysis.

Documentation Supplies. The costs for digital cameras used to document field activities were not included in project costs. These were considered project-specific costs that would not be needed in all cases. In addition, these items can be used for multiple projects. Similarly, the cost of supplies (logbooks, copies, etc.) used to document field activities was not included in the analysis because they are project specific.

Health and Safety Equipment. Costs for rental of the mercury vapor analyzer and the purchase of PPE were considered site specific and, therefore, were not included as costs in the economic analysis. Safety glasses and disposable gloves were required for sample handlers and would likely be required in most cases. However, these costs are not specific to any one vendor or technology. As a result, these costs were not included in the economic analysis.

Mobilization and Demobilization. Costs for mobilization and demobilization were considered site specific, and not factored into the economic analysis. Mobilization and demobilization costs actually impact laboratory analysis more than field analysis. When a field economic analysis is performed, it may be possible to perform a single mobilization and demobilization. During cleanup or remediation activities, several mobilizations, demobilizations, and associated downtime costs may be necessary when an off-site laboratory is used because of the wait for analytical results.

7.2 DMA-80 Costs

This subsection presents information on the individual costs of capital equipment, supplies, support equipment, labor, and IDW disposal for the DMA-80.

7.2.1 Capital Equipment Cost

During the demonstration, the DMA-80 was operated for approximately 3 days and was used to analyze 173 samples. Table 7-1 summarizes the DMA-80 capital costs for the three procurement options: rental, lease, and

purchase. Also shown are estimated costs for an optional printer and analytical balance. Figure 7-1 shows the relative costs for the basic capital equipment. These costs reflect the basic DMA-80 system (with standard auto sampler) and the controller/monitor. No options (e.g., balance or printer) and no supply or shipping costs are included. As would be expected, this chart clearly shows that leasing is the most cost-effective option (in terms of capital costs), followed by rental, for short-term projects. As project duration (or use on multiple projects) approaches two years, the purchase option is the most cost-effective. These scenarios cover only capital cost, not the cost of optional or user-supplied equipment, supplies, support equipment, labor, and IDW disposal.

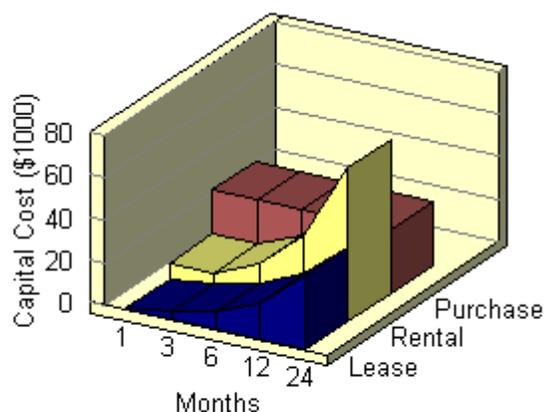


Figure 7-1. Capital equipment costs.

The DMA-80 sells for \$30,000, including the 40-position auto-sampler, the controller (with mouse, keyboard, display monitor, and software), and related electrical connections. Also included are:

- 40 stainless steel weigh boats
- Plastic tubing for compressed oxygen connections
- An instruction manual

Table 7-1. Capital Cost Summary for the DMA-80

Item	Quantity	Unit Cost (\$)	Total Cost for Selected Project Duration				
			1-Month	3-Month	6-Month	12-Month	24-Month
Purchase DMA-80	1	\$30,000	\$30,000	\$30,000	\$30,000	\$30,000	\$30,000
Monthly Rental of DMA-80 ^a	1	\$3,000	\$9,000	\$9,000	\$18,000	\$36,000	\$72,000
Monthly Lease of DMA-80 ^b	1	\$1,450	\$1,450	\$4,350	\$8,700	\$17,400	\$34,800
Purchase Balance (Optional) ^c	1	\$1,500	\$1,500	\$1,500	\$1,500	\$1,500	\$1,500
Purchase Printer (Optional) ^c	1	\$150	\$150	\$150	\$150	\$150	\$150

a Ten percent of purchase price with a three month minimum.

b \$1,450 per month (24-month lease with \$1 buyout).

c A balance is required, but may be provided by the user. A printer is optional; it may also be provided by the user.

These items are considered supplies and are discussed in Subsection 7.2.2. Compressed oxygen is required, but must be obtained from a local supplier, along with the appropriate regulator and cylinder mounting brackets (see Subsection 7.2.2). A balance is also required and can be purchased (no rental or lease) from Milestone for \$2950. A balance can be purchased from a laboratory supply company for approximately \$1500 to \$2,500, depending upon model (www1.fishersci.com, 2003). The lowest cost, \$1500, was used in this cost analysis. Alternatively, the user can supply a balance with a 9-pin connector to interface with the DMA-80 (Milestone, 2003). The costs presented in Figure 7-1 do not reflect the cost of the balance because it is optional equipment and can be provided by the user (it may already be owned). A printer can be purchased for approximately \$150; however, as with the balance, no printer costs are included in the cost analysis (or Figure 7-1) because this equipment is optional and may be supplied or already owned by the user. Balance and printer costs are shown in Table 7-1.

7.2.2 Cost of Supplies

Supplies used during the demonstration included solid SRMs, compressed oxygen, micro spatulas, and a mercury trap. NIST soil SRMs sell for \$250 each; typically both a high and a low standard will be required for many applications, for a total cost of \$500. If sediments are analyzed, a NIST sediment SRM may be obtained for \$150. No costs for a sediment SRM are included in this analysis. These standards have a life-expectancy of one to three years (one year is assumed for this cost analysis).

Welding grade compressed oxygen is used as a carrier gas for the DMA-80. It can be obtained from a local source and prices will vary. For this cost analysis, a price of \$0.04/L was used. An 80 ft³ (2,265 L) cylinder will last for

approximately 19 days, assuming 10 hours of constant operation of the DMA-80 at a flow rate of 200 mL/min. A regulator is required to reduce the flow rate to 200 mL/min. Purchase of a flow regulator and cylinder brackets is estimated at \$200, which is a one-time cost that can be spread over the entire term for longer projects. Alternatively, the cost can be included with the cylinder rental cost as was done for this analysis. Table 7.2 summarizes the costs for the carrier gas, assuming the same number of samples are run per day during each period and a total cost of \$0.04/L, which equals \$5/day of operation. The rental cost for a mounting bracket and pressure regulator is included in the \$0.04/L.

Table 7-2. Carrier Gas Cost Summary

Item	Months				
	1	3	6	12	24
Flow Regulator	NA	NA	NA	NA	NA
Mounting Bracket	NA	NA	NA	NA	NA
Oxygen	\$105	\$315	\$630	\$1,260	\$2,520
Total Cost	\$105	\$315	\$630	\$1,260	\$2,520

A mercury trap was also required during the demonstration and would likely be needed for most field applications. The trap consisted of a polyethylene drying tube (Fisher Scientific # SN:09-242C or equivalent), glass wool (Fisher Scientific # SN:11-390 or equivalent), and potassium permanganate as the reactive ingredient to “trap” mercury. The polyethylene drying tube costs \$20 for a package of 12; it is assumed that this will provide a sufficient supply for

up to 2 years. The glass wool costs \$35 for 454 g, enough for approximately 100 traps. Therefore, the cost is approximately \$0.35 per trap. The glass wool should be changed out every 3 months, so annual costs are \$1.40. Potassium permanganate costs \$60 for 500 g (enough for 6 traps), or \$10 per trap. Annual costs are \$40. Total trap costs are presented in Table 7-3.

Table 7-3. Mercury Trap Costs

Item	Months				
	1	3	6	12	24
Drying Tube	\$20	\$20	\$20	\$20	\$20
Glass Wool	\$35	\$35	\$35	\$35	\$35
KMnO4	\$10	\$10	\$20	\$40	\$80
Total	\$65	\$65	\$75	\$95	\$135

Two to three micro spatulas are normally required to prevent cross contamination and allow time for cleaning. A set of three micro spatulas costs \$10, and would be expected to last at least two years.

7.2.3 Support Equipment Cost

Milestone was provided with a 10x10 foot tent for protection from inclement weather during the demonstration. It was also provided with one table and two chairs for use during sample preparation and analytical activities. The rental cost for the tent (including detachable sides, ropes, poles, and pegs) was \$270 per week. The rental cost for the table and two chairs for one week totaled \$6. Total support equipment costs were \$276 per week for rental.

For longer projects, purchase of support equipment should be considered. Two folding chairs would cost approximately \$40. A 10x10 foot tent would cost between \$260 and \$1,000, depending on the construction materials and the need for sidewalls and other accessories (e.g., sand stakes, counter weights, storage bag, etc.). A cost of \$800 was used for this cost analysis. A folding table would cost between \$80 and \$250, depending on the supplier. For purposes of this cost analysis, \$160 was used. Total purchase costs for support equipment are estimated at \$1,000.

7.2.4 Labor Cost

One field technician was required for 3 days during the demonstration to complete sample analyses and prepare a data summary. Based on a labor rate of \$300 per day,

total labor cost for application of the DMA-80 was \$900 for the 3-day period. Labor costs assume qualified technicians are available locally, and that no hotel or per diem costs are applicable. Table 7-4 summarizes labor costs for various operational periods. The costs presented do not include supervision and quality assurance because these would be associated with use of any analytical instrument and are a portion of the overhead multiplier built into the labor rate.

Table 7-4. Labor Costs

Item	Months				
	1	3	6	12	24
Technician	\$6,300	\$18,900	\$37,800	\$75,600	\$151,200
Supervisor	NA	NA	NA	NA	NA
Quality Control	NA	NA	NA	NA	NA
Total	\$6,300	\$18,900	\$37,800	\$75,600	\$151,200

7.2.5 Investigation-Derived Waste Disposal Cost

Milestone generated PPE waste, decontaminate solution waste, and excess soil waste. The PPE waste was charged to the overall project due to project constraints. The minimum waste volume is a 5-gallon container. Mobilization and container drop-off fees were \$1,040; a 5-gallon soil waste drum was \$400, and a 5-gallon liquid waste drum was \$400. (These costs were based on a listed waste stream with hazardous waste number U151). The total IDW disposal cost was \$1,840. These costs may vary significantly from site to site, depending on whether the waste is classified as hazardous or nonhazardous and whether excess sample material is generated that requires disposal. Table 7-5 presents IDW costs for various operational periods, assuming that waste generation rates were similar to those encountered during the demonstration.

7.2.6 Summary of DMA-80 Costs

The total cost for performing mercury analysis is summarized in Table 7-6. This table reflects costs for projects ranging from one to 24 months. The rental option was used for estimating the equipment cost.

However, because the minimum rental for the DMA-80 is 3 months, the total cost is inflated by the high capital cost. Additionally, capital costs for rental exceed those for

purchase at approximately 10 months, so rental is no longer as cost-effective for projects exceeding this duration. Finally, a lease agreement may be a cost-effective alternative as compared to either rental or purchase for projects lasting less than 21 months. At that point, equipment purchase may be more-cost-effective; however, the decision on which purchase option to utilize should be made on a case-by-case basis.

Table 7-5. IDW Costs

Item	Months				
	1	3	6	12	24
Drop Fee	\$1,040	\$3,120	\$6,240	\$12,480	\$24,960
Disposal	\$400	\$1,200	\$2,400	\$4,800	\$9,600
Total	\$1,440	\$4,320	\$8,640	\$17,280	\$34,560

Table 7-6. Summary of Rental Costs for the DMA-80

Item	Quantity	Unit	Unit Cost (\$)	Total Cost for Selected Project Duration ^a				
				1-Month	3-Month	6-Month	12-Month	24-Month
Capital Equipment								
Monthly Rental of DMA-80	1	NA	\$3,000	\$3,000	\$9,000	\$18,000	\$36,000	\$72,000
Supplies								
Micro Spatula (set of 3) ^b	1	set	\$10	\$10	\$10	\$10	\$10	\$10
Solid SRM ^c	2	each	\$250	\$500	\$500	\$500	\$1,000	\$1,500
Mercury Trap (all components)	1	each	NA	\$65	\$65	\$75	\$95	\$135
Compressed Oxygen ^d	1	L	\$0	\$105	\$315	\$630	\$1,260	\$2,520
Total Supply Cost	—	-----	-----	\$680	\$890	\$1,215	\$2,365	\$4,165
Support Equipment ^e								
Table (optional) - weekly	1	each	\$5	\$20	\$60	\$120	\$160	\$160
Chairs (optional) - weekly	2	each	\$1	\$10	\$25	\$40	\$40	\$40
Tent (for inclement weather only) - weekly	1	each	\$270	\$800	\$800	\$800	\$800	\$800
Total Support Equipment Cost	--	-----	-----	\$830	\$885	\$960	\$1,000	\$1,000
Labor								
Field Technician (person day)	1	hour	\$38	\$6,300	\$18,900	\$37,800	\$75,600	\$151,200
IDW								
Drop Fee	NA		\$1,040	\$1,040	\$3,120	\$6,240	\$12,480	\$24,960
Disposal	NA	week	\$400	\$400	\$1,200	\$2,400	\$4,800	\$9,600
Total IDW Costs	—	-----	-----	\$1,440	\$4,320	\$8,640	\$17,280	\$34,560
Total Cost				\$12,250	\$33,995	\$66,615	\$132,245	\$262,925

a Other than unit costs, all costs are rounded to the nearest \$5.

b For solid samples and SRMs.

c Only for use with solid samples; assumes two SRMs are required (a low and a high standard) with a life expectancy of 1 year (some standards will have longer shelf lives). Liquid standards are also available and are generally less expensive.

d Assumes rental of the cylinder, regulator, and mounting bracket for all time periods, plus the cost of oxygen consumed.

e Rental costs were used through the 3-month period for chairs and the 6-month period for the table. Purchase costs were used for longer periods. Purchase costs for the tent were used for all periods.

Table 7-7 summarizes costs for the actual demonstration. Note that the 3-month rental cost of the DMA-80 was used for capital costs.

Table 7-7. DMA-80 Costs by Category

Category	Category Cost (\$)	Percentage of Total costs
Instrument Cost	\$3,000	48.3%
Supplies	\$590	9.5%
Support Equipment	\$280	4.5%
Labor	\$900	14.5%
IDW Disposal	\$1,440	23.2%
Total	\$6,210	100.0%

The cost per analyses based upon 173 samples when renting the DMA-80 is \$35.90 per sample. The cost per analysis for the 173 samples, excluding instrument cost is \$18.55 per sample.

7.3 Typical Reference Method Costs

This Section presents costs associated with the reference method used to analyze the demonstration samples for mercury. Costs for other project analyses are not covered. The referee laboratory utilized SW-846 Method 7471B for all soil and sediment samples. The referee laboratory performed 421 analyses over a 21-day time period.

A typical mercury analysis cost, along with percent moisture for dry-weight calculation, is approximately \$35. This cost covers sample management and preparation, analysis, quality assurance, and preparation of a data package. The total cost for 173 samples at \$35 would be \$6,035. This is based on a standard turnaround time of 21 calendar days. The sample turnaround time from the laboratory can be reduced to 14, 7, or even fewer calendar days, with a cost multiplier between 125% to 300%, depending upon project needs and laboratory availability. This results in a cost range from \$6,035 to \$18,105. The laboratory cost does not include sample packaging, shipping, or downtime caused to the project while awaiting sample results.

Chapter 8

Summary of Demonstration Results

As discussed previously in this ITVR, the Milestone DMA-80 was evaluated by having the vendor analyze 173 soil and sediment samples. These 173 samples consisted of both medium- and low-concentration field samples from three sites, SRMs, and spiked field samples. Table 8-1 provides a breakdown of the numbers of these samples for each sample type and concentration range or source. Collectively, these samples provided the different matrices, concentrations, and types of mercury needed to perform a comprehensive evaluation of the DMA-80.

8.1 Primary Objectives

The primary objectives of the demonstration were centered on evaluation of the field instrument and performance in relation to sensitivity, accuracy, precision, time for analysis, and cost. Each of these objectives was discussed in detail in previous chapters and is summarized in the following paragraphs. The overall demonstration results suggest that the experimental design was successful for evaluation of the Milestone DMA-80. Quantitative results were reviewed and this instrument was found to be very comparable to standard analyses performed by the laboratory and the collected data provide the evidence to support this statement.

The two primary sensitivity evaluations performed for this demonstration were the MDL and PQL. Following procedures established in 40 CFR Part 136, the MDL for the DMA-80 is likely between 0.049 and 0.068 mg/kg. The equivalent MDL for the referee laboratory is 0.0026 mg/kg. The calculated MDL is only intended as a statistical estimation and not a true test of instrument sensitivity.

The PQL for the DMA-80 is likely somewhere around 0.082 mg/kg based upon the analysis of several low concentration standard reference materials. Both the MDL and PQL were

determined for soils and sediments and the instrument may be capable of measuring lower concentrations for aqueous samples; however, this is not tested during the demonstration. The referee laboratory PQL, determined as part of the laboratory analysis was 0.005 mg/kg based upon a lower calibration standard. The %D is < 10%.

Accuracy was evaluated by comparison to SRMs and comparison to the referee laboratory analysis for field samples. This included spiked field samples for evaluation of additional concentrations not otherwise available. Milestone data compared to SRM values were within expected accuracy determinations. ALSI data compared to SRM values were also within expected accuracy determinations. (DMA-80 results were within SRM 95% prediction intervals 93% of the time, and referee laboratory results were within SRM 95% prediction intervals 96% of the time.) Comparison of the Milestone to the referee laboratory data for all field and spiked samples (including SRMs) based upon hypothesis testing at the $\alpha = 0.01$ level suggest that the two data sets are not dissimilar. Additional aggregate analysis for all collected data also suggests that the two data sets are not dissimilar.

The number of Milestone average values less than 30% different from the referee laboratory results or SRM reference values; however, was 16 of 30 different sample lots. Only 2 of 30 Milestone average results have relative percent differences greater than 100% for this same group of samples; however, when making the comparison between Milestone and ALSI data, and taking into account the possible bias associated with both sets of data, this comparison may be within reasonable expectations for considering these two separate analyses to be equivalent. Therefore, it could be concluded that the Milestone DMA-80 was within the expected accuracy for analysis of mercury in soil comparable to laboratory Method 7471B.

Precision was determined by analysis of replicate samples. The single most important measure of precision provided, overall average RSD, is 23.7% for the referee laboratory compared to the Milestone average RSD of 19.4%. Both of these RSDs are within the predicted 25% RSD objective for precision. The precision of the Milestone field instrument is therefore very comparable to laboratory precision, and within expected precision variation for soil and sediment matrices. Precision was not affected by sample concentration or matrix.

Time measurements were based on the length of time the operator spent performing all phases of the analysis, including setup, calibration checks, and sample analysis (including all reanalysis). Milestone analyzed 173 samples in 1,330 minutes over four days, which averaged to 7.7 minutes per sample result. Based on this, an operator could be expected to analyze 62 samples (8 hours x 60 minutes ÷ 7.7 minutes/sample) in an 8-hour day.

Cost of the Milestone sample analyses included capital, supplies, labor, support equipment, and waste disposal. The cost per sample was calculated both with and without the cost of the instrument included. This was performed because the first sample requires the instrument purchase, and as the sample number increases, the cost per sample would decrease. A comparison of the field Milestone cost to off-site laboratory cost was not made. To compare the field and laboratory costs correctly, it would be necessary to include the expense to the project while waiting for analyses to return from the laboratory (potentially several mobilizations and demobilizations, stand-by fees, and other aspects associated with field activities). Table 8-2 summarizes the results of the primary objectives.

8.2 Secondary Objectives

Table 8-3 summarizes the results of the secondary objectives.

Table 8-1. Distribution of Samples Prepared for Milestone and the Referee Laboratory

Site	Concentration Range	Soil	Sample Type		SRM
			Sediment	Spiked Soil	
Carson River (Subtotal = 75)	Low (1-500 ppb)	7	10	7	7
	Mid (0.5-50 ppm)	9	0	14	21
	High (50->1,000 ppm)	0	0	0	0
Puget Sound (Subtotal = 57)	Low (1 ppb - 10 ppm)	26	0	14	17
	High (10-500 ppm)	0	0	0	0
Oak Ridge (Subtotal = 41)	Low (0.1-10 ppm)	17	3	7	14
	High (10-800 ppm)	0	0	0	0
Subtotal		72	13	42	70

Table 8-2. Summary of DMA-80 Results for the Primary Objectives

Demonstration Objective	Evaluation Basis		Performance Results	
			DMA-80	Reference Method
Instrument Sensitivity	MDL.	Method from 40 CFR Part 136.	Between 0.049 and 0.068 mg/kg	0.0026 mg/kg
	PQL.	Low concentration SRMs or samples.	Approximately 0.082 mg/kg	0.005 mg/kg
Accuracy	Comparison to SRMs, field, and spiked samples covering the entire range of the instrument calibration.		Milestone's DMA-80 is within expected accuracy for laboratory analysis. Milestone's field instrument is very comparable to the referee laboratory analytical method, 7471B.	
Precision	Determined by analysis of replicate samples at several concentrations.		Overall RSD was computed to be 19.4% compared to the referee laboratory RSD of 23.7%. This is a combined measure of precision which includes sampling and aliquoting variations. Milestone's precision is comparable to the laboratory analysis and is not affected by matrix or concentration.	
Time per Analysis	Timed daily operations for 4 days and divided the total time by the total number of analyses.		One technician performed all setup, calibration checks, sample preparation and analysis, and equipment demobilization. Individual analyses took 5 minutes each, but the total time per analysis averaged approximately 8 minutes per sample.	
Cost	Costs were provided by Milestone and independent suppliers of support equipment and supplies. Labor costs were estimated based on a salary survey. IDW costs were estimated from the actual costs encountered at the Oak Ridge demonstration.		The cost per analyses based upon 173 samples, when renting the DMA-80, is \$35.90 per sample. The cost per analyses for the 173 samples, excluding capital cost, is \$18.55 per sample. The total cost for equipment rental and necessary supplies during the demonstration is estimated at \$6,210. The cost breakout by category is: capital costs, 48.3%; supplies, 9.5%; support equipment, 4.5%; labor, 14.5%; and IDW, 23.2%.	

Table 8-3. Summary of DMA-80 Results for the Secondary Objectives

Demonstration Objectives	Evaluation Basis	Performance Results
Ease of Use	Field observations during the demonstration.	The DMA-80 is easy to operate, requiring one field technician with a basic knowledge of chemistry acquired on the job or in a university, and training on the DMA-80.
Health and Safety Concerns	Observation of equipment, operating procedures, and equipment certifications during the demonstration.	No significant health and safety concerns were noted during the demonstration. The only potential health and safety concerns identified were the generation of mercury vapors and the use of oxygen as the carrier gas. The vendor recommends and can provide a mercury filter; oxygen can be safely handled using standard laboratory procedures.
Portability of the Device	Review of device specifications, measurement of key components, and observation of equipment setup and tear down before, during, and after the demonstration.	The DMA-80 was not easily portable (by hand) due to its size and weight. It was easy to set up and can be taken anywhere that a small van or SUV can go. The instrument is better characterized as mobile rather than field portable.
Instrument Durability	Observation of equipment design and construction, and evaluation of any necessary repairs or instrument downtime during the demonstration.	The DMA-80 was well designed and constructed for durability.
Availability of Vendor Instruments and Supplies	Review of vendor website and telephone calls to the vendor after the demonstration.	The DMA-80 is readily available for lease, or purchase. DMA-80 rental is available on a limited basis. Spare parts and consumable supplies can be added to the original DMA-80 order or can be received within 24 to 48 hours of order placement. Supplies and standards not provided by Milestone are readily available from laboratory supply firms.

Chapter 9 Bibliography

- Anchor Environmental. 2000. *Engineering Design Report, Interim Remedial Action Log Pond Cleanup/Habitat Restoration Whatcom Waterway Site*, Bellingham, W.A. Prepared for Georgia Pacific West, Inc. by Anchor Environmental, L.L.C., Seattle, WA. July 31, 2000.
- Confidential Manufacturing Site. 2002. Soil Boring Data from a Remedial Investigation Conducted in 2000.
- Milestone, 2003. *Getting Ready for the DMA-80 Direct Mercury Analyzer from Milestone Inc.*, Instrument SOP. June, 2003.
- Rothchild, E.R., R.R. Turner, S.H. Stow, M.A. Bogle, L.K. Hyder, O.M. Sealand, H.J. Wyrick. 1984. *Investigation of Subsurface Mercury at the Oak Ridge Y-12 Plant*. Oak Ridge National Laboratory, ORNL/TM-9092.
- U.S. Environmental Protection Agency. 1994. Region 9. *Human Health Risk Assessment and Remedial Investigation Report - Carson River Mercury Site* (Revised Draft). December 1994.
- U.S. Environmental Protection Agency. 1995. *Contaminants and Remedial Options at Selected Metal-Contaminated Sites*. July 1995. Washington D.C., EPA/540/R-95/512.
- U.S. Environmental Protection Agency. 1996. *Test Methods for Evaluating Solid Waste, Physical/Chemical Methods*, SW-846 CD ROM, which contains updates for 1986, 1992, 1994, and 1996. Washington D.C.
- U.S. Environmental Protection Agency. 1998. Unpublished. *Quality Assurance Project Plan Requirements for Applied Research Projects*, August 1998.
- U.S. Department of Energy. 1998. *Report on the Remedial Investigation of the Upper East Fork of Poplar Creek Characterization Area at the Oak Ridge Y-12 Plant*, Oak Ridge, TN. DOE/OR/01-1641&D2.
- U.S. Environmental Protection Agency. 2002a. Region 9 Internet Web Site, www.epa.gov/region9/index.html.
- U.S. Environmental Protection Agency. 2002b. *Guidance on Data Quality Indicators*. EPA G-5i, Washington D.C., July 2002.
- U.S. Environmental Protection Agency. 2003. *Field Demonstration and Quality Assurance Project Plan - Field Analysis of Mercury in Soil and Sediment*. August 2003. Washington D.C., EPA/600/R-03/053.
- Welding Supply, 2003. Price List for Oxygen and Pressure Regulators, www.weldingsupply.com. July 2003.
- Wilcox, J.W., Chairman. 1983. *Mercury at Y-12: A Summary of the 1983 UCC-ND Task Force Study*. Report Y/EX-23, November 1983.
- www.fishersci.com, 2003.
- www.milestonesci.com, 2003.

Appendix A Milestone Comments

In the present work, the method detection limit (MDL) was determined from the data collected during the Demonstration exercise.

To demonstrate the versatility and to validate the performance of the instrument, samples were processed, as provided by the EPA and in the order presented by the EPA. As a result, low concentration samples were processed after high concentration samples, and vice versa. As in most analytical instruments, a memory effect

can occur when low level samples are processed immediately after high level samples. This effect can be completely mitigated by operator technique.

When the MDL is determined by running 7 replicates of a low level soil standard, in sequence, the MDL is found to be 8 ppb. To obtain optimum results for such low level samples in the field, good operating technique would require the technician to process several blanks prior to running the low level samples.

This appendix was written solely by Milestone. The statements presented in this appendix represent the developer's point of view and summarize the claims made by the developer regarding the DMA-80. Publication of this material does not represent EPA's approval or endorsement of the statements made in this appendix; performance assessment and economic analysis results for the DMA-80 are discussed in the body of the ITVR.

Appendix B Statistical Analysis

Two separate hypothesis tests were used to compare the referee laboratory samples to the vendor tested samples. This appendix details the equations and information for both of these statistical analyses. For purposes of this appendix, we have chosen to call the test comparing sample populations using a separate calculation for each sample lot the “hypothesis test,” and the statistical comparison of the entire sample set (all 28 separate sample lots) analyzed by the vendor and the laboratory the “unified hypothesis test,” also known as an “aggregate analysis” for all of the sample lots.

Hypothesis Test

A hypothesis test is used to determine if two sample populations are significantly different. The analysis is performed based on standard statistical calculations for hypothesis testing. This incorporates a comparison between the two sample populations assuming a specified level of significance. For establishing the hypothesis test, it was assumed that both sample sets are equal. Therefore, if the null hypothesis is rejected, then the sample sets are not considered equal. This test was performed on all sample lots analyzed by both Milestone and the referee laboratory. H_0 and H_a , null and alternative hypothesis respectively, were tested with a 0.01 level of significance (LOS). The concern related to this test is that, if two sample populations have highly variable data (poor precision), then the null hypothesis may be accepted because of the test’s inability to exclude poor precision as a mitigating factor. Highly variable data results in wider acceptance windows, and therefore, allows for acceptance of the null hypothesis. Conclusions regarding this analysis are presented in the main body of the report.

To determine if the two sample sets are significantly different, the absolute value of the difference between the

laboratory average \bar{x}_L and the vendor average \bar{x}_V is compared to a calculated μ . When the absolute value of the difference is greater than μ , then the alternate hypothesis is accepted, and the two sets (laboratory and vendor) are concluded to be different.

To calculate μ , the variances for the laboratory data set and the vendor data set are calculated by dividing their standard deviations by the number of samples in their data set. The effective number of degrees of freedom is then calculated.

$$f = \frac{(V_L + V_V)^2}{\left(\frac{V_L^2}{n_L + 1}\right) + \left(\frac{V_V^2}{n_V + 1}\right)} - 2$$

Where:

- f = effective number of degrees of freedom
- V_L = variance for the laboratory results
- n_L = number of samples for the laboratory data set
- V_V = variance for the vendor results
- n_V = number of samples for the vendor data set.

The degrees of freedom (f) is used to determine the appropriate “t” value and used to calculate μ at the 0.01 level of significance using the following:

$$\mu = t_{1-\{0.005/2\}} \sqrt{V_L + V_V}$$

Unified Hypothesis Test

For a specified vendor, let Y_{ij} be the measured Hg concentration for the j^{th} replicate of the i^{th} sample for $i=1,2,\dots,I$ and $j=1,2,\dots,J_i$. Let $X_{ij} = \log(Y_{ij})$, where log is the logarithm to the base 10. Define $\bar{x}_{i\log}$ to be the average over all log replicates for the i^{th} sample given by:

$$\bar{X}_{i\log} = J_i^{-1} \log \sum_{j=1}^{J_i} X_{ij}$$

Denote the estimate of the variance of the log replicates for the i^{th} sample to be:

$$s^2 = \left(\sum_{i=1}^I (J_i - 1) \right)^{-1} \log \sum_{i=1}^I \sum_{j=1}^{J_i} (X_{ij} - \bar{X}_{i\log})^2$$

Now for the reference laboratory, let Y'_{ij} be the measured Hg concentration for the j^{th} replicate of the i^{th} sample for $i=1,2,\dots,I'$ and $j=1,2,\dots,J'_i$. Denote the reference laboratory quantities X'_{ij} , \bar{x}'_i , and s'^2 defined in a manner similar to the corresponding quantities for the vendor.

Assumptions: Assume that the vendor measurements, Y_{ij} , are independent and identically distributed according to a lognormal distribution with parameters μ_i and σ^2 . That is, $X_{ij} = \log(Y_{ij})$ is distributed according to a normal distribution with expected value μ_i and variance σ^2 . Further, assume that the reference laboratory measurements, Y'_{ij} , are independent and identically distributed according to a lognormal distribution with parameters μ'_i and σ'^2 .

The null hypothesis to be tested is:

$$H_0: \mu_i = \mu'_i + \delta, \text{ for some } \delta \text{ and } i = 1, \dots, I$$

against the alternative hypothesis that the equality does not hold for at least one value of i .

The null hypothesis H_0 is rejected for large values of:

$$\chi^2_{I-1} = \frac{\sum_{i=1}^I (\bar{X}_{i\log} - \bar{X}'_{i\log} - \delta)^2 + (J_i^{-1} + J'^{-1}_i)}{s^2_{pool}}$$

Where χ^2_{I-1} is approximately a chi-square random variable with $(I-1)$ degrees of freedom:

$$\delta = I^{-1} \log \sum_{i=1}^I (\bar{X}_{i\log} - \bar{X}'_{i\log})$$

and

$$s^2_{pool} = \frac{s^2 \log \sum_{i=1}^I (J_i - 1) + s'^2 \log \sum_{i=1}^{I'} (J'_i - 1)}{\sum_{i=1}^I (J_i - 1) + \sum_{i=1}^{I'} (J'_i - 1)}$$

Critical values for the hypothesis test are the upper percentile of the chi-square distribution with $(I-1)$ degrees of freedom obtained from a chi-square table.

Results of Unified Hypothesis Test for Milestone

Milestone and ALSI both supplied multiple assays on replicates derived from a total of 28 different sample lots, either field materials or reference materials. The Milestone and ALSI data from these assays formed the basis of this assessment.

The statistical analysis is based on log-transformed (logarithm base 10) data and uses a chi-square test for equality of Milestone and ALSI population means for given a sample lot. Equality of variances is assumed. A description of the statistical procedure is provided below.

Initially, the hypotheses tested was that, on average, Milestone and ALSI would produce the same results within a given sample lot. This hypotheses is stated as:

$$H_{10}: (\text{Milestone lot log mean}) = (\text{ALSI lot log mean})$$

H_{10} was strongly rejected in that the chi-square statistic was 140.2, which exceeds the upper 99th percentile of the chi-square distribution with 28 degrees of freedom having value of 48.3.

The null hypotheses was rejected in part because Milestone results tended to exceed those from ALSI for the same sample lot. To explore this effect, the null hypothesis was revised to included a bias term in the form of:

$$H_{20}: (\text{Milestone lot log mean}) = (\text{ALSI lot log mean}) + (\text{delta}).$$

Where delta is a single value that does not change from one sample lot to another, unlike the lot log means. H_{20} was rejected, in that the chi-square statistic was 63.9, which exceeded the upper 99th percentile of the chi-square distribution with 27 degrees of freedom with value of 47.0. In this analysis, delta was estimated to be 0.12 in logarithmic (base 10) space, which indicates an average upward bias for Milestone of $10^{0.12}=1.318$ or about 32%.

Milestone analytical results for sample lot 37 were large relative to the concentration provided with the sample reference material, and the Milestone data for sample lot 37 made a substantial contribution to the chi-square

statistic.

Accordingly, excluding sample lot 37 from the data set resulted in a chi-square statistic of 42.1, which does not exceed the upper 99th percentile of the chi-square distribution with 26 degrees of freedom with value 45.6. So, with excluding sample lot 37 data, one fails to reject H_{20} at the 99th percent level. In this analysis, delta was estimated to be 0.11 in logarithmic (base 10) space, which indicates an average upward bias of $10^{0.11}=1.288$ or about 29%.

Note further, that excluding sample lots 37 and 18 resulted in accepting H_{20} at the 95% level, and excluding sample lots 37, 18 and 39 resulted in accepting H_{20} at the 90% level. Summary information on these analyses is provided in Table 6-5.

Table B-1. Summary of Unified Hypothesis Test

Hypothesis	Total Sample Lots	Excluded Lot	DF	s^2_{pool}	Delta	Chi-square	P-value
H_{10}	28	None	28	0.01562	0.0000	140.211	0.000000
H_{20}	28	None	27	0.01562	0.1230	63.901	0.000079
H_{20}	28	37	26	0.01369	0.1104	42.085	0.024058
H_{20}	28	37, 18	25	0.01411	0.1172	36.124	0.069744
H_{20}	28	37, 18, 39	24	0.01410	0.1228	31.304	0.145220

DF = Degrees of Freedom
 s^2 = variance