COMMONWEALTH OF MASSACHUSETTS
EXECUTIVE OFFICE OF ENVIRONMENTAL AFFAIRS
DEPARTMENT OF ENVIRONMENTAL PROTECTION
ONE WINTER STREET, BOSTON, MA 02108 617-292-5500

MITT ROMNEY
Governor

KERRY HEALEY
Lieutenant Governor

STEPHEN R. PRITCHARD
Secretary

ROBERT W. GOLLEDGE, Jr.
Commissioner

MassDEP

Report on Results of the 2004-2005 Double-Blind Laboratory Evaluation Program

December 13, 2005

Massachusetts Department of Environmental Protection
1 Winter Street
Boston, MA 02180

http://www.mass.gov/dep

Executive Summary

During 2004 and 2005, the Massachusetts Department of Environmental Protection (MassDEP) conducted a large double-blind laboratory evaluation study, involving 19 commercial laboratories that provide the majority of analytical support services to parties assessing and cleaning up hazardous waste sites in Massachusetts. A "double-blind" study is one in which a laboratory is unaware that they have been sent samples that contain known concentrations of contaminants. The study was undertaken by MassDEP as part of a multi-year/multi-component data enhancement effort, in order to obtain a *direct, real world* sense of data quality and reliability in its waste site cleanup program.

MassDEP contracted with a well-known laboratory Proficiency Testing company to prepare test samples. To maintain the confidentiality of the study, the company set up mock consulting firms to send out samples and pay for analyses. Each laboratory was shipped a soil sample and groundwater sample spiked with measured concentrations of 5 common Volatile Organic Compounds (VOCs). This procedure was repeated on 3 different occasions -- in July, September, and November of 2004 -- at identical spiking concentrations.

In addition to these 19 commercial laboratories, double-blind samples were also delivered to the MassDEP state analytical laboratory (the Wall Experiment Station), by an agency employee, under the pretense of being samples from a confidential enforcement case.

MassDEP believes the results of this study are very encouraging. The vast majority of the laboratories evaluated were able to consistently quantify most analytes within 20% of the actual value. This excellent result is well within the most stringent acceptance criteria in use by the industry.

In a few cases, false positive or false negative results were reported, particularly with respect to vinyl chloride in water, which is known to be a problematic analyte. MassDEP is conducting further review of analytical data generated by the study to attempt to determine the reasons for these results.

Given these findings, MassDEP believes the public can have confidence in the integrity of the commercial laboratory community, and in the accuracy of the analytical data used to confirm cleanup of sites contaminated with Volatile Organic Compounds (VOCs), which are among the most pervasive and problematic pollutants at hazardous waste sites.

Table of Contents

Table of Contents (continued)

List of Figures

List of Tables

1.0 BACKGROUND

In Massachusetts, the cleanup of contaminated sites is regulated by the Massachusetts Department of Environmental Protection (MassDEP) under a privatized program initiated in 1993. In the last 12 years, over 20,000 sites have been assessed, remediated, and closed-out under this system, by privately funded Licensed Site Professionals that are obligated to follow the performance and cleanup standards specified in 310 CMR 40.0000, the Massachusetts Contingency Plan (MCP).

Since the inception of the program, concern has existed over the quality of analytical data used to support cleanup decisions. While MassDEP has a certification program for laboratories conducting drinking water and wastewater analyses, it does not at present evaluate or certify laboratories for the analysis of soil and groundwater samples from contaminated sites. Consequently, all assessment, cleanup, and closure decisions are based upon analytical test data from laboratories that are not specifically approved or monitored for this work. Moreover, the highly competitive nature of the analytical services industry in New England led some to suspect that poor - perhaps even fraudulent - performance was common.

To address these concerns, MassDEP initiated a comprehensive *Data Quality Enhancement Program* in the late 1990s. (http://www.mass.gov/dep/about/qaqcdocs.htm). With substantial input and contributions from the laboratory community, the agency generated a series of documents that provide additional detail and specification on the conduct of EPA SW-846 Test Methods, together with general sampling and analytical Quality Assurance/Quality Control requirements. While these efforts have provided guidance and additional clarity for laboratories, data users, and regulators on the production of high quality analytical data, a quantifiable and direct assessment of how well this system was working was needed.

As a result, MassDEP implemented a Double-Blind Laboratory Evaluation Program (Program) during the spring of 2004. This effort, detailed below, is believed to be one of the largest projects of its kind ever conducted in the United States.

2.0 OBJECTIVES

The primary objectives of this study were to improve and ensure confidence in the data relied upon by the waste site cleanup program by:

1. **Complementing and Extending the Data Quality Enhancement Program** – MassDEP has devoted considerable effort over the last five years to promote and ensure the production of reliable analytical data, producing numerous work products and policies. The overall and specific results and data from this effort will help MassDEP determine the scope and direction of future initiatives.

2. **Providing a Quantifiable Assessment of Data Quality** – While the Data Quality Enhancement Program has created the infrastructure and provided the tools for the

production and documentation of high quality data, a double-blind testing effort is the most direct way to determine if these tools and procedures are being used as intended, and producing the desired results.

3. **Providing for Market Deterrence and Correction** – By design, MassDEP's privatized cleanup program reacts to market-driven incentives. Conducting and publishing the results of this and future double-blind efforts will provide a market incentive for laboratories to maintain robust quality assurance programs, and provide a counter-balance away from competitive forces that focus only on providing the lowest-cost services.

3.0 SCOPE

The scope of this Program was to evaluate analytical services at 20 laboratories, including MassDEP's in-house laboratory, the Wall Experiment Station (WES). These laboratories were selected because they collectively analyze an estimated 75% of all samples related to assessment and remediation of sites in Massachusetts. The focus of the Program was on the analysis of Volatile Organic Compounds (VOCs), the most common contaminants of interest at sites across the state.

All laboratories were instructed to follow MassDEP's "MCP methods", which modify and clarify EPA's SW-846 Test Methods and replace analytical and quality control "recommendations" with "requirements", and provide detailed specification and performance standards on items that are otherwise left in SW-846 to the discretion of individual analysts. The MCP methods were developed by MassDEP in 2003 with significant input from the laboratory community, and are used at virtually all sites in Massachusetts at the present time. All laboratories that use these procedures are required to certify under pains and penalty of perjury that they have followed and have met all required procedures and standards, or, if they did not, to explicitly disclose and explain exceptions. For complete details see http://www.mass.gov/dep/about/qaqcdocs.htm.

4.0 DESIGN AND EXECUTION

The design and execution of the Program involved selecting a contractor, the 20 laboratories to be evaluated, the types and concentrations of contaminants, and other sample preparation requirements.

An important feature of this study was the decision to ship 3 rounds of samples over a 4-6 month period, *containing the same analytes at the same concentrations.* In this manner, each laboratory got "3 bites at the apple". While a poor performance during a single round could be attributable to a variety of factors and circumstances - including the possibility of problems with the sample itself - consistent data outliers over multiple rounds would tend to be indicative of more systemic and/or pervasive operational and/or equipment issues at the laboratory facility.

4.1 Selection of Contractor

In accordance with state requirements, MassDEP used a competitive bidding process to select the company that would assist in conducting the program.

After receiving signed confidentiality agreements from solicited bidders, a "Request for Response" was issued in March 2004. Because complete secrecy is essential for the success of any double-blind study, the bid package placed a premium on experience performing double-blind evaluations with "third party" billing and specifically required bidders to document their experiences with laboratory coordination and confidentiality issues for similar double-blind projects.

In May 2004, the contract was awarded to Environmental Resource Associates (ERA) of Arvada, Colorado, as the most qualified Proficiency Testing (PT) provider. It is noted that ERA is the only private provider accredited by both the National Institute of Standards and Technology (NIST) and the American Association for Lab Accreditation (A2LA). Under this contract, ERA was responsible for all aspects of sample preparation, laboratory coordination, sample shipping, payment (i.e., "third party" billing) and all other administrative activities associated with the project.

4.2 Selection of Laboratories

Based upon budget and project parameters, as well as an institutional knowledge of the analytical service providers in New England, a decision was made to include 20 laboratories in the Program.

The selection of laboratories for the study was based upon the volume of work they conduct on MCP-related work in Massachusetts. Under the guise of an information gathering exercise for educational and outreach purposes, MassDEP field staff were asked to list those laboratories that, in their experience, conduct most of the analytical testing at sites within their region of the state. This list was cross-checked against a systematic examination of site cleanup reports submitted to the agency to ensure that the labs with the highest volume of samples were included. The final list of laboratories selected for inclusion in the Double-Blind study is contained in Table 4-1. Collectively, it is estimated that these laboratories provide analytical support services at approximately 75% of all contaminated sites in Massachusetts.

4.3 Selection of Contaminants and Spiking Concentrations

Each of the three "sampling events" consisted of sending one whole-volume water sample and one whole-volume soil sample to each of the 20 laboratories. The objective was to spike common VOC contaminants in soil and/or groundwater at concentrations that should be readily identifiable and quantifiable (i.e., approximately 10 to 100 times the analyte's Reporting Limit). An effort was made to select and spike analytes in a manner that looked realistic in order to not raise suspicions among study participants. A summary of the analytes and spiking concentrations chosen for this study, along with MassDEP risk-based cleanup standards, is shown in Table 4-2.

Table 4-1
Laboratories Selected for MassDEP Double-Blind Laboratory Evaluation Program

LABORATORY	LOCATION
Accutest Laboratories	Marlborough, MA
Alpha Analytical	Westborough, MA
AMRO Environmental	Merrimack, NH
Chemserve	Milford, NH
Con-Test Analytical	East Longmeadow, MA
Eastern Analytical	Concord, NH
ESS Laboratory	Cranston, RI
GeoLabs	Braintree, MA
Groundwater Analytical	Buzzards Bay, MA
Katahdin Analytical	Westbrook, ME
Maxymillian Technologies	Lanesborough, MA
New England Chromachem	Salem, MA
New England Testing	Providence, RI
Phoenix Environmental	Manchester, CT
Premier Lab	Dayville, CT
Spectrum Analytical	Agawam, MA
STL Westfield	Westfield, MA
Toxikon Corp.	Bedford, MA
Wall Experiment Station (MassDEP)	Lawrence, MA
Woods Hole Analytical	Raynham, MA

Table 4-2
Double-Blind Sample Analytes and Spiking Concentrations

Water			Soil		
ANALYTE	**DESIGN TARGET µG/L OR PPB**	GW-1 STANDARD[1] µG/L OR PPB	ANALYTE	**DESIGN TARGET µG/G OR PPM**	S-1/GW-1 STANDARD[2] µG/G OR PPM
Benzene	**25**	5	Benzene	**20**	10
1,1,1-Trichloroethane (1,1,1-TCA)	**150**	200	1,1,1-Trichloroethane (1,1,1-TCA)	**40**	30
Trichloroethylene (TCE)	**35**	5	Trichloroethylene (TCE)	**15**	0.4
cis-1,2,-Dichloroethylene (*cis*-1,2,-DCE)	**100**	70	Tetrachloroethylene (PCE)	**10**	0.5
Vinyl Chloride (VC)	**20**	2	Methyl-t-Butyl Ether (MtBE)	**5**	0.3

[1] Applicable in drinking water resource areas
[2] Applicable in residential settings overlying drinking water resource areas

4.4 Preparation and Shipment of Samples

Whole-volume samples were prepared by ERA using analytically verified stock standard solutions and/or neat materials. All of the stocks used in the preparation of test samples were analyzed against at least two other independent sources to ensure the accuracy of spiking concentrations. Where available, a NIST Standard Reference Material (SRM) was used as one of these sources.

All volumetric glassware used in the preparation process was calibrated to "Class A" tolerances. All balances used were calibrated and traced to NIST weights. Notes for each sample were recorded by the chemist preparing the sample and reviewed by an independent chemist or manager.

Spiking solutions were prepared in methanol. Water samples were prepared by the volumetric addition and zero headspace mixing of the spiking solution into acidified reagent water. This results in residual concentrations of methanol in the water sample (approximately 50 mg/L), which is of potential concern with respect to maintaining the confidentiality of the study (i.e., significant methanol presence in a purported "real world" sample could trigger suspicion). However, since the specified VOC testing method (MCP Method 8260B) requires laboratories to spike samples with internal and surrogate standards that are contained in a methanol solution, it was considered unlikely that the original methanol contribution would be discernable.

The preparation of soil samples was more complicated. Each evaluated laboratory provided the mock consulting company (i.e., ERA) with a pre-weighed vial containing varying amounts of methanol (5, 10, 15 or 20 ml). ERA then calculated the amount of methanol spiking solution that would need to be added to the vial to create the required target sample concentrations of indicated contaminants. A gas-tight syringe was then used to withdraw methanol from the vial, in the exact volume of the calculated spiking solution addition. Subsequently, de-ionized water was added to the vial to approximate a 5% soil moisture content, followed by the addition of the appropriate mass of dry soil (i.e., 5, 10, 15, or 20 grams). Finally, the methanol-based spiking solution was added to the methanol/water/soil mixture in the sample vial using a gas-tight syringe. The vial was then capped and shaken thoroughly to mix the contents.

In order to allow the evaluated laboratories to determine moisture content and report soil data on a "dry weight" basis, as required by MCP Method 8260B, ERA also dispensed un-spiked soil into un-preserved vials (i.e., no methanol). This soil was prepared at a moisture content of 5% by weight.

Analytes and target concentrations were not changed between rounds. However, because ERA needed to prepare a new spiking solution after the first round, there were very slight differences in some of the final spiking concentrations between Round 1 and Rounds 2 & 3.

4.5 Implementation and Follow-Through

Three mock consulting firms were created to contract analytical services with the 19 participating private laboratories. A general "script" was prepared by MassDEP for use by these "firms"

concerning a desire to test a sample for an undisclosed site undergoing assessment. The most important directive given to ERA, however, was to maintain the appearance of a normal transaction, including, as appropriate, negotiating costs. All payments made to the private laboratories were from the mock consulting companies, as were all communications and correspondence.

Providing samples to the MassDEP Wall Experiment Station (WES) required a different approach. In this case, an agency employee received the shipment of samples from ERA, and personally transported them to WES under standard MassDEP Chain of Custody form and procedures. Laboratory personnel were informed that the samples originated from an undisclosed location that was the subject of a confidential agency investigation and enforcement action.

The three rounds of samples were prepared and shipped in July 2004, September 2004, and November 2004, respectively. ERA provided its final summary of results, original laboratory reports/invoices, and signed Analytical Report Certification Forms to MassDEP in April 2005.

5.0 RESULTS AND OBSERVATIONS

The results from this study are organized into two categories: (1) overall laboratory results, and (2) individual laboratory results. Overall results indicate laboratory performance as a whole along with general trends, observations and comparisons between compounds and matrices. Individual laboratory results are a "snapshot in time" of how a particular facility performed for these specific samples and analytical test methodologies, and may or may not be indicative of longer-term performance on other samples using similar methodologies. Individual results for each of the 20 laboratories are presented in alphabetical order in Appendix A.

5.1 Overall Laboratory Results

Overall laboratory results were compiled both numerically and graphically. The results were assessed numerically using "percent difference" expressed as the average of the absolute values of the percent differences between the "Assigned" (or "True") Values and each of the 60 results for each compound (i.e., 20 labs and 3 rounds yield 60 results per compound per water sample and 60 results per soil sample). These results are shown in Table 5-1.

In addition, results were compiled in graphical form to readily assess and compare the measured values of analytes for all three rounds. These results are shown in Figures 5-1 through 5-4.

The actual numerical results for water and soil are depicted in Figures 5-1 and 5-3, respectively. However, a more useful graphical presentation is to normalize all results to the same scale to allow even comparisons between different analytes and media. Therefore, the percent differences between the assigned values and each actual lab result were calculated and are depicted in Figures 5-2 and 5-4 for water and soil, respectively.

Table 5-1
Overall Laboratory Results

Water			Soil		
ANALYTE	ASSIGNED (TRUE) VALUE µG/L OR PPB	**AVERAGE PERCENT DIFFERENCE**	ANALYTE	ASSIGNED (TRUE) VALUE µG/G OR PPM	**AVERAGE PERCENT DIFFERENCE**
Benzene	24.5	**10.1**	Benzene	19.8	**24.1**
1,1,1-Trichloroethane (1,1,1-TCA)	149	**10.8**	1,1,1-Trichloroethane (1,1,1-TCA)	39.9	**25.7**
Trichloroethylene (TCE)	35.2	**8.7**	Trichloroethylene (TCE)	15.1	**24.4**
cis-1,2,-Dichloroethylene (*cis*-1,2,-DCE)	101	**15.8**	Tetrachloroethylene (PCE)	10.2	**24.1**
Vinyl Chloride (VC)	20.4	**33.7**	Methyl-t-Butyl Ether (MtBE)	4.96	**18.0**
Overall Average →		**15.8%**	Overall Average →		**23.3%**

5.2 Discussion of Overall Laboratory Results

Based upon these overall results and percent differences, the following findings were observed and considered by MassDEP:

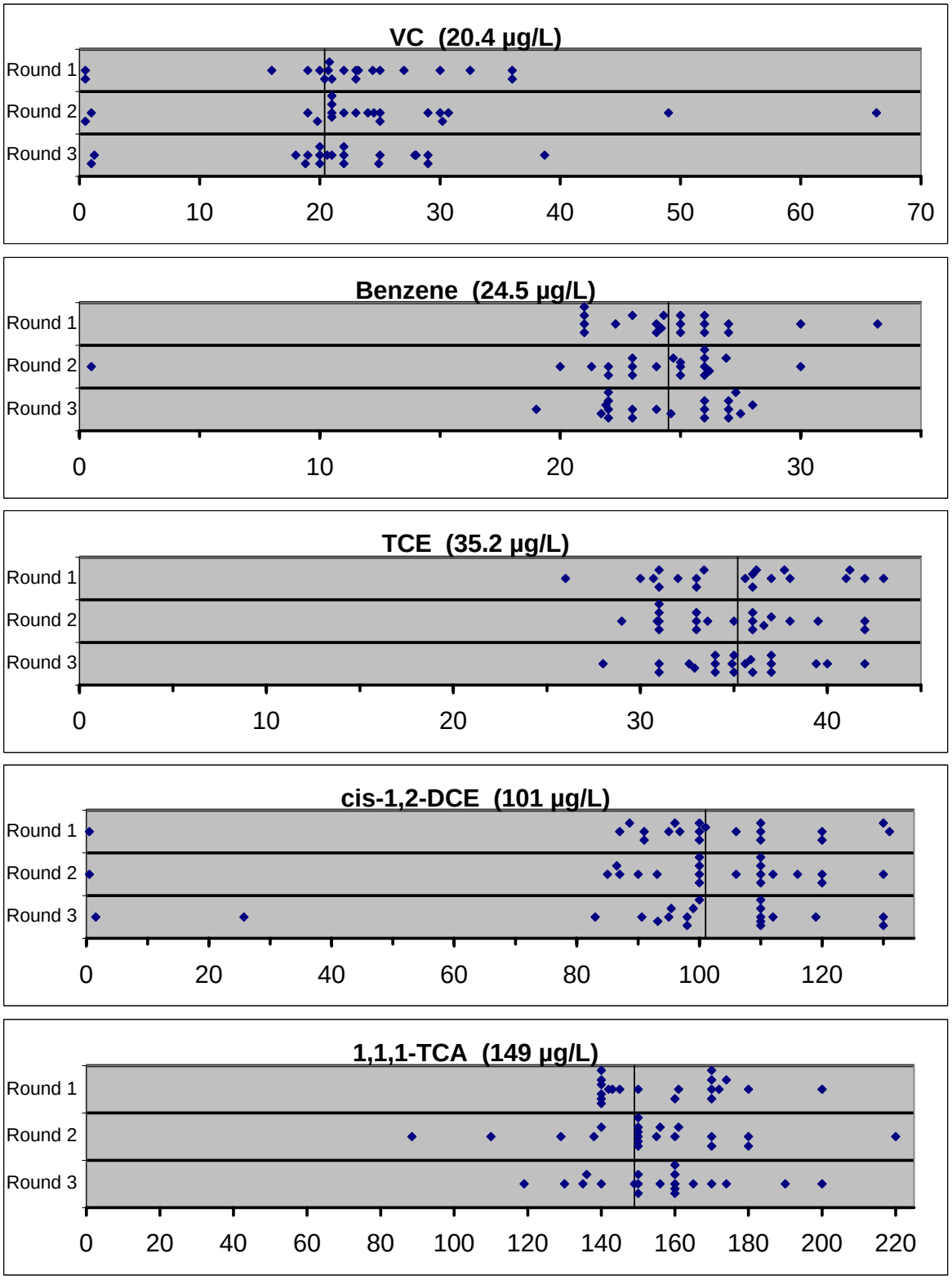
5.2.1 Variability Between Rounds - There was a noticeable increase in the variability of the results between the first two rounds, particularly for soil. As a whole, the overall laboratory performance was not as strong in Round 2 as it was in Round 1.

5.2.2 Accuracy of Water and Soil Matrices – The overall average differences for water and soil were 15.8% and 23.3%, respectively. With the exception of vinyl chloride (VC), all analytes in water were quantified with greater accuracy than any of the analytes in soil. Other than the VC results and with a few select but extreme exceptions, almost all of the individual laboratory results for water were within +/- 20% of the assigned (or "true") value. Fewer, though still a majority, of the soil results were within +/- 20% of the assigned (or "true") value.

5.2.3 Variability of Water and Soil Matrices - The overall variability (or "scatter") of the soil results for all three rounds is greater than that for water. This was expected, given the increased complexity in the preparation and analysis of soil VOC samples, which increases the possibility (and compounding) of error and positive or negative bias, and is consistent with ERA's historical database and other available industry information. Of methodological interest is the effect of such variables as:

- *Hold times (partitioning effects and/or methanol loss):* Hold time (the number of days between sample collection and sample analysis) is known to be an issue with "real world" soil VOC samples due to (a) an increased extraction of analytes from the soil into the

Figure 5-1
MassDEP Double-Blind Laboratory Evaluation Program
WATER SAMPLES
VOCs by MCP Method 8260/5030 (µg/L or ppb)



µg/L or ppb

Figure 5-2
MassDEP Double-Blind Laboratory Evaluation Program
WATER SAMPLES
Percent Difference from Assigned Values

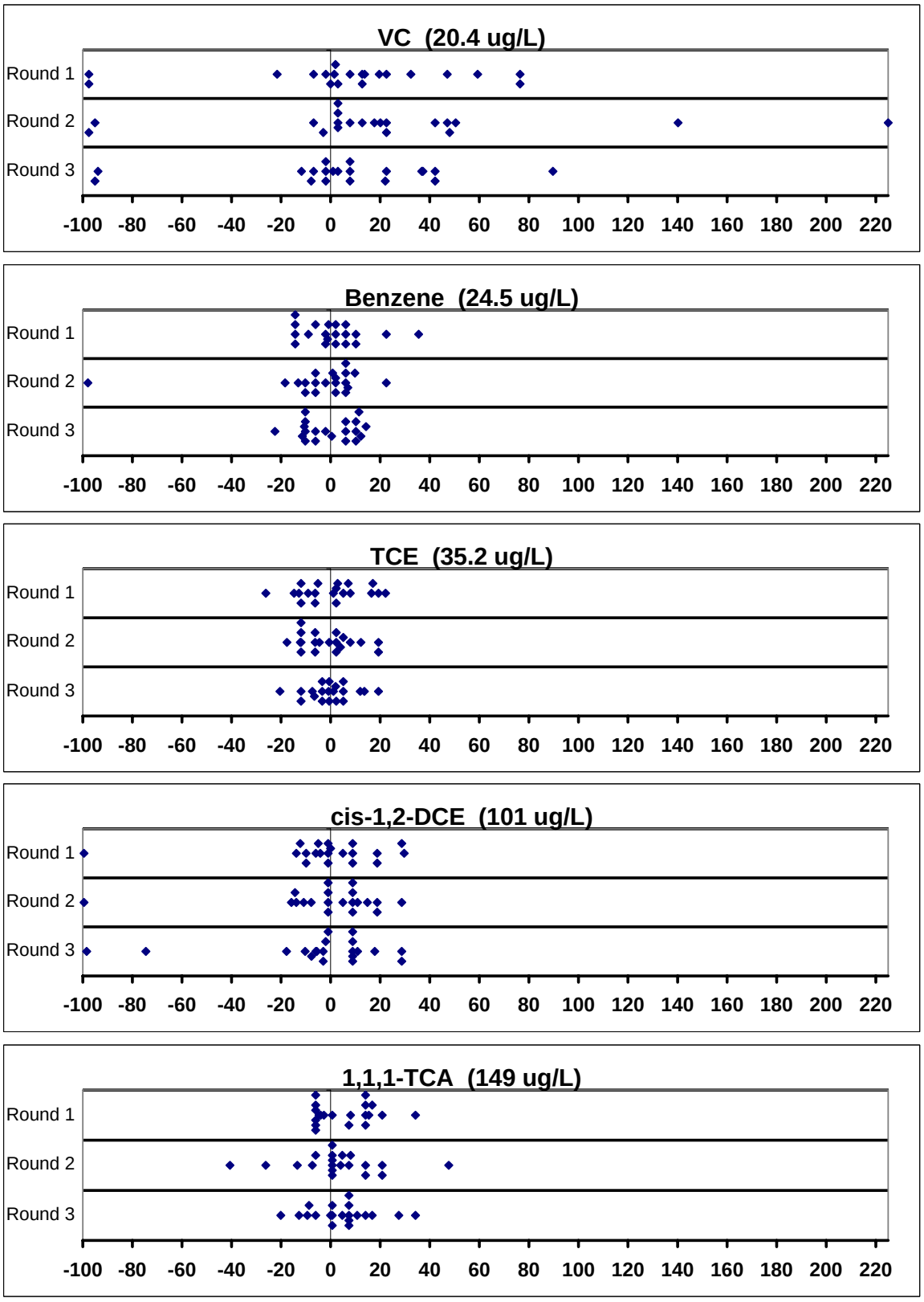


Figure 5-3
MassDEP Double-Blind Laboratory Evaluation Program
SOIL SAMPLES
VOCs by MCP Method 8260B/ 5035 (µg/g or ppm)

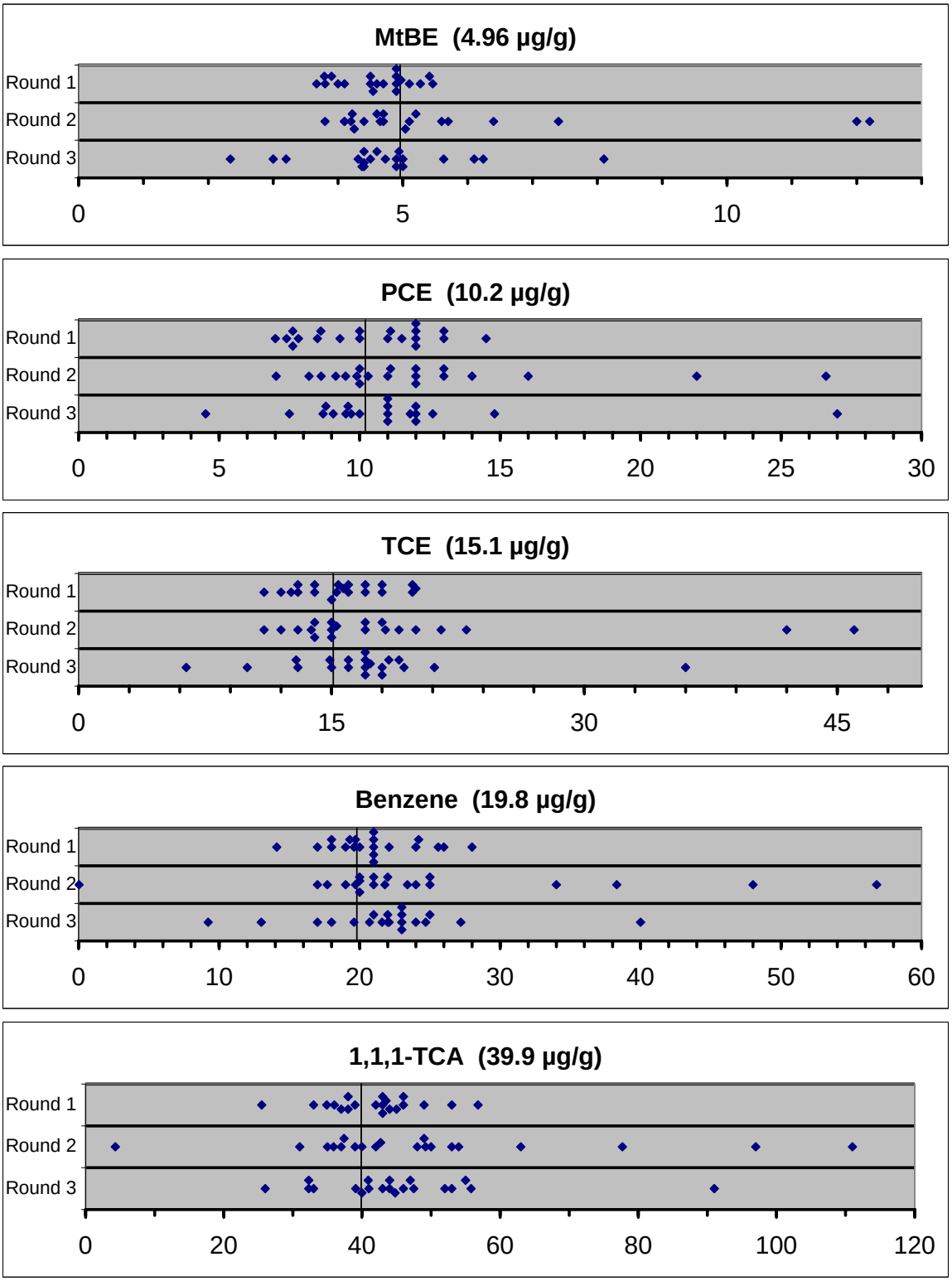
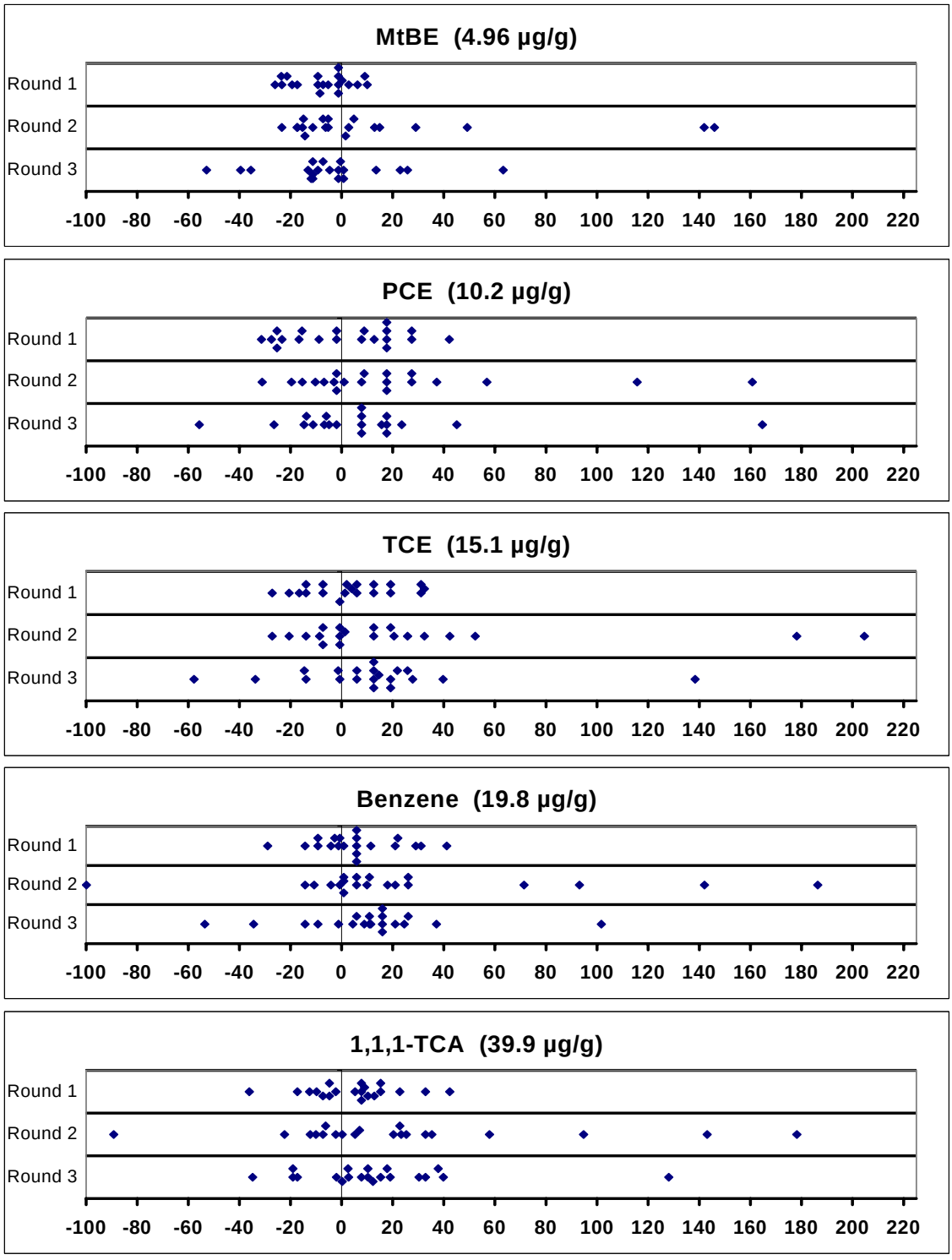
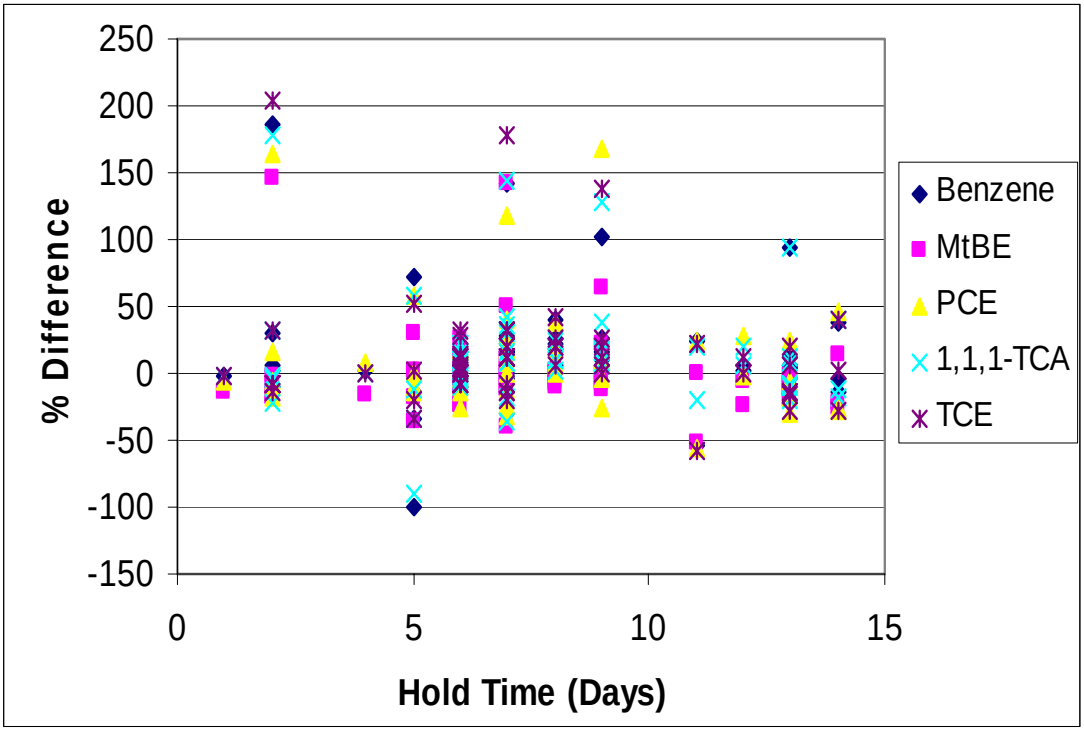


Figure 5-4
MassDEP Double-Blind Laboratory Evaluation Program
SOIL SAMPLES
Percent Difference from Assigned Value



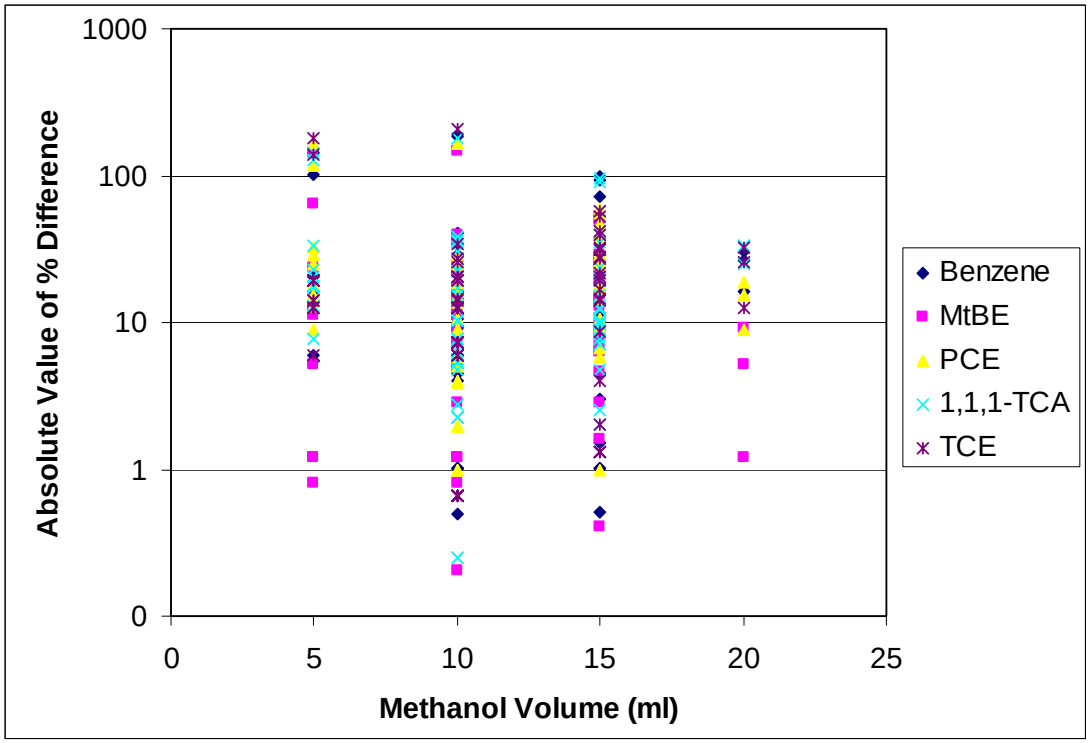
methanol preservative over time and/or (b) an increased possibility for methanol loss due to evaporation over time. Variability in extraction over time is unlikely to be an issue in this study since analytes were spiked directly into a methanol solution and would be unlikely to partition into the (originally uncontaminated) soil sample. Rather, for this study, loss of methanol from the sample due to evaporation could conceivably lead to a positive bias. However, when the difference for each individual soil result was plotted versus its individual hold time, no such trend is observed, and no direct correlation is evident (e.g., the maximum "R squared" number was less than 0.04). This relationship is shown graphically in Figure 5-5 below.

Figure 5-5
Percent Difference vs Hold Time (Soil)



- *Methanol and sample volumes:* Smaller methanol and soil sample volumes could lead to proportionally larger analytical variability in both positive and negative directions due to measurement errors. In addition, any loss of methanol from the sample due to evaporation could be magnified in smaller samples (i.e., loss of 1 ml of methanol from a 5 gram soil/5 ml methanol sample is more significant than loss of 1 ml from a 20 gram soil/20 ml methanol sample). However, when the absolute value of difference for each individual soil result was plotted versus methanol volume, no such trend is observed, and no direct correlation is evident (e.g., the maximum "R squared" number was less than 0.02). This relationship is shown graphically in Figure 5-6.

Figure 5-6
Percent Difference vs Methanol Volume (Soil)

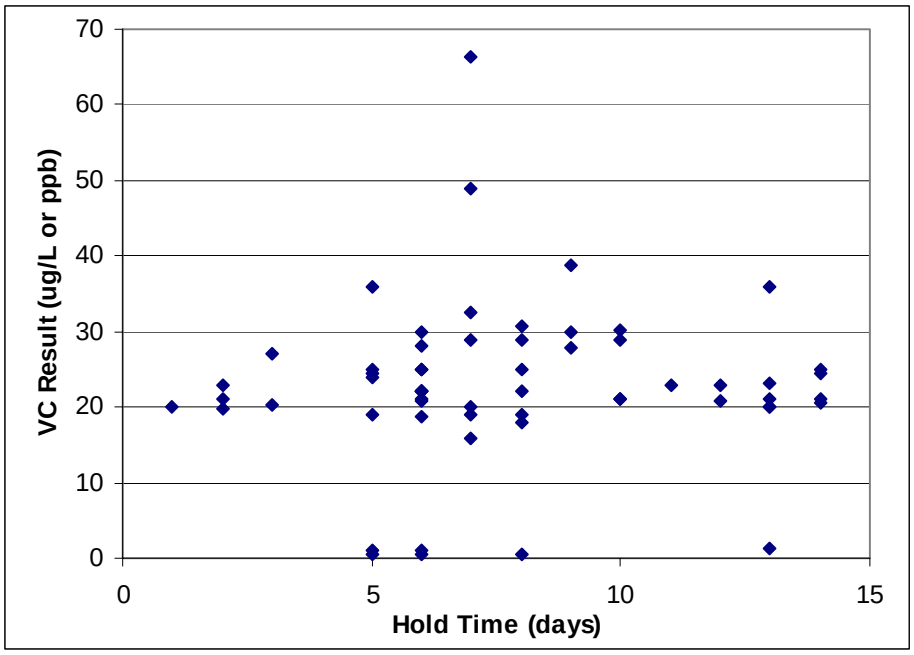


- *Equipment condition:* Finally, there are even more mundane explanations for the variability in soil results, including the possibility that some laboratories may analyze soil VOC samples on dedicated GC systems (separate from the instrumentation used for water samples), which tend to see more "dirty" samples, and thus may experience more systemic "carry over" and chromatographic "noise".

***5.2.4 Vinyl Chloride* (VC)** - The average difference for VC in water was 33.7%, greater than any difference for any analyte in either water or soil. Moreover, there were six false negative (non-detect) results for VC in water. Except for these six false negatives, most of the laboratories over-quantified their VC result. Given the number of laboratories that were able to produce accurate results over all three rounds, however, it is clearly possible to adequately identify and quantify this compound, at the spiking concentration used in this study.

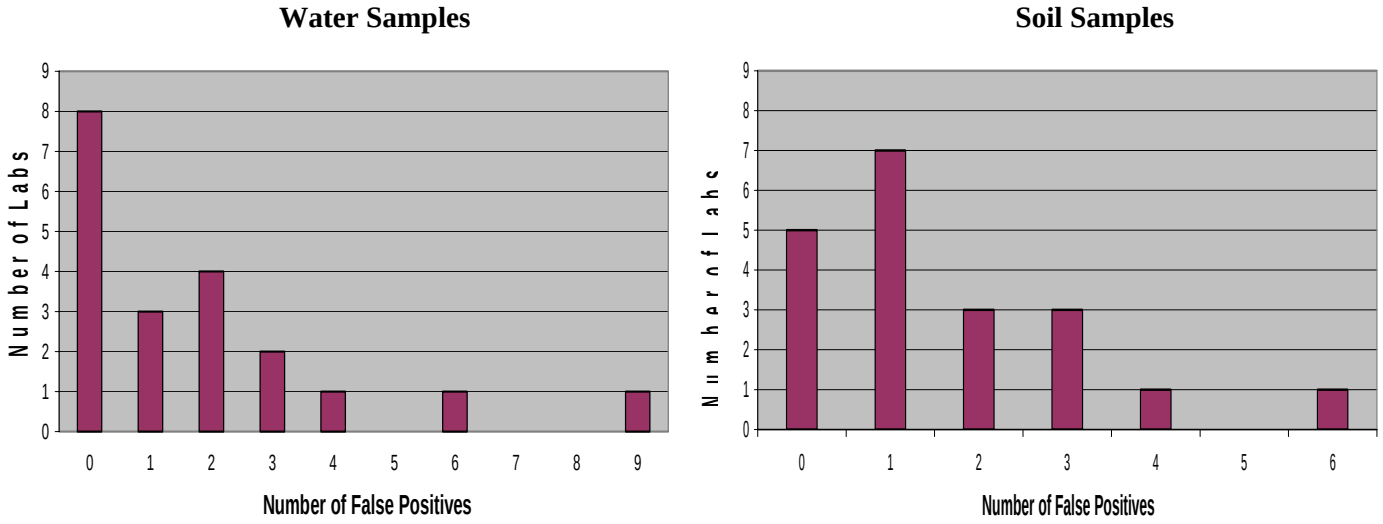
Because of its high volatility, some have speculated that VC results will decline over time from its true concentration (i.e., the faster the sample is analyzed, the higher the VC result will be). However, when the difference for each individual VC result in water was plotted versus its individual hold time, no such trend is observed, and no direct correlation is evident (e.g., "R squared" was less than 0.02). This relationship is shown graphically in Figure 5-7.

Figure 5-7
Vinyl Chloride Result vs Hold Time (Water)
Assigned Value = 20.4 µg/L



5.2.5 False Positives – 12 laboratories reported one or more false positives in the water samples and 15 laboratories reported one or more false positives in the soil samples, as detailed in Appendix A, and summarized in Figure in Figure 5-8.

Figure 5-8
False Positive Detections in Water and Soil Samples (All Three Rounds)



Over all sample rounds, 11 analytes were falsely reported to be present in samples at concentrations above 1 µg/L in water and/or 1 µg/g in soil, as detailed in Table 5-2. Another 13 analytes were reported present at concentrations less than 1 µg/L in water and/or 1 µg/g in soil. Chloromethane was the most frequently reported false positive, with 14 detections (water samples only) at concentrations ranging to 20.7 µg/L

Table 5-2
False Positive Detections Above 1 µg/L Water and/or 1 µg/g Soil (All Three Rounds)

Analyte	Water Sample Detects		Soil Sample Detects	
	# Detects	Conc Range µg/L	# Detects	Conc Range µg/g
1,1, -Dichloroethene	5	3J[a] - 11	2	2 - 4.31
1,2 -Dichloroethane			1	2.32
1,4 -Dioxane			1	9.5
Bromoform	2	2.1-26		
Carbon Tetrachloride			1	7.2
Chloroethane	1	12.4		
Chloromethane	14	2J[a] – 20.7		
Dichlorodifluoromethane	1	9.23		
Methylene Chloride			2	1.4J[a] – 4.5
Naphthalene			1	5.5
Trans-1, 2-Dichloroethylene	1	2J		

[a] "J" signifies detection below Reporting Limit

Most false positive detections were at low concentrations, and would be unlikely to significantly impact site assessment or remedial decisions. However, for most compounds, this would appear to be a preventable problem, with good laboratory practice (e.g., storing solvents and standards in locations separate from sample storage and analytical equipment areas). *The reporting of chloromethane in the aqueous samples may be an exception, in that this compound may actually be present in some samples, as a reaction/breakdown product of the acid preservative (HCl) and/or as a thermal degradation product of other analytes.* This possibility will be further pursued by MassDEP as a methodological issue of interest.

It is unclear how pervasive cross-contamination problems are within the industry. This double-blind study is perhaps instructive in this regard, given an assumption that most laboratories would likely employ especially robust "housekeeping" procedures during a single-blind study effort, knowing that they are being tested in this regard.

5.2.6 Mis-identified Compounds - Other than for VC, most of the false negatives appeared to be misidentified compounds (e.g., one lab reported *trans*-1,2-DCE at almost the exact assigned value for *cis*-1,2-DCE, and another lab reported bromoform instead of benzene).

5.2.7 Prices - The price per sample charged by the laboratories during this Program varied from $80 to $185, including surcharges for "MCP deliverables" billed by some of the labs. However, it is important to note that prices quoted by any one lab can vary depending on the

client, the market, their marketing strategy, their volume of work and/or a number of other factors.

5.3 Individual Laboratory Results

Individual laboratory results are listed by laboratory in alphabetical order in Appendix A.

6.0 ACCEPTANCE LIMITS

Given the results obtained from this study, how does one go about evaluating how well laboratories performed? If 0% is the ideal percent difference, is a result of 15.8% average percent difference for water samples and 23.3% average percent difference for soil samples a good result?

Because there is no single method used by the laboratory community to rate the results of a study like this one, MassDEP compared the results of this study against 3 different types of acceptance limits, described more fully below. Most results of the laboratories evaluated in this study met all three sets of limits, which is indeed a good result.

In determining acceptance criteria, two points must be noted. First, there are no mandated or universally applied standards to evaluate the quality of analytical testing data of this nature. Second, the difficulty of the overall task being evaluated must be considered (i.e., laboratories are attempting to identify and quantify very small amounts of chemicals, in the low parts per billion range. For perspective, note that one part per billion is equivalent to 1 inch in 16,000 miles.)

Environmental analytical data are commonly assessed using various criteria or "acceptance limits" based on a variety of state, federal, industry and/or individual laboratory standards or guidelines. These criteria can be absolute (i.e., acceptance is benchmarked to a specified value or percent difference) or relative (i.e., acceptance is based upon a statistical evaluation of multiple analyses and/or multiple laboratories). For general informational purposes, three of the most relevant approaches to the designation of acceptance limits are described below.

1) MassDEP Compendium of Analytical Methods (CAM):

The CAM is a detailed set of analytical procedures, based upon EPA SW-846 Test Methods (http://www.mass.gov/dep/about/qaqcdocs.htm). A product of the *MassDEP Data Quality Enhancement Program*, these procedures remove all ambiguities in the EPA test methods, and provide clear QA/QC performance standards and metrics.

For the test method used by laboratories during this study (i.e., MCP Method 8260B), the CAM requires the use of Laboratory Control Spikes and Surrogate Spikes to assess method accuracy. In both cases, the percent recoveries of spiked analytes must be +/- 30% of their true value. On the basis of this absolute, *intra*-laboratory assessment metric, quantification of PT sample analytes within +/- 30% of their true value, on an *inter*-laboratory basis, would be considered quite good.

2) Proficiency Test (PT) Sample Provider (ERA)

Companies that specialize in the production, distribution, and evaluation of Proficiency Test (PT) samples often establish "in-house" performance metrics. The PT provider used by MassDEP in this Program, Environmental Resource Associates (ERA), has developed and maintains its own library of methodological acceptance limits, based on years of experience in conducting studies, performance data from independent refereed laboratory studies, and data from USEPA Water Pollution (WP), Water Supply (WS), and Contract Laboratory Program (CLP) studies. ERA uses different databases for water and for soil. These limits represent approximate 95% confidence intervals based on large independent data sets (i.e., they were not calculated from the data obtained in this Program or any other *individual study*.)

3) USEPA and Industry Standard

Numerous organizations, including the USEPA and the National Environmental Laboratory Accreditation Conference (NELAC) commonly define acceptance limits on a *relative basis*, by judging a laboratory's performance in comparison to the performance of a collective group of laboratories in a *common study*. Typically, an individual laboratory's performance for any given PT sample analyte is considered acceptable if it is within +/- two standard deviations from the mean value reported by the collective group of laboratories. For a normal distribution, this approximately equals the 95% confidence interval (i.e., 95% of the results fall within this range).

The ranges for these three different types of acceptance limits for this Program are summarized in Tables 6-1 and 6-2 below. For illustrative purposes only, the ranges shown for the "Industry Standard" limits are based on averages of all three rounds, rather than on separately calculated ranges for each of the three rounds. The important points, however, are that (a) the acceptance ranges from both ERA and "Industry Standard" criteria, and presumably from any other ***inter***-laboratory criteria, are in some cases wider (or less conservative) than the range of +/- 30% from MassDEP CAM; and (b) most results from this Program fall within all three of these acceptance limits.

Table 6-1
"Acceptance Limit" Ranges for Water (µg/L or ppb)

Analyte		Benzene	1,1,1-TCA	TCE	*cis*-1,2-DCE	VC
True Concentration (µg/L)		24.5	149	35.2	101	20.4
Acceptable	MassDEP CAM (+/- 30%)	17.2 - 31.9	104 – 194	24.6 – 45.8	70.7 – 131.3	14.3 – 26.5
	ERA	18.7 – 29.6	106 – 181	25.2 – 42.9	76.5 – 125	10.9 – 31.9
	EPA/Industry Standard	19.2 – 29.9	113 – 198	27.3 – 42.7	63.7 – 139.7	10.2 – 40.5

Table 6-2
"Acceptance Limit" Ranges for Soil (µg/g or ppm)

Analyte		Benzene	1,1,1-TCA	PCE	TCE	MtBE
True Conc (µg/g)		19.8	39.9	10.2	15.1	4.96
Acceptable	MassDEP CAM (+/- 30%)	13.4 – 31.9	27.9 – 51.9	7.1 – 13.3	10.6 – 19.6	3.5 – 6.6
	ERA	15.5 – 24.7	29.9 – 50.9	7.20 – 13.9	11.5 – 19.0	3.34 – 6.70
	EPA/Industry Standard	9.7 – 36.3	16.5 – 74.8	3.8 – 18.8	5.6 – 28.9	2.2 – 7.8

A variety of other and more detailed numerical analyses may be performed on the results of this Program using the complete data summary, Individual Laboratory Results, shown in Appendix A

7.0 CONCLUSIONS

On the basis of the information and data presented and discussed above, the following conclusions are offered:

- Overall laboratory performance was generally very good;
- Most results from this program fall well within all three sets of acceptance limits against which they were compared;
- MtBE recovery and quantification in soil samples was very good, with the overall lowest average percent difference (18%) of any soil analyte. This is significant, given problems that can exist with purging and with trap desorption efficiencies;
- Vinyl Chloride is a particularly problematic analyte in water samples for some laboratories, even at concentrations well above analytical Reporting Limits. This issue will be further considered and addressed as part of the agency's *Data Quality Enhancement Program*;
- False positive detections were seen to be a relatively minor - though seemingly preventable - problem. Specifically, most laboratories reported at least one analyte that was not added to the PT sample by ERA (i.e., "false positive"), presumably because of sample and/or laboratory cross-contamination issues. In all sample rounds (i.e., 60 water samples and 60 soil samples), 11 analytes were falsely reported in aqueous samples at concentrations above 1 µg/L and/or soil samples above 1 µg/g; another 13 analytes were falsely reported at concentrations less than 1 µg/L and/or 1 µg/g. *The reporting of chloromethane in the aqueous samples may be an exception, however, in that this compound may actually be present in some samples, as a reaction/breakdown product of the acid preservative (HCl) and/or as a thermal degradation product of other analytes.*

This possibility will be further pursued by MassDEP as a methodological issue of interest; and

- This was a useful exercise, affirming the feasibility and utility of a state agency to engage in what many agree is the next horizon in analytical quality assurance/quality control: large-scale, systematic double-blind testing programs. In addition to providing useful information on specific technical and operational matters, such studies also serve as a powerful market incentive to promote better work, to the extent they become a regular and expected occurrence. MassDEP intends to leverage these forces and conduct similar programs in the future.

REFERENCES

ERA; *Environmental Proficiency Testing and Quality Control Standards*; Environmental Resource Associates; Arvada, Colorado; 2005.

MassDEP; *Massachusetts Contingency Plan*, 310 CMR 40.0000, June 2003 and as amended. (http://www.mass.gov/dep/cleanup/laws/310cmr40.pdf)

MassDEP**;** *Compendium of Quality Assurance and Quality Control Requirements and Performance Standards for Selected Analytical Methods*, 2003, (http://www.mass.gov/dep/about/qaqcdocs.htmT).

USEPA; *2003 NELAC Standard*; EPA/600/R-04/003; National Environmental Laboratory Accreditation Conference; June, 2003.

USEPA; *National Standards for Water Proficiency Testing Studies*; NERL-Ci-0045; December 30, 1998.

Appendix A

Individual Laboratory Results

Table A-1: Individual Laboratory Results
Round #1 WATER (µg/L or ppb)

	Analytes spiked					*Analytes reported as false positives*					
	Benzene	**cis-1,2-Dichloroethylene**	**1,1,1-Trichloroethane**	**Trichloroethylene**	**Vinyl chloride**	*Chloroethane*	*Chloromethane*	*Dichlorodifluoromethane*	*trans-1,2-Dichloroethylene*	*1,1-Dichloroethene*	*Chloroform*
Assigned Values	**24.5**	**101**	**149**	**35.2**	**20.4**						
Participating Labs	**Units = µg/L**										
Accutest Laboratories	25.0	101	172	37.7	23.2					3.7	
Alpha Analytical	21	96	150	31	36						
AMRO Environmental	25	110	160	37	27		3.8				
Chemserve	23	100	140	26	19						
Con-Test Analytical	24.2	96.8	143	36.2	20.8						
Eastern Analytical	27	120	170	36	30						
ESS Laboratory	21.0	88.6	145	30.7	20.4						
GeoLabs	33.2	87.0	142	41.2	20.7		20.7	9.23			
Groundwater Analytical	24	100	170	33	23						
Katahdin Analytical	26	110	140	32	20				2J	3J	0.5J
Maxymillian Technologies	24.3	ND (1.0)	140	35.6	24.4	12.4			103		
New England Chromachem	30	131	174	43	ND (1.0)						
New England Testing	26	110	170	38	ND (1.0)						
Phoenix Environmental	24	95	140	33	23						
Premier Lab	21	91	140	31	21						
Spectrum Analytical	22.3	106	161	33.4	32.5		4.4				
STL Westfield	21	91	140	30	16						
Toxikon Corp.	27	130	200	42	36						
Wall Experiment Station[1]	25	120	170	36	22						
Woods Hole Analytical	26	100	180	41	25						
Study Mean	24.8	104	157	35.2	24.4						
Standard Deviation	3.1	13	17.7	4.5	5.7						

[1] All Wall Experimental Station results flagged with 'B' by laboratory. B = Analyte detected in LB, LRB, and/or trip blank or no trip blank was collected.

Table A-2: Individual Laboratory Results
Round #1 SOIL (µg/g or ppm)

	Analytes spiked						Analytes reported as false positives							
	% Solids	Benzene	MTBE	Tetrachloroethylene	1,1,1-Trichloroethane	Trichloroethylene	*1,4-Dioxane*	*1,1-Dichloroethene*	*Bromomethane*	*tert-Butlybenzene*	*sec-Butlybenzene*	*n-Butlybenzene*	*Naphthalene*	*Methylene chloride*
Assigned Values	**95.0%**	**19.9**	**4.96**	**10.4**	**39.9**	**15.1**								
Participating Labs	**%**	**Units = mg/Kg dry wt.**												
Accutest Laboratories	91.1	14.1	4.97	7.82	25.5	12.0		0.320						
Alpha Analytical	95	19	4	8.5	36	13								
AMRO Environmental	95.0	26.0	4.90	12.0	53.0	20.0		0.920						
Chemserve	94.2	21	4.9	7	43	13	9.5							
Con-Test Analytical	94.8	22.1	4.54	8.62	43.0	12.6		4.31						
Eastern Analytical	95	24	4.9	12	49	16								
ESS Laboratory	96	19.7	3.79	11.1	43.5	15.4								
GeoLabs	95	24.2	5.41	11.5	46	19.8								
Groundwater Analytical	94	21	4.9	13	37	16								
Katahdin Analytical[1]	94.3	17.0	3.9J	7.4	33.0	11.0		0.490J						1.40JB
Maxymillian Technologies	94.1	19.3	3.67	7.62	34.9	15.3			0.163	0.059	0.059	0.084	0.658	
New England Chromachem	81	19.6	5.27	7.62	38.0	15.7								
New England Testing[1]	95.13	28.0	4.50	11.0	44.0	18.0								
Phoenix Environmental	94	20.0	5.10	12.0	46.0	17.0							5.50	
Premier Lab	94.8	18.0	3.80	10.0	39.0	14.0								
Spectrum Analytical	94.9	25.6	5.46	14.5	56.8	19.8								
STL Westfield	94.7	21.0	4.50	12.0	43.0	17.0								
Toxikon Corp.	94.2	18.0	4.10	9.30	38.0	14.0								
Wall Experiment Station[2]	94	21	4.6	10	42	15								
Woods Hole Analytical	94.5	21	4.7	13	45	18								
Study Mean	93.8	21.0	4.63	10.3	41.8	15.6								
Standard Deviation	3.2	3.3	0.55	2.2	7.0	2.6								

[1] Katahdin Analytical and New England Testing data for 1,1,1-TCA, Benzene, TCE, and PCE taken from diluted runs, 'E' flagged on original run.

[2] All Wall Experimental Station results flagged 'B' by laboratory. B = Analyte detected in LB, LRB, and/or trip blank or no trip blank was collected.

Table A-3: Individual Laboratory Results
Round #2 WATER (µg/L or ppb)

	Analytes spiked					Analytes reported as false positives						
	Benzene	**cis-1,2-Dichloroethylene**	**1,1,1-Trichloroethane**	**Trichloroethylene**	**Vinyl chloride**	*Chloroethane*	*Chloromethane*	*Dichlorodifluoromethane*	*trans-1,2-Dichloroethylene*	*1,1-Dichloroethene*	*Bromoform*	*Chloroform*
Assigned Values	**24.5**	**101**	**149**	**35.2**	**20.4**							
Participating Labs	**Units = µg/L**											
Accutest Laboratories	21.3	93.1	138	30.9	21.0					6.3		
Alpha Analytical	24	100	160	36	21							
AMRO Environmental	20	87	150	29	22							
Chemserve	26	110	180	33	30		4					
Con-Test Analytical	26.2	116	88.5	33.0	30.2							
Eastern Analytical	30	130	220	42	49							
ESS Laboratory	26.9	106	156	39.5	24.5							
GeoLabs	23.0	86.5	129	33.6	19.8							
Groundwater Analytical	22	100	140	31	21					11		
Katahdin Analytical	26	120	180	42	29				0.4J	3J		0.6J
Maxymillian Technologies	25.0	ND(1.0)	150	36.0	30.7		2.04		106			
New England Chromachem	ND (1)	112	161	37	25						26	
New England Testing	25	110	170	31	ND(1.0)							
Phoenix Environmental	23	100	150	31	21							
Premier Lab	26	100	150	35	24							
Spectrum Analytical	24.7	110	155	36.6	66.3							
STL Westfield	22	85	110	31	19		1.0					0.51
Toxikon Corp.	26	110	170	38	25							
Wall Experiment Station	25	120	150	36	23							
Woods Hole Analytical	23	90	150	33	ND(2.0)		3.1		2.2		2.1	
Study Mean	24.5	105	153	34.7	27.9							
Standard Deviation	2.3	13	26.8	3.7	11.8							

Table A-4: Individual Laboratory Results
Round #2 SOIL (µg/g or ppm)

	Analytes spiked						Analytes reported as false positives								
	% Solids	Benzene	MTBE	Tetrachloroethylene	1,1,1-Trichloroethane	Trichloroethylene	*Acrylonitrile*	*1,3-Dichloropropane*	*1,2,3-Trichloropropane*	*Carbon tetrachloride*	*2-Butanone*	*Carbon disulfide*	*Bromoform*	*Bromomethane*	*Naphthalene*
Assigned Values	**95.0%**	**19.8**	**4.96**	**10.1**	**39.9**	**15.1**									
Participating Labs	**%**	**Units = mg/Kg dry wt.**													
Accutest Laboratories	95.7	17.7	5.7	10.3	37.4	15									
Alpha Analytical	95	25	7.4	13	54	20									
AMRO Environmental	95.0	34	6.4	16	63	23									
Chemserve	94.3	21	4.6	8.2	39	14									
Con-Test Analytical	94.8	38.3	4.25	7.03	77.7	11.0									
Eastern Analytical	95	24	5.6	12	53	18									
ESS Laboratory	95	19.7	4.22	8.63	35.9	13.8									
GeoLabs	94	56.8	12.2	26.6	111	46.0					1.01				
Groundwater Analytical	95	19	4.4	10	37	14									
Katahdin Analytical	94.7	48	12	22	97	42									
Maxymillian Technologies	95	21.8	4.65	11.1	42.7	18.2								0.140	
New England Chromachem	92	ND(0.050)	5.04	9.15	4.31	15.3							18.5		
New England Testing	94.5	17	4	10	35	12									
Phoenix Environmental	95	22	5	12	49	17									
Premier Lab	95.2	20	4.2	11	40	15				7.2					
Spectrum Analytical	94.2	23.4	5.1	14	49.2	21.5	0.19	0.0678	0.183						0.132
STL Westfield	94.6	20	4.1	9.5	31	13									
Toxikon Corp.	94.4	25	4.7	12	50	19									
Wall Experiment Station	94	20	4.7M	9.9M	42	15M									
Woods Hole Analytical	94.8	21	3.8	13	48	17						0.65			
Study Mean	94.6	26.0	5.67	12.4	49.8	19.2									
Standard Deviation	0.7	10.8	2.43	4.8	23.4	9.3									

Table A-5: Individual Laboratory Results
Round #3 WATER (µg/L or ppb)

	Analytes spiked					Analytes reported as false positives						
	Benzene	cis-1,2-Dichloroethylene	1,1,1-Trichloroethane	Trichloroethylene	Vinyl chloride	Chloroethane	Chloromethane	Dichlorodifluoromethane	trans-1,2-Dichloroethylene	1,1-Dichloroethene	Bromoform	Chloroform
Assigned Values	24.5	101	149	35.2	20.4							
Participating Labs	Units = µg/L											
Accutest Laboratories	27.5	112	174	35.6	38.7							
Alpha Analytical	19	83	130	28	20							
AMRO Environmental	23	98	160	34	21		3.6					
Chemserve	26	110	150	35	20		3					
Con-Test Analytical	27.3	90.6	135	39.4	27.9		3.9					
Eastern Analytical	26	110	160	35	29							
ESS Laboratory	23.0	93.2	149	32.9	20							
GeoLabs	21.7	25.7	119	35.9	18.8		6.40					
Groundwater Analytical	22	100	170	31	< 2.5							
Katahdin Analytical	27	130	190	37	29		2J		0.5J			0.8J
Maxymillian Technologies	24.6	1.54	156	34.9	24.9		4.11		101			
New England Chromachem	28	119	165	40	22							
New England Testing	22	95	160	34	19							
Phoenix Environmental	27	110	160	37	28							
Premier Lab	26	110	150	36	22							
Spectrum Analytical	21.9	95.4	136	32.6	20.6		2.3		1			
STL Westfield	22	99	140	31	18							
Toxikon Corp.	27	130	200	42	25							
Wall Experiment Station[1]	22	110	150	34	22							
Woods Hole Analytical	24	98	160	37	<2		3.0					
Study Mean	24.4	96	156	35.1	23.7							
Standard Deviation	2.6	31	19.3	3.3	5.2							

[1] All Wall Experimental Station results flagged with 'B' by laboratory. B = Analyte detected in LB, LRB, and/or trip blank or no trip blank was collected.

Table A-6: Individual Laboratory Results
Round #3 SOIL (µg/g or ppm)

	Analytes spiked						Analytes reported as false positives				
	% Solids	Benzene	MTBE	Tetrachloroethylene	1,1,1-Trichloroethane	Trichloroethylene	tert amyl ether	1,1-Dichloroethene	1,2-Dichoroethane	m,p-Xylene	Methylene chloride
Assigned Values	**95.0%**	**19.8**	**4.96**	**10.1**	**39.9**	**15.1**					
Participating Labs	**%**	**Units = mg/Kg dry wt.**									
Accutest Laboratories	95.6	20.7	4.37	9.59	32.3	12.9		0.517	2.32		
Alpha Analytical	95	24	4.6	10	44	18					
AMRO Environmental	95.4	23	4.5	11	53	17		0.63		0.034	
Chemserve	95.1	17	4.4	8.8	33	13					
Con-Test Analytical	95.3	9.22	2.34	4.52	32.3	6.38		0.295			
Eastern Analytical	96	23.0	5.00	11.0	47.0	17.0					
ESS Laboratory	94	19.6	4.31	9.51	39.1	14.9		0.0568			
GeoLabs	96	22.1	4.73	11.8	40.9	19.3					
Groundwater Analytical	95	25.0	6.1	11.0	55.0	19.0					
Katahdin Analytical[1]	95.1	22.0	4.4	7.5	46.0	16.0					0.24J
Maxymillian Technologies	95.4	27.2	5.63	14.8	55.8	21.1					
New England Chromachem	95	21.6	6.24	9.06	44.8	17.3					
New England Testing[1]	100	23.0	4.9	12.0	52.0	18.0					
Phoenix Environmental	96	23.0	3.0	12.0	43.0	17.0		2.00			
Premier Lab	95.6	22.0	4.9	12.0	44.0	17.0					
Spectrum Analytical	95.3	24.7	4.94	12.6	47.5	18.4					
STL Westfield	95.3	21.0	5.0	11.0	41.0	16.0		0.83			
Toxikon Corp.	95.2	13.0	3.2	8.70	26.0	10.0	0.380				
Wall Experiment Station[2]	95	18	4.4	9.7	40	15					
Woods Hole Analytical	95.3	40	8.1	27	91	36					4.5B
Study Mean	95.5	22.0	4.75	11.2	45.4	17.0					
Standard Deviation	1.1	5.9	1.22	4.3	13.3	5.6					

[1] Katahdin Analytical and New England Testing data for 1,1,1-TCA, Benzene, TCE, and PCE taken from diluted runs, 'E' flagged on original run.

[2] All Wall Experimental Station results flagged 'B' by laboratory. B = Analyte detected in LB, LRB, and/or trip blank or no trip blank was collected. MTBE, trichloroethene, and tetrachloroethene results flagged 'M' by laboratory indicating analyte concentration between the MDL and RDL.