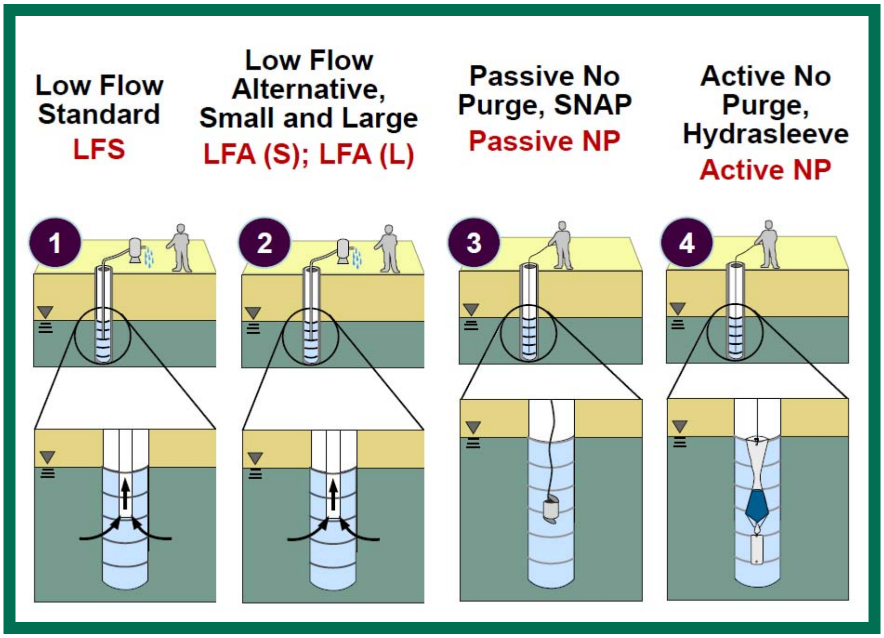
ESTCP
Cost and Performance Report

(ER-201209)



Methods for Minimization and Management of Variability in Long-Term Groundwater Monitoring Results

June 2015

ENVIRONMENTAL SECURITY
TECHNOLOGY CERTIFICATION PROGRAM

U.S. Department of Defense

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
25/06/2015

2. REPORT TYPE
Cost and Performance Report

3. DATES COVERED *(From - To)*
Sept 2012 - March 2016

4. TITLE AND SUBTITLE
Methods for Minimization and Management of Variability in Long-Term Groundwater Monitoring Results

5a. CONTRACT NUMBER
W912HQ-12-C-0055

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
McHugh, T.E., Kulkarni, P.R., Newell, C.J., and S.L. Britt

5d. PROJECT NUMBER
ER-201209

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
GSI Environmental, Inc. 2211 Norfolk, Suite 1000, Houston, TX 77098
ProHydro, Inc. 1011 Fairport Road, Fairport, NY 14450

8. PERFORMING ORGANIZATION REPORT NUMBER
ER-201209

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Environmental Security Technology Certification Program
4800 Mark Center Drive, Suite 17D08
Alexandria, VA 22350-3605

10. SPONSOR/MONITOR'S ACRONYM(S)
ESTCP

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
ER-201209

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The purpose of this project was to: 1) validate sample collection methods and procedures that minimize variability in groundwater monitoring results; and 2) validate improved methods to optimize monitoring frequency and assess long-term concentration trends that better account for short-term variability in groundwater monitoring results. All three objectives of the project were met. The demonstration results indicated that the sample method (except active no purge) has only a modest impact on monitoring variability and concentration, suggesting sampling methods should be selected based on factors such as cost, ease of implementation, and sample volume requirements rather than concerns regarding data quality. At both sites, low flow standard (purging to parameter stability) and low flow alternative (small volume) showed the lowest variability. The results were consistent between the two sites except for the active no purge (HydraSleeve) method, which was more variable at the California site than the Texas site. Low flow small volume purge and passive no-purge (SNAP sampler) were the two best sampling methods based on the combined goals of minimizing monitoring cost and minimizing variability in monitoring results.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Tom McHugh
				47	19b. TELEPHONE NUMBER *(Include area code)* 713-522-6300

Reset

COST & PERFORMANCE REPORT
Project: ER-201209

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (continued)

Page

LIST OF FIGURES

Page

LIST OF TABLES

ACRONYMS AND ABBREVIATIONS

1,1-DCE	1,1-dichloroethene
1,2-DCA	1,2-dichloroethane
AFCEC	Air Force Civil Engineer Center
ASTM	ASTM International
bgs	below ground surface
cis-1,2-DCE	*cis*-1,2-dichloroethylene
EDC	ethylene dichloride
ESTCP	Environmental Security Technology Certification Program
ft	foot/feet
Freon-113	trichlorotrifluoroethane
ITRC	Interstate Technology & Regulatory Council
L/min	liters per minute
LFA (L)	low flow alternative, large volume purge
LFA(S)	low flow alternative, small volume purge
LFS	low flow standard
LTM	long-term monitoring
µg/L	micrograms per liter
mg/L	milligrams per liter
mL/min	milliliters per minute
mm	millimeters
NP	no purge
NRC	National Research Council
PCE	tetrachloroethylene
PI	principal investigator
QA	quality assurance
RPD	relative percent difference
SERDP	Strategic Environmental Research and Development Program
SOP	standard operating procedure

ACRONYMS AND ABBREVIATIONS (continued)

TCE	trichloroethylene
trans-1,2-DCE	*trans*-1,2-dichloroethylene
USEPA	U.S. Environmental Protection Agency
VC	vinyl chloride
VOA	volatile organic analysis
VOC	volatile organic compound

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial and technical support provided by the Environmental Security Technology Certification Program (ESTCP) for this project. In particular, we would like to thank Andrea Leeson and the ESTCP Support Office staff for their valuable technical and operational assistance.

The authors also wish to acknowledge current and past GSI Environmental employees for their contributions to this project. They include: Shawn M. Paquette and James A. Kearley (site selection and coordination); Julia M. Small (groundwater sampling and data analysis); and Ben T. Medina (groundwater sampling).

This page left blank intentionally.

EXECUTIVE SUMMARY

OBJECTIVES OF THE DEMONSTRATION

The National Research Council (NRC) has estimated that annual monitoring costs over $100,000,000 at Department of Defense facilities across the country (NRC, 2013). This cost includes on-going monitoring of roughly 40,000 groundwater monitoring wells. A primary purpose of this monitoring is to determine the long-term reduction in contaminant concentrations due to natural attenuation or active remediation. However, short-term variability in contaminant concentrations limits our ability to accurately quantify contaminant attenuation rates, increasing our monitoring costs, and limiting our ability to make appropriate site management decisions. The purpose of this project was to: 1) validate sample collection methods and procedures that minimize variability in groundwater monitoring results; and 2) validate improved methods to optimize monitoring frequency and assess long-term concentration trends that better account for short-term variability in groundwater monitoring results. All three objectives of the project were met:

1. Task 1: Validate the use of alternative field sampling procedures for the collection of groundwater samples in order to minimize variability in groundwater monitoring results.
2. Task 2: Develop and validate an improved method to optimize monitoring frequency by evaluating the site-specific short-term variability and long-term attenuation rate.
3. Task 3: Develop and validate an improved method to identify long-term concentration trends that better account for the potentially confounding effects of short-term variability.

The main report focused on the field demonstration of alternative field sampling procedures (Task 1) while Appendix E of ER-201209 Final Report documented the development and demonstration of the improved data analysis methods (Tasks 2 and 3).

TECHNOLOGY DESCRIPTION

The overall objective of the Task 1 demonstration was to validate sample collection procedures that minimize variability in monitoring results. The program provided a direct comparison of the short-term variability associated with commonly used sampling methods for volatile organic compounds (VOC), including: 1) three variations of low-flow purge; 2) SNAP Sampler (passive no purge); and 3) HydraSleeve (active no purge).

The field demonstration was conducted at eight monitoring wells at each of the two demonstration sites; in Texas and in California. Each sampling method was used six times with a total of 96 samples per method and a total of 480 groundwater samples for the demonstration program from both sites. Four VOCs were consistently detected in the samples from the Texas site while 10 were consistently detected in the wells from the California site. The resulting 3,262 data points were used to evaluate the effect of sample method on short-term variability in the monitoring results and statistical bias (i.e., difference in concentration between methods).

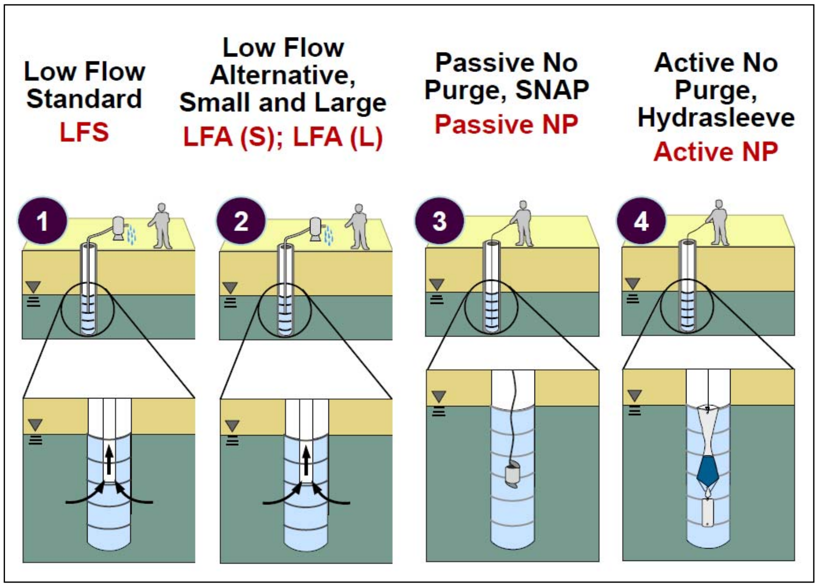


Figure ES-1 Sampling methods investigated

DEMONSTRATION RESULTS

The demonstration results indicated that the sample method (except active no purge) has only a modest impact on monitoring variability and concentration, suggesting sampling methods should be selected based on factors such as cost, ease of implementation, and sample volume requirements rather than concerns regarding data quality. At both sites, low flow standard (purging to parameter stability) and low flow alternative (small volume) showed the lowest variability. The results were consistent between the two sites except for the active no purge (HydraSleeve) method, which was more variable at the California site than the Texas site.

Although low flow alternative (large volume) and passive no purge (SNAP samplers) yielded slightly more variable groundwater monitoring results than low flow standard, this increase in variability would have little impact on the number of events needed to characterize the long-term concentration trend. However, the increased variability with the active no purge method would increase the number of sampling events required to characterize long-term concentration trend in the well. Figure ES-2 depicts the relative cost versus variability of the five sample methods. **Low flow small volume purge and passive no-purge (SNAP sampler) were the two best sampling methods based on the combined goals of minimizing monitoring cost and minimizing variability in monitoring results.**

Although statistically-significant differences in concentration were observed between methods, the average bias was small for all methods. This finding is consistent with a number of previous studies on the effect of sample method on contaminant concentration, although some prior studies have suggested a low bias for the active no purge method.

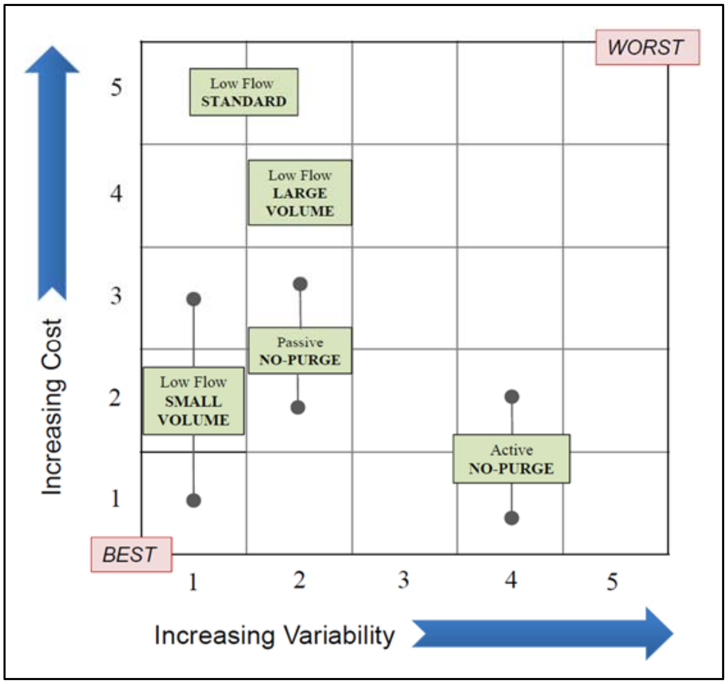


Figure ES-2 Semi-quantitative analysis of sampling methods
Gray dots indicate range of costs for shallow and deep wells.

IMPLEMENTATION ISSUES

All tested methods are mature technologies, have few end-user concerns, are straightforward to master, and can be easily applied without substantial implementation issues at most sites. Extensive peer-reviewed literature and guidance exists for all methods. Both the no purge sample methods and the alternative (i.e., fixed volume) low flow purge methods were found to be more cost effective than the standard method of low flow purge to parameter stability, with little clear benefits in data quality. The no purge methods result in little to no generation of purge waste and, therefore, may be more strongly favored at sites where management of purge waste is a logistical challenge or is expensive. Sample volume constraints for the no purge methods are the principal implementation concern where certain analyte suites require large water volumes. For those sites, the low flow alternative methods may be more applicable.

This page left blank intentionally.

1.0 INTRODUCTION

The purpose of this project was to: 1) validate sample collection methods and procedures that minimize variability in groundwater monitoring results; and 2) validate improved methods to optimize monitoring frequency and assess long-term concentration trends that better account for short-term variability in groundwater monitoring results. The specific goals of the project are as follows:

1. Task 1: Validate the use of alternative field sampling procedures for the collection of groundwater samples in order to minimize variability in groundwater monitoring results.
2. Task 2: Develop and validate an improved method to optimize monitoring frequency based on a site-specific evaluation of the short-term variability and long-term attenuation rate.
3. Task 3: Develop and validate an improved method to identify long-term concentration trends that better account for the potentially confounding effects of short-term variability.

This Cost and Performance Report focuses on the field demonstration of alternative field sampling procedures (Task 1). Tasks 2 and 3 resulted in the development of a spreadsheet-based tool that details the improved methods for evaluation of long-term groundwater monitoring data. The development and validation of this tool are documented in Appendix E of the Final Report.

1.1 BACKGROUND

Long-term monitoring (LTM) programs need to generate high-quality data by selecting monitoring points in appropriate locations and a sampling frequency that is adequate to monitor and evaluate trends at the site (U.S. Environmental Protection Agency [USEPA], 2004; Air Force Civil Engineer Center [AFCEC], 2006). To ensure data quality, limits on analytical variability were measured using laboratory duplicate samples (e.g., a relative percent difference [RPD] of 20%) and limits on sampling variability using field duplicates (e.g., an RPD of 30%) were established. If these data quality objectives are met, then the remaining variability in monitoring results is generally accepted as inherent to the nature of any monitoring system. However, for many monitoring programs, this remaining variability is much higher than the objectives for sampling and analytical variability; therefore, this high variability makes it more difficult to evaluate protection of receptors and remediation progress. Often, the only recommended course of action is to conduct more intensive monitoring because larger amounts of data are necessary to compensate for the high variability, and to identify true spatial and temporal trends in the groundwater plume.

For the purpose of this report, the “short-term variability” was defined as increases and decreases in contaminant concentrations in groundwater unrelated to the long-term reduction in source strength related to the effects of natural contaminant attenuation or active site remediation. The short-term variability typically has a time scale of less than 3 months and accounts for 60 to 70 percent (%) of the total variability in groundwater monitoring results with the long-term reduction in source strength accounting for the remaining 30 to 40% of total variability (McHugh et al., 2011). This short-term variability significantly limits the ability to understand the plume response to active remediation, source treatment, or natural attenuation.

Short-term variability distorts the long-term attenuation rate estimated from the monitoring data and the true long-term source attenuation rate. Inaccuracy in long-term monitoring trends may delay proper data interpretation and decision-making. At a minimum, variability increases monitoring costs by increasing the number of wells, sampling frequency, and data evaluation time needed to understand plume behavior. However, in many cases, variability unrelated to the true long-term concentration trend results in incorrect conclusions regarding plume stability or remedy effectiveness. In these cases, project costs can be dramatically increased by decisions to implement more aggressive remedies or to maintain frequent sampling schedules. Monitoring variability also greatly complicates the development and introduction of innovative groundwater monitoring technologies such as field-based sensors or new sampling techniques. This variability limits the ability to evaluate the accuracy, precision, and comparability of the new monitoring methods relative to the existing methods. This project utilized the improved understanding from Strategic Environmental Research and Development Program (SERDP) projects ER-1704 and ER-1705 to validate a suite of tools to minimize and manage groundwater monitoring variability.

1.2 OBJECTIVE OF THE DEMONSTRATION

The objective of this field demonstration was to validate the use of alternative groundwater sampling procedures to minimize short-term variability in groundwater monitoring records. For this purpose, the demonstration compared monitoring results obtained using standard current low-flow sampling procedures to the results obtained using alternative low-flow sampling procedures and improved procedures using two no-purge sampling technologies that are increasing in popularity. Although the demonstration did not show that the alternative methods significantly reduced monitoring variability, the results showed that two of the alternative monitoring methods (i.e., low flow small volume purge and passive no-purge [SNAP sampler]) can reduce groundwater monitoring costs with no significant change in monitoring variability compared to standard low flow purge sampling.

1.3 REGULATORY DRIVERS

As part of the regulatory clean-up process, monitoring of site contaminants in groundwater is typically required from the time that monitoring wells are initially installed during site assessment until regulatory closure is attained. The goals of LTM programs include: 1) guarding against the migration of chemicals away from the defined areas of impact (i.e., to protect receptors); and 2) monitoring the progress of groundwater remediation programs. Most commonly, the relevant regulatory requirements for site monitoring programs are qualitative rather than quantitative, leaving the regulatory project manager significant discretion with respect to the required number of monitoring wells and monitoring frequency. In other words, the regulations typically contain a general requirement to collect “sufficient” data to meet the program goal, leaving the responsible party to negotiate the number of wells and monitoring frequency with the regulator.

Short-term variability in groundwater monitoring results creates a significant barrier to the design and implementation of efficient LTM programs. Short-term variability increases both the amount of time and the amount of data needed to accurately characterize the long-term trend. When long-term trends are characterized without properly accounting for the potential confounding effects of short-term variability, the analysis may result in incorrect conclusions regarding plume stability or remediation progress. The development of alternative sampling procedures to reduce short-term variability and improve data analysis methods that better account for short-term variability will improve the efficiency and utility of LTM programs.

This page left blank intentionally.

2.0 TECHNOLOGY

The purpose of this project was to demonstrate: 1) alternative groundwater sampling procedures that reduce short-term variability in groundwater monitoring results; and 2) improved data analysis methods that better account for the confounding effects of short-term variability on the long-term concentration trend. The following sections describe Task 1, the field demonstration program.

2.1 TECHNOLOGY DESCRIPTION

This demonstration project: 1) compared monitoring variability associated with different low-flow and no-purge groundwater sampling methods; and 2) validated the use of alternative sampling procedures with these methods in order to minimize this variability. The alternative sampling procedures were compared against standard low-flow sampling procedures.

The technology for the demonstration consisted of five sampling methods as follows:

1) Low Flow Standard;
2a) Low Flow Alternative (Small Volume);
2b) Low Flow Alternative (Large Volume);
3) Passive No-Purge (SNAP Samplers); and
4) Active No-Purge (HydraSleeve).

A detailed description of the sampling methods and procedures used for the demonstration is provided in Section 5.

The hypothesis for the field demonstration program was that the alternative sampling methods would reduce the short-term variability in groundwater monitoring results compared to the low flow standard method.

2.2 ADVANTAGES AND LIMITATIONS OF THE TECHNOLOGY

2.2.1 Advantages of the Technology

Short-term variability in groundwater monitoring results complicates the attainment of LTM objectives by increasing the amount of data needed to accurately characterize the long-term concentration trend and by increasing the likelihood of incorrect conclusions regarding the long-term trend (e.g., concluding that the concentration is increasing when the true long-term trend is decreasing).

The use of low variability sampling methods and sampling procedures that minimize the variability associated with the selected method will improve the quantitative and qualitative evaluation of LTM results. For quantitative evaluations, a reduction in short-term variability will reduce the number of measurements and the evaluation time required to identify a statistically significant long-term concentration trend. For qualitative evaluations, a reduction in short-term variability will reduce the occurrence of anomalous apparent concentration trends and will make it easier to accurately determine the long-term trend through visual inspection of the monitoring results. The attainment of more visually obvious concentration trends will make it easier for stakeholders to agree on remedy effectiveness and plume stability conditions.

2.2.2 Limitations of the Technology

The technologies used for validation were not expected to eliminate all sources of contaminant data variability in groundwater monitoring results. The intent was to reduce variability, not eliminate it. This is the primary limitation of the technology. SERDP Project ER-1705 identified four general categories of variability in groundwater monitoring results:

- ***Variability Source 1: Signal Variability.*** Changes in constituent concentration within the bulk groundwater in the vicinity of the monitoring well. These changes may be due to source remediation or may reflect variations in groundwater flow direction, water table fluctuation, or other short-term changes in the fate and transport of volatile organic compounds (VOC) from the source to the monitoring point that are not directly related to the long-term trend.
- ***Variability Source 2: Aquifer and Well Dynamics.*** When constituent concentrations are stratified within the aquifer, flow dynamics within the monitoring well and the impact of the sampling method on those flow dynamics can impact the monitoring results.
- ***Variability Source 3: Sample Collection and Handling.*** VOCs, by nature, move readily from water to air. As a result, VOC loss during sample collection and handling can contribute to variability between samples and loss of accuracy in monitoring. In conventional groundwater sampling for VOCs, the water sample is poured into a sampling vial and shipped to an off-site lab in an ice chest. Other constituents may also be affected by sample collection and handing procedures. For example, metals results can be affected by the amount of sediment in the sample.
- ***Variability Source 4: Sample Analysis.*** Monitoring accuracy depends on the accuracy, precision, and reproducibility of the laboratory analysis. However, prior studies have found that analytical variability is a small component of overall monitoring variability.

The sample collection methods and procedures served to reduce Variability Sources 2 (aquifer and well dynamics) and 3 (sample collection and handling). If these two sources of variability are not the main sources of short-term monitoring variability, then the improved methods and procedures will have a limited effect on the overall variability in the monitoring results. The magnitude of short-term monitoring variability varies between monitoring wells (McHugh et al., 2011); therefore, the effectiveness of the alternative methods for reducing variability is also expected to vary. However, it is expected that the alternative sampling methods and procedures can be implemented without increasing monitoring costs and in many cases may actually reduce costs. As a result, any reduction in monitoring variability will provide benefit without cost.

3.0 PERFORMANCE OBJECTIVES

The overall objective of the demonstration was to validate sample collection procedures that minimize variability in groundwater monitoring results. In addition, the demonstration provided a direct comparison of the short-term variability associated with three commonly used sampling methods: 1) low-flow purge, 2) SNAP sampler (passive no purge), and 3) HydraSleeve (active no purge). The objective of the field demonstration was met by:

1) Applying three sampling methods with alternative procedures to eight monitoring wells at each of two demonstration sites;
2) Applying low-flow purge to parameter stability using standard procedures as the reference sampling method for each monitoring well;
3) Conducting six rounds of sampling using each sampling method; and
4) Comparing the short-term variability associated with each sampling method.

Specific performance objectives are summarized in Table 1.

Table 1 Performance objectives

Performance Objective	Data Requirements	Success Criteria	Results
Quantitative Performance Objectives			
Attainment of a complete dataset that supports multi-variate statistical analyses.	A balanced dataset based on analytical results for each planned primary sample: 480 total samples Two demonstration sites; five sampling method/procedure combinations; eight monitoring wells per site; and six sampling events per method.	Analytical results for >95% of planned primary samples (i.e., analytical results for >456 samples).	Objective met: The dataset consisted of analytical results from 478 samples out of a planned 480 sample (99.6% completeness).
Attainment of analytical results representative of constituent concentrations in the collected groundwater samples.	Results from laboratory analysis of groundwater samples. Associated QA results (e.g., laboratory QA results, duplicate analyses) to demonstrate acceptable laboratory performance.	For >90% of analyses: • Precision: RPD < 30% for field duplicate % samples; RPD <20 for laboratory duplicate results • Accuracy: standard laboratory accuracy • Sensitivity: < 1 µg/L for all VOCs.	Objective met: • 90% of field duplicate RPD values for the combined dataset were below the RPD criteria of 30%. • All lab duplicate RPD values were below the RPD criteria of 20%. • >90% of samples met laboratory accuracy criteria. • Met sensitivity criteria.

QA= quality assurance
µg/L = micrograms per liter

Table 1 Performance objectives (continued)

Performance Objective	Data Requirements	Success Criteria	Results
Quantitative Performance Objectives (continued)			
Demonstration of reduced short-term variability for one or more of the sampling methods with alternative procedures compared to the reference method.	A statistically-significant difference in short-term variability between sampling methods with lower variability in the datasets obtained using the alternative methods.	A statistically-significant difference ($p<0.05$) in variability between all datasets using Levene's test, analysis of variance (or a non-parametric equivalent), and a statistically-significant difference ($p<0.05$) in variability for a pair-wise comparison of individual alternative method dataset versus the reference method dataset using a t-test (or a non-parametric equivalent). See Section 6.0 for details.	Objective met: • Levene's test using the entire dataset shows a statistically significant difference in the variances for the sets of normalized residuals for the different sample methods ($p<0.001$). Levene's test also showed a statistically significant difference between methods for the two sites evaluated separately. • A pair-wise comparison of each alternative sample method to low-flow standard indicated that low-flow alternative (large volume) and active no purge (HydraSleeve) were significantly more variable than low-flow standard at the California site ($p<0.05$); and, low-flow alternative (large volume) and passive no purge (SNAP) were significantly more variable than low-flow standard at the Texas site.
Qualitative Performance Objectives			
Collection of representative groundwater samples.	Implementation of each sampling method using the appropriate reference sampling procedures or alternative sampling procedures in accordance with the sample method SOP.	Documentation of appropriate implementation of each sample method in accordance with the SOP for >95% of samples.	Objective met: Documentation of appropriate implementation of each sample method recorded for 100% of samples.
Ease of implementation of the alternative sampling procedures.	Field experience implementing the groundwater sampling procedures.	Validated SOPs for the alternative sampling procedures that can be implemented by field sampling personnel with a typical level of qualifications and experience.	Objective met: SOPs for each method included in the final report.

SOP= standard operating procedures

4.0 SITE/PLATFORM DESCRIPTION

The improved sample collection methods were demonstrated at two field demonstration sites at eight monitoring wells each. One field demonstration site was in the Houston, Texas area. The second field demonstration site was in Los Angeles, California.

The goal of the site selection process was to identify monitoring wells representative of those typically using for long-term monitoring of contaminant plumes. The following criteria were used to identify the selected sites:

1. Access to site for duration of demonstration;
2. Historical monitoring data;
3. One or more contaminants detected during >80% of historical monitoring events;
4. Well diameter between 1 inch and 4 inches;
5. Well screen length between 5 feet (ft) and 20 ft; and
6. Site located close to principal investigator (PI) or co-PI to minimize mobilization costs.

4.1 DEMONSTRATION SITE #1: HOUSTON, TEXAS

4.1.1 Site Location and History

The selected demonstration site #1is in northwest Houston, and is the location of a former manufacturing plant; the site is not currently active. Affected groundwater was detected in 1992, and a groundwater recovery and treatment system has been operating since May 1997.

4.1.2 Site Geology/Hydrogeology

Layers of silt, sandy clay, and clay are present from approximately 0 to 14 ft below ground surface (bgs), after which a layer of fine silty sand extends to 52 ft bgs, and is an unconfined aquifer. The water table in the aquifer is at approximately 29 ft bgs. Two more layers of sand (150 – 170 ft bgs) and a deeper aquifer (220 – 600 ft bgs) exist, and are separated by layers of clay.

4.1.3 Contaminant Distribution and Selected Wells

The affected groundwater plume extends approximately 950 ft in length across the property with the following constituents: tetrachloroethylene (PCE), trichloroethylene (TCE), cis-1,2-dichloroethylene (cis-1,2-DCE), vinyl chloride (VC), and 1,1-dichloroethene (1,1-DCE).

The following table highlights the key construction information for the eight wells selected for the field demonstration, as well as the historical contaminant range at each well.

Table 2 Key information on selected wells at Demonstration Site #1

Well ID	Well Diameter (inches)	Screen Length (ft)	Screen Depth Interval (ft bgs)	Key Contaminants	Historical Contaminant Range (mg/L)
MW-02A	2	10	30 - 40	PCE TCE cis-1,2-DCE 1,1-DCE	0.002 – 0.11
MW-06	2	10	25 - 35		0.002 – 0.5
MW-13	2	10	27 - 37		0.02 – 0.15
MW-15	2	10	25 - 35		0.002 – 0.01
MW-23A	2	10	28 - 38		0.006 – 0.02
MW-25A	2	10	28 - 38		0.001 – 0.003
MW-26	2	10	27.5 – 37.5		<0.00014 – 0.02
TW-01	2	10	27 - 37		0.005 – 0.03

in.= inch
mg/L= milligram per liter

4.2 DEMONSTRATION SITE #2: LOS ANGELES, CALIFORNIA

4.2.1 Site Location and History

Demonstration site #2 is located in Santa Fe Springs, California, near Los Angeles, and is the location of a former chemical repackaging facility. The site is currently an auto repair and staging lot. Affected groundwater was detected in the late 1980s, and a soil vapor recovery and treatment system is the only on-site treatment currently operating.

4.2.2 Site Geology/Hydrogeology

The site is located in the flood plain of the San Gabriel River system, south and east of Los Angeles. From approximately 0 to 40 ft bgs interbedded sands, silts, and gravels are present. Consistently distributed tight clay exists from about 40 to 45 ft, isolating the shallow water table from the deeper aquifer. Below 45 ft, a fairly consistent medium sand is present to approximately 80 ft. Saturation and water level in the deeper zone fluctuates, but during initial site characterization water levels were about 55 ft deep.

4.2.3 Contaminant Distribution and Selected Wells

The affected groundwater plume extends throughout the property from both on-site and off-site sources, with the following primary constituents: PCE, TCE, cis-1,2-DCE, 1,2-dichloroethane (1,1-DCA), 1,1-DCE, 1,4-dioxane, ethylene dichloride (EDC), Chloroform, trans-1,2-dichloroethylene (trans-1,2-DCE), and Trichlorotrifluoroethane (Freon-113). The following table highlights the key construction information for the eight wells selected for the field demonstration, as well as the historical contaminant range at each well.

Table 3. Key information on selected wells at Demonstration Site #2.

Well ID	Well Diameter (inches)	Screen Length (ft)	Screen Depth Interval (ft bgs)	Key Contaminants	Historical Contaminant Range (mg/L)
MW-13	2	10	52 - 62	PCE TCE cis-1,2-DCE 1,1-DCA 1,1-DCE 1,4-Dioxane EDC Chloroform trans-1,2-DCE Freon-113	0.002 – 0.45
MW-14	2	10	55 - 65		0.002 – 0.7
MW-15	2	10	54 - 64		0.002 – 0.9
MW-17	2	10	56 - 66		0.002 – 0.08
MW-20	2	10	57 - 67		0.002 – 0.11
MW-21	2	10	53 - 63		0.002 – 2.3
MW-23	4	10	71 – 81		0.002 – 0.14
MW-24	4	10	67 - 77		0.002 – 0.21

This page left blank intentionally.

5.0 TEST DESIGN

The overall objective of the demonstration was to validate sample collection procedures that minimize variability in groundwater monitoring results. In addition, the demonstration provided a direct comparison of the short-term variability associated with three commonly used sampling methods: 1) low-flow purge, 2) SNAP sampler (passive no purge), and 3) HydraSleeve (active no purge). The sample collection methods with improved sampling procedures were demonstrated in eight monitoring wells at each of two field demonstration sites (16 wells total).

5.1 CONCEPTUAL EXPERIMENTAL DESIGN

The sampling program consisted of three main types of sampling methods: 1) low-flow purge, 2) passive no purge (SNAP sampler), and 3) active no purge (HydraSleeve).

5.1.1 Sample Method/Procedure Combinations

The low-flow purge sampling method was implemented conventionally as the reference method with two variations of the method to include improved sampling procedures (fixed small volume purge and fixed large volume purge). The two no-purge sampling methods were implemented using those devices' standard sampling procedures modified to include the improved procedures that are relevant to each specific method. The sampling methods/procedures are summarized in Table 4 and Figure 1 below.

Table 4 Summary of sampling methods/procedures

Sampling Method	Sampling Procedures	Sampling Method/Procedure Combination Used
Low-flow purge	Reference method (standard)	1) Low flow standard
	Fixed small volume purge (alternative)	2a) Low flow alternative, small volume
	Fixed large volume purge (alternative)	2b) Low flow alternative, large-volume
Passive no-purge	SNAP sampler	3) Passive no-purge (SNAP sampler)
Active no-purge	HydraSleeve	4) Active no-purge (HydraSleeve)

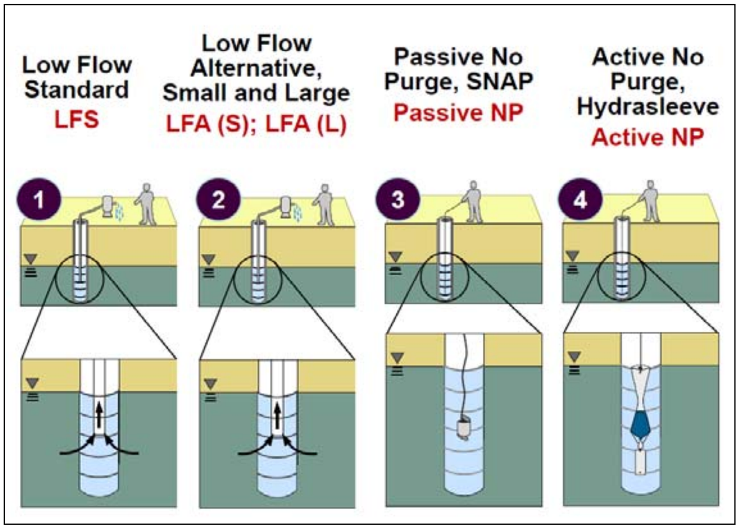


Figure 1 Summary of sampling methods used in demonstration program

Although this project used specific sampling methods to represent general categories of sample collection methods, it is expected that the project results are applicable to other methods within the category. For example, the project team would expect the project findings obtained using the SNAP sampler to be applicable to other passive no purge methods, such as a passive diffusion bag sampler.

5.1.2 Demonstration Sites, Monitoring Wells, and Rounds of Sampling

The field program was conducted at two demonstration sites. At each site, each sample method/procedure combination was used to collect a groundwater sample from each of eight monitoring wells. Except for low-flow alternative (small volume) and low-flow alternative (large volume), each sampling method was implemented during separate sampling events, with 10 to 20 days between each sample event. Low-flow alternative (small volume) and low-flow alternative (large volume) were implemented sequentially during a single sample event.

Each round of sampling consisted of four events in which all method/procedure combinations were implemented, and six rounds of sample collection were conducted. Each round of sampling was completed over a period of approximately 60 days (i.e., four sampling events with 10 to 20 days between each event), resulting in a total of approximately 1 year to complete the six rounds of sampling. The sampling program yielded a dataset of 480 groundwater samples (i.e., five sample method/procedure combinations, eight wells, and six rounds of sampling at each of two demonstration sites).

5.2 BASELINE CHARACTERIZATION

As discussed in Section 4, the selection of demonstration sites and specific monitoring wells within each site was based on the identification of several factors. As such, no additional baseline characterization was conducted prior to executing the demonstration.

5.3 TREATABILITY OR LABORATORY STUDY RESULTS

No treatability or laboratory studies were conducted as part of this field demonstration.

5.4 FIELD TESTING

At each of the two demonstration sites, the field program was implemented through 24 field sampling events (i.e., six rounds of sampling with four sampling events for each round (see Table 5).

Each sample method/procedure combination was used to collect a groundwater sample from each of eight monitoring wells. Except for the low-flow alternative (fixed small volume purge and fixed large volume purge), each sampling method was implemented during separate sampling events, with 10 to 20 days between each sample event. The time period between events was designed to allow the well and surrounding aquifer to re-stabilize to a natural/ambient state, so that each sampling method was not impacted by activities of the previous events.

Tables 5 and 6 summarize the number of sample rounds, and the general schedule of sampling events per sampling round. The methods were applied sequentially, with approximately 10-20 days between each sampling event.

Table 5 Field testing schedule

Sampling Event	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Program Week	**2**	**4**	**6**	**8**	**10**	**12**	**14**	**16**	**18**	**20**	**22**	**24**	**26**	**28**	**30**	**32**	**34**	**36**	**38**	**40**	**42**	**44**	**46**	**48**
Sample Method/Procedure																								
1) Low-flow standard	X				X				X				X				X				X			
2a) and 2b) Low-flow alternative procedure (small and large volumes)		X				X				X				X				X				X		
3) Passive no-purge (SNAP)			X				X				X				X				X				X	
4) Active no-purge (HydraSleeve)				X				X				X				X				X				X

Table 6 Summary of Task 1 field demonstration program

Number of Demonstration Sites	Number of Wells per Site	Sample Method/Procedure	Sample Rounds	Total Samples
2	8	Reference method: low-flow, standard procedure	6	96
		Low-flow, alternative procedure (small volume purge)	6	96
		Low-flow, alternative procedure (large volume purge)	6	96
		Passive no-purge (SNAP sampler)	6	96
		Active no-purge (HydraSleeve)	6	96
			Total number of samples collected for primary data set	480

Note: Sample count does not include field and lab duplicates or other QA/quality control samples.

5.5 SAMPLING METHODS

The purpose of this technology demonstration was to compare the short-term variability associated with three common groundwater sampling methods, and to evaluate whether or not implementing methods using alternative sampling procedures will reduce short-term variability.

As a result, the technology for the demonstration consisted of five sampling methods as follows:

1) Low flow standard,
2a) Low flow alternative (small volume),
2b) Low flow alternative (large volume),
3) Passive no-purge (SNAP samplers), and
4) Active no-purge (HydraSleeve).

5.5.1 General Sampling Method Types Overview

The general sampling method types can be categorized as:

1. Low-flow sampling,
2. Passive no-purge, and
3. Active no-purge.

5.5.1.1 Low-Flow Sampling

Low-flow sampling involves use of a pump (either an above-ground peristaltic pump or a down-hole electric pump) to remove water from the monitoring well at a low flow rate (<1 liters per minute [L/min]) to minimize drawdown and disturbance of the well. As most commonly implemented today, water is purged from the monitoring well until field parameters (e.g., temperature, pH, specific conductance, and either dissolved oxygen or oxidation-reduction potential) stabilize and then the groundwater sample is collected.

5.5.1.2 Passive No-Purge (SNAP Sampler)

Passive no-purge sampling involves placement of a sampling device (either a diffusion bag [ITRC, 2004], or a SNAP sampler [ITRC, 2007]) into the monitoring well approximately 2 weeks prior to the sampling event and allowing the sampler to equilibrate with the water in the monitoring well (Figure 2). After the equilibration period, the sampling device is closed (if needed) and the sample is removed from the well. The resulting sample is representative of water in the well under ambient flow conditions.

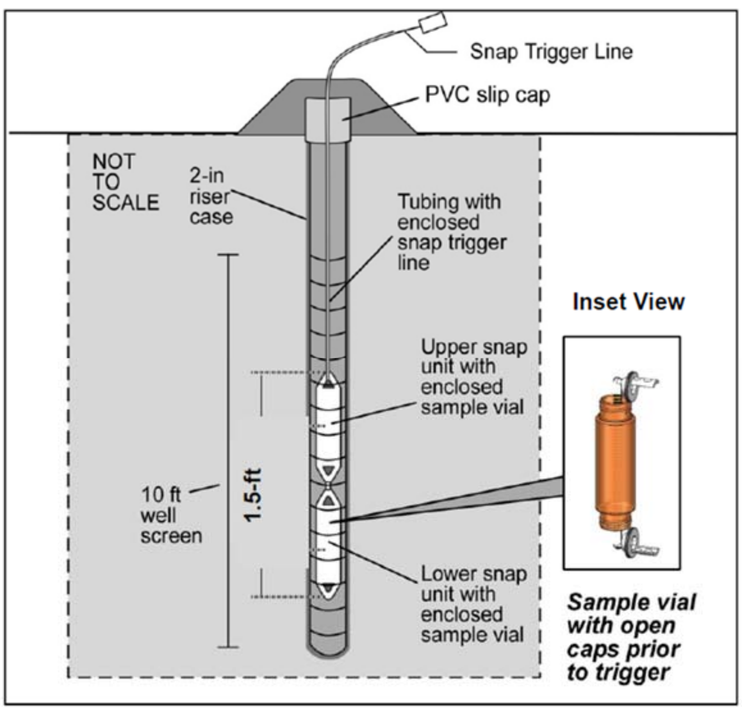


Figure 2 Use of SNAP sampler to collect groundwater samples

5.5.1.3 Active No-Purge (HydraSleeve)

Active no-purge sampling involves active sample collection from an unpurged monitoring well. The HydraSleeve is an active groundwater sampling device that is filled by pulling the sampler upwards through the screened interval of the monitoring well (Figure 3). The HydraSleeve sampler is installed in the monitoring well approximately 2 weeks prior to sample collection to allow the sampler to equilibrate with the groundwater and to avoid the mixing that would occur if the sampler were installed immediately prior to sample collection.

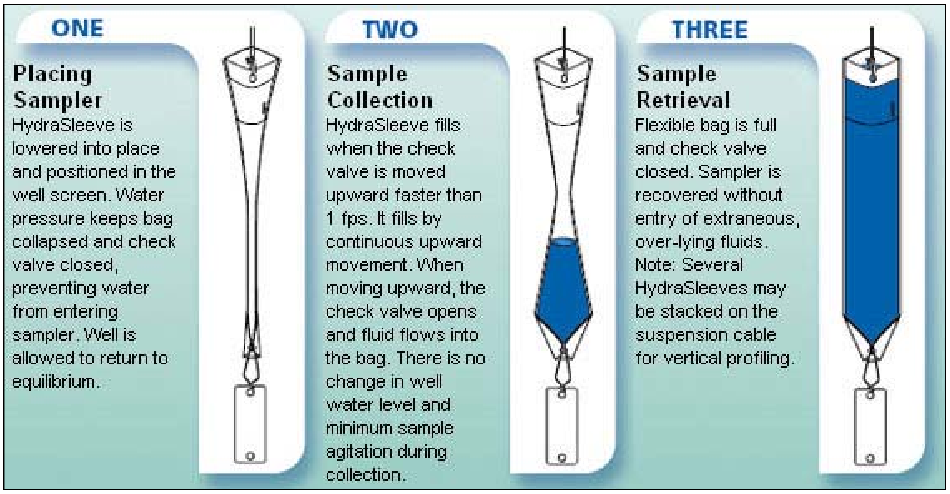


Figure 3 HydraSleeve sampling system
(Source: http://www.HydraSleeve.com)

5.5.2 Standard and Alternative Sampling Procedures

The specific procedures utilized during sample collection can vary somewhat between sampling teams, but are based on the procedures recommended in applicable guidance documents. These guidance documents often do not reflect the most recent innovations and improvements in sampling procedures, and implementation in this study is designed to incorporate some variations inherent to the procedures themselves. This field demonstration will demonstrate the ability of alternate sampling procedures to reduce short-term variability in groundwater monitoring results. Standard and improved sampling procedures are summarized in Table 7. Although some of these improved procedures are already utilized by several field sampling teams, it is uncommon for all of the improved procedures to be used together.

Table 7 Summary of standard and alternative sampling procedures

Procedure Element	Standard Procedure	Alternative Procedure
Equipment installation	Install low-flow sampling equipment immediately prior to sample collection.	Install sampling equipment at least 2 weeks before sample collection.
Sample collection elevation	Collect sample from approximately the same elevation within the well screen during each sample event.	Mark sampling equipment to ensure that sample is collected from the same elevation (+/- 1 inch) during each sample event.
Purge volume	Purge volume based on field parameter stability. Purge volume varies between sample events.	Constant purge volume for each sample event (3 L for smaller volume purge, 18 L for larger volume purge)
Pumping rate	If purge rate is >250 mL/min, then pumping rate lowered to 250 mL/min for sample collection.	Constant pumping rate used for purging and sample collection, up to 1000 mL/min.

Vial filling method and rate	Allow water to flow down the inside wall of the volatile organic analysis (VOA) vial.	Insert tube into vial and fill vial from bottom. Pull tube up as vial fills, keeping tube below water level in vial.
Removal of bubbles from vials	Check for bubbles in filled vials. Reopen and top-off vial to remove any bubbles larger than 1 millimeter (mm) in diameter.	Accept any vial at least 95% full (i.e., < 2 mL headspace). If vial contains > 2 mL headspace, discard and fill new vial.

L=liters
mL/min=milliliters/min
VOA= volatile organic analysis
Mm=millimeter

Each element of the standard and alternative procedure is discussed in more detail in the Final Report.

5.5.3 Five Sampling Methods and Procedures

The general sampling method types, as well as standard and improved procedures, were combined to create five sampling method/procedure combinations that were used throughout the field program demonstration. The five methods and their respective sampling procedures are summarized in Table 8 below.

Table 8. Summary of sampling methods and procedures implemented in field program.

	1) Low-Flow Standard	2a) and 2b) Low-Flow Alternative, Small- and Large Volume	3) Passive No-Purge (SNAP Samplers)	4) Active No-Purge (HydraSleeve)
Equipment	Install day of sampling	Install dedicated equipment	Install dedicated equipment	Install dedicated equipment
Intake Depth	Approximately constant	Precise, constant sample depth	Precise, constant sample depth (sample top vial only)	Water column of 1.0 to 1.5 times the length of sampler (2.5 ft) (GeoInsight, 2010)
Well Purge	Purge to parameter stability	Fixed Volume: Small (3 L) and Large (18 L)	None	None
Flow Rate	Varies between purge and sample, <250 mL/min	Constant during purge and sample, <1000 mL/min	None	None
Vial Fill	Side pour method	Bottom fill method	None	Bottom fill method
Vial Bubbles	Remove >1 mm bubbles	>2 mL headspace, replace vial	>2 mL headspace, replace vial	>2 mL headspace, replace vial

5.6 SAMPLING RESULTS

Sample results were analyzed in two different ways. First, the effect of sample method on short-term variability in constituent concentrations was addressed. The hypothesis for the field program was that alternative sampling methods would reduce the short-term variability in measured constituent concentrations compared to the reference method, low-flow standard.

Secondly, the statistical bias between sample methods (i.e., the difference in concentrations) was evaluated.

The results from this demonstration, combined with the results from SERDP projects ER-1704 and ER-1705, support a conceptual model that short-term variability in groundwater monitoring results is mostly attributable to small-scale spatial variability in contaminant concentrations within an aquifer (see Figure 4), and varying degrees of ambient mixing with the well screen between sampling events. This conceptual model is supported by the following findings:

- *Field Duplicate Variability is Not Significant:* The results from both ER-1705 and this project showed little variation in field duplicate concentrations (i.e., typically less than 10%). This indicates that laboratory analytical variability is small relative to other sources of variability in monitoring results.
- *Few Important Differences Between Sample Methods:* The results from this field demonstration show that no-purge and low flow purge sample methods yield monitoring results of similar quality when evaluated in terms of short-term variability and statistical bias, with the exception of Active No Purge at some sites. For most methods, variability associated with sample collection procedures is small relative to other sources of short-term variability.
- *Concentrations vary with Purge Volume:* The results from ER-1704, ER-1705, and this project show that contaminant concentrations can vary with purge volume; however, the magnitude and pattern of change varies from well to well. The change in concentration with purge volume exceeds two-fold in approximately 10% of wells and the direction of change appears to be random. Contaminant concentration may either increase or decrease with purge volume and in some wells may increase and then decrease (or decrease then increase). Contaminant concentrations may not stabilize when purge parameters stabilize.
- *Concentrations vary over Short Time Periods:* The results from both ER-1705 and this project show that contaminant concentrations can vary over short time periods (i.e., days to weeks). The concentration change on a time scale of days to weeks is much higher than the field duplicate variability and somewhat higher than the purge variability. The variation in concentration over short time periods is mostly time independent (i.e., the magnitude of change is largely independent of the time between sampling events).

This page left blank intentionally.

6.0 PERFORMANCE ASSESSMENT

A summary of the performance objectives for this demonstration, along with an overview of technology performance, was presented in Section 3. This section includes a detailed assessment of technology performance based on the quantitative data presented in Section 5. Following completion of the sampling and analysis program, the data were reviewed to determine whether the success criteria for each performance objective have been met. The evaluation of each individual performance objective is discussed below, with references to relevant supporting results in Section 5.

6.1 ATTAINMENT OF A COMPLETE DATASET

The field program consisted of six sample events for each of five sample methods implemented at eight monitoring wells at each of two demonstration sites (Texas and California). The complete sampling program was expected to yield 480 groundwater samples (not including field duplicates). For the Texas site, no sample was recovered using the HydraSleeve in MW-15 for Sampling Event 2 and MW-13 for Sampling Event 6, and logistical constraints prevented collection of replacement samples. As a result, the dataset consisted of analytical results from 478 samples (99.6% completeness).

The complete dataset consists of chemical concentration results from two sites, summarized as follows:

Table 9 Summary of complete dataset from two demonstration sites

Site	Number of Wells	Sample Methods Tested[1]	Number of Sample Events per Method	Number of Samples Collected	Number of Chemicals Detected[2]	Total Concentration Measurements
Texas	8	5	6	238*	4	862*
California	8	5	6	240	10	2400
					Total	**3262**

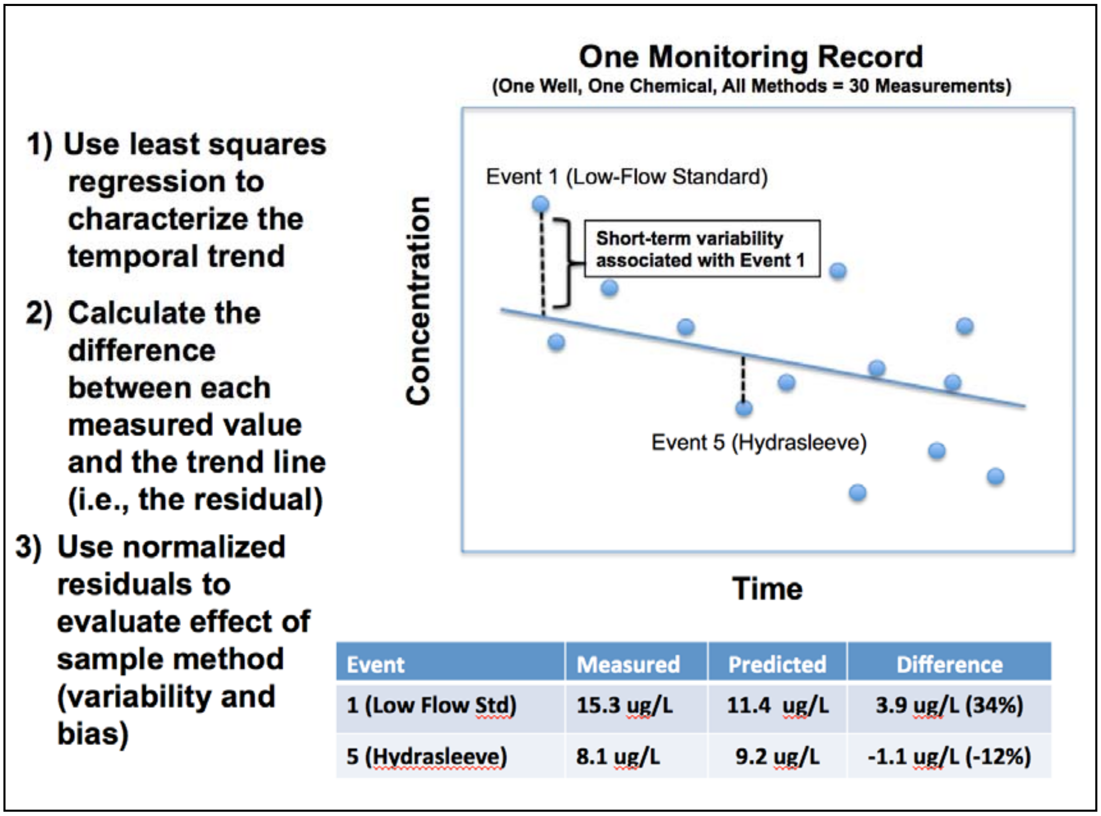
Notes:
(*) Missing data from the Texas site includes: i) TCE in Well MW-25A (DF=3%), cis-1,2-DCE in Well MW-25A (DF=0%), 1,1-DCE in Well MW-25A (DF=0%); and ii) no sample was recovered from recovered using the HydraSleeve in MW-15 for Sampling Event 2 and MW-13 for Sampling Event 6.
1. The sample methods tested include: low-flow standard, low-flow alternative (3L), low-flow alternative (18L), passive no-purge (SNAP), and active no-purge (HydraSleeve).
2. These chemicals were detected in >90% of the events/methods for at least one well at the site.

6.1.1 Data Clean-Up

Of the 3262 concentration measurements that were retained for further processing and statistical analyses, 37 were non-detect results. For these 37 data points, the detection limit was substituted for the non-detect result.

6.1.2 Data Processing: Short-Term Variability Component for Individual Measurements

The hypothesis for Task 1 was that alternative sample methods would reduce the short-term variability in groundwater monitoring results. In order to test this hypothesis, the project team needed to quantify the short-term variability associated with each concentration measurement. For each concentration measurement, the short-term variability was defined as the difference between the measured concentration and the long-term concentration (see Figure 4).



Event	Measured	Predicted	Difference
1 (Low Flow Std)	15.3 ug/L	11.4 ug/L	3.9 ug/L (34%)
5 (Hydrasleeve)	8.1 ug/L	9.2 ug/L	-1.1 ug/L (-12%)

Figure 4 Method to quantify the short-term variability component for each concentration measurement

6.1.3 Data Processing: Paired Measurements

The short-term variability has also been evaluated by analyzing the change in concentration between paired measurements. For paired measurements, the change in concentration between the sample pair was calculated as:

Concentration Change = (Higher Concentration / Lower Concentration) x 100 – 100%

Using this calculation method, the change in concentration is always expressed as a positive value. This calculation is similar to RPD, except that the maximum possible value for RPD is 200% while there is no upper bound value for concentration change.

6.2 EFFECT OF SAMPLE METHOD ON SHORT-TERM VARIABILITY IN CONSTITUENT CONCENTRATIONS

The hypothesis for the field program was that alternative sampling methods would reduce the short-term variability in measured constituent concentrations compared to the reference method, low flow standard.

6.2.1 Evaluation of Sample Methods Based on the Short-Term Variability Factor for Individual Measurements

A pair-wise comparison of each alternative sample method to low-flow standard indicated that low-flow alternative (large volume) and active no purge (HydraSleeve) were significantly more variable than low-flow standard at the California site ($p<0.05$); and low-flow alternative (large volume) and passive no purge (SNAP) were significantly more variable than low-flow standard at the Texas site (Figure 5). Although the differences in variability were statistically significant, further analysis indicates that only the variability in the active no purge (HydraSleeve) method is likely to increase the amount of monitoring data needed to characterize the long-term change in concentration (see Section 6.3.1).

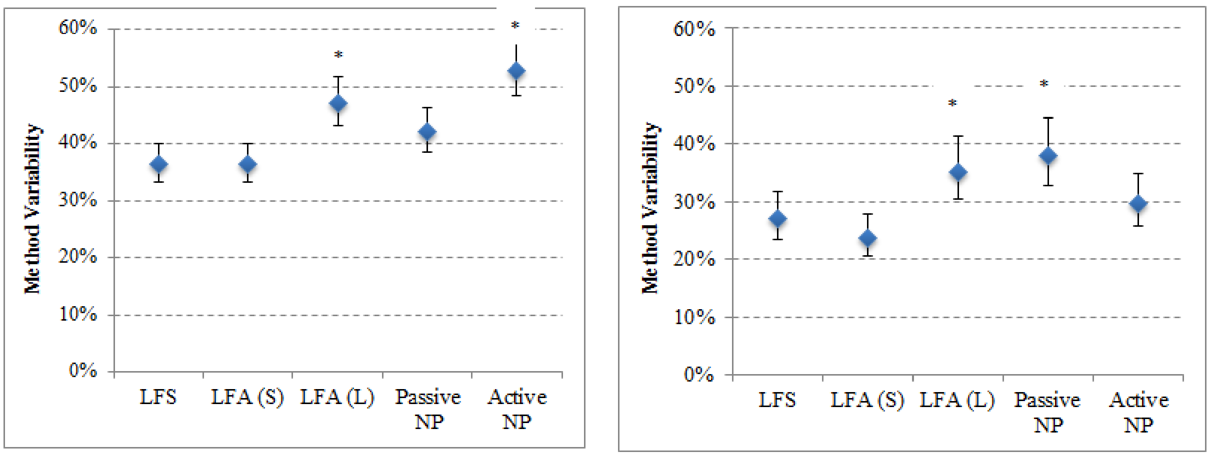


Figure 5. Short-term variability by sample method: results for individual sites.
The graphs show the standard deviation of the normalized residuals (short-term variability factors) for each sample method.The error bars show the 95% confidence interval for the standard deviation.* = method variability is significantly higher than low flow standard ($p<0.05$).LFS = Low Flow Standard,LFA (L) = Low Flow Alternative, Large Volume Purge,LFA(S) = Low Flow Alternative, Small Volume Purge,Passive NP = Passive No Purge (SNAP Sampler),Active NP = Active No Purge (HydraSleeve).

6.2.2 Evaluation of Sample Methods Based on the Variability between Paired Measurements

The evaluation of variability between paired samples provides a second method to evaluate the effect of sample method on short-term variability and also of other factors contributing to the overall short-term variability. For comparison, the project team also looked at the difference in concentration between field duplicate samples as well aslow flow alternative (small volume) samples and low flow alternative (large volume) samples collected on the same day. As shown in Figure 6, the difference in concentration between sample events was similar for all sample methods except Active NP (HydraSleeve). For low flow standard, low flow alternative (small volume), low flow alternative (large volume), and passive no purge (SNAP), the median concentration change for paired samples ranged from 20% to 24% and the 90th percentile concentration ratio ranged from 90% to 130%. However, for active no purge (HydraSleeve), the median concentration ratio was 43% and the 90th percentile was 500%. This higher variability for active no purge (HydraSleeve) was observed at the California site but not the Texas Site. The SERDP study ER-1705 also found that the active no purge (HydraSleeve) method yielded results that were more variable than those obtained using the low flow standard method when used at the Hill Air Force Base site. This prior finding suggests that the higher variability associated with the active no purge method is not unique to the California site.

The difference in concentration between sample events was similar to the difference in concentration between low flow alternative (small volume) and low flow alternative (large volume) samples collected on the same day (median concentration difference was 11% and 90th percentile was 100%), but was much larger than the difference in concentration between field duplicates (median difference was 6% and 90th percentile was 30%).

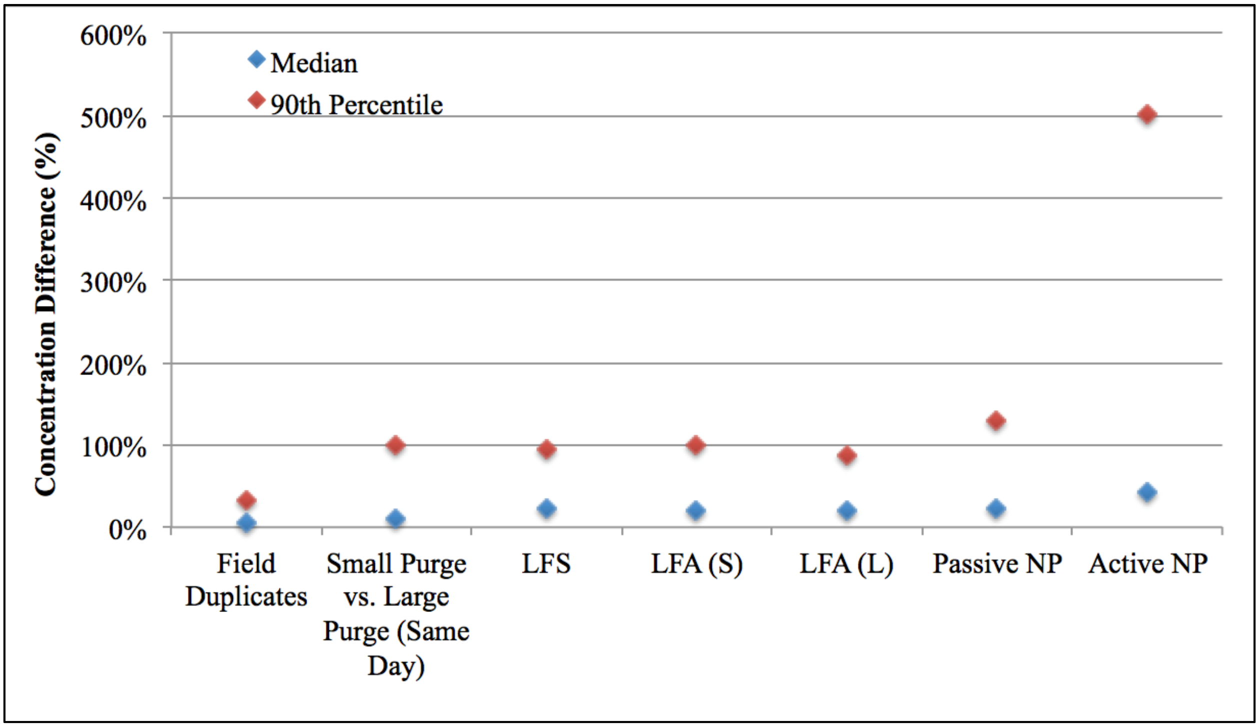


Figure 6. Effect of sample method on variability between paired concentration measurements.

6.2.3 LFS = Low Flow Standard, LFA (L) = Low Flow Alternative, Large Volume Purge, LFA(S) = Low Flow Alternative, Small Volume Purge, Passive NP = Passive No Purge (SNAP Sampler), Active NP = Active No Purge (HydraSleeve).Impact of Sample Method Variability on Evaluation of Long-Term Concentration Trends

For each of the five sampling methods, the Task 1 demonstration program yielded a dataset of six sampling events from a total of 16 monitoring wells with 4 to 10 contaminants detected in each monitoring well. These five datasets were used to evaluate how the differences in short-term variability between the sample collection methods would affect the ability to characterize the LTM trend (Table 10).

Table 10 Effect of sample method on amount of monitoring required to characterize the long-term concentration trend

Sampling Method	Short-Term Variability (log scale)[1]	Quarterly Monitoring Events[2]	Increase Relative to Low Flow Std.[3]
Low flow standard	0.45	28	N/A
Low flow alternative small volume	0.47	28	0%
Low flow alternative large volume	0.50	30	7%
Passive NP	0.52	30	7%
Active NP	0.81	39	39%

Notes:
1) Short-term variability factor for Tier 2 Optimization tool; calculated as the standard deviation of the natural log of the residuals for each monitoring record.
2) Number of quarterly monitoring events required to characterize a long-term concentration trend with medium accuracy for a monitoring well with a true attenuation rate of 0.14 yr-1 (half life of 5 years) and a short-term variability factor equal to that measured for the specific sampling method.
3) Percent increase in monitoring (relative to low flow standard.) required to characterize the long-term concentration trend with the same level of accuracy.
4) See Appendix E of Final Report for an explanation of how the "amount of monitoring data required" was determined.

The results of this analysis indicate that the small differences in variability between low flow standard, low flow alternative (small volume), low flow alternative (large volume), and passive no purge (SNAP) have little effect on the amount of data needed to characterize the long-term monitoring trend. However, the variability associated with active no purge (HydraSleeve) results in a 39% increase in the amount of data needed to characterize the long-term trend. As shown in Figure 6, the active no purge method resulted in some individual measurements that were very different from the average concentration. These large errors have a correspondingly large effect on the ability to accurately characterize the long-term trend. As a result, the variability associated with the active no purge method had a larger effect on the amount of data needed to characterize the long-term trend than the variability associated with the other sampling methods.

6.3 EVALUTION OF CONCENTRATION DIFFERENCE BETWEEN SAMPLE METHODS

The primary goal of this field demonstration was to evaluate the effect of sample methods on short-term variability in monitoring results. However, the study design also allows for an evaluation of statistical bias between sample methods (i.e., the difference in concentrations).

6.3.1 Overall Statistical Bias between Methods

The project team utilized the demonstration dataset to determine whether individual sample methods yielded VOC concentration results that were statistically higher (i.e., biased high) or lower (i.e., biased low) relative to the dataset as a whole.

Statistical Test Results: A t-test was used to evaluate whether individual sample methods were biased low or high relative to the full dataset. A method was determined to be biased high (or low), if the average statistical bias was different from zero at the 95% confidence level.

Overall, the biases between methods were low. The average statistical bias typically ranged from +20% to -15% (see Figure 7). The most pronounced differences between methods were observed at the Texas site where the two no purge methods showed a statistical bias about +20% and the three purge methods showed a statistical bias of about -12%. However, this result was driven largely by three monitoring wells where the no purge concentrations were consistently higher than the purge concentration. This well-specific difference between the sample methods appeared to be more important than any well independent differences in statistical bias between the methods. In other words, although we observed statistically-significant differences in VOC concentration between sample methods, the magnitude of these concentration differences would likely not be important for site decision-making.

California Site **Texas Site**

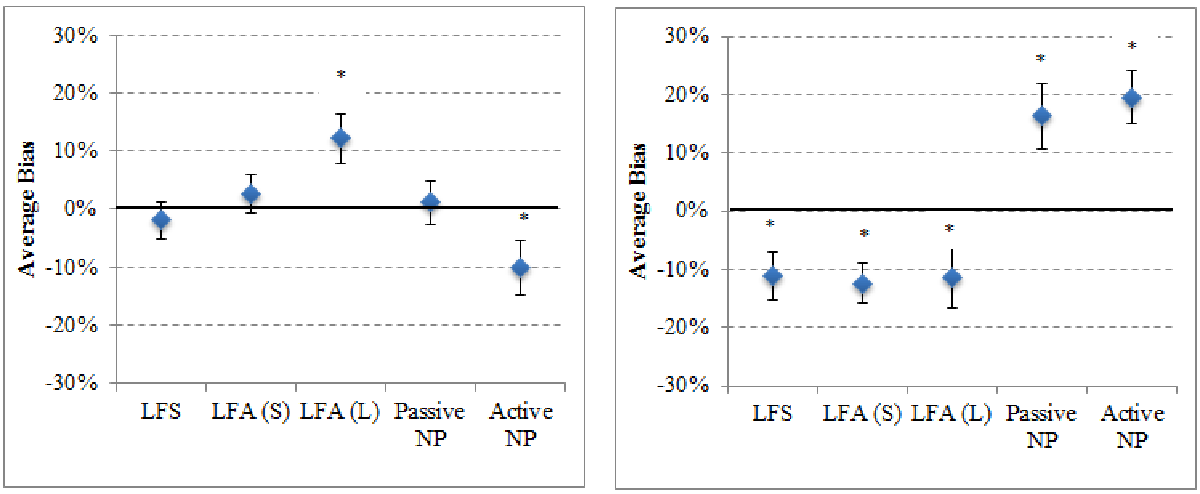


Figure 7 Difference in contaminant concentration by sample method: results for individual sites

The graphs show the average of the normalized residuals (i.e., average statistical bias) for each sample method. The error bars show the 95% confidence interval for normalized residual. * = method bias is significantly different from zero ($p<0.05$). LFS = Low Flow Standard, LFA (L) = Low Flow Alternative, Large Volume Purge, LFA(S) = Low Flow Alternative, Small Volume Purge, Passive NP = Passive No Purge (SNAP Sampler), Active NP = Active No Purge (HydraSleeve).

7.0 COST ASSESSMENT

Costs incurred during the field program demonstration for each of the four groundwater sampling methods were tracked and analyzed: low flow standard, low flow alternative, passive no purge (SNAP), and active no purge (HydraSleeve). Incurred costs for the field program demonstration were then extrapolated in order to estimate costs for implementing each technology at a standard site.

7.1 COST MODEL

The field demonstration included five different sampling methods, each implemented at two sites. Key cost elements that were tracked included: 1) project planning and preparation, 2) field implementation, and 3) data evaluation and reporting.

7.1.1 Cost Element: Project Planning and Preparation

Project planning for the field demonstration included site selection, review of existing site data, attainment of site access, and detailed work plans for all sampling events. Additionally, supplies such as submersible pumps, SNAP samplers, and HydraSleeve samplers were purchased prior to field mobilization.

7.1.2 Cost Element: Field Program

Costs for the field program include labor hours sample collection during sampling events. Additionally, equipment rental, purchase of replacement parts as well as sample analysis was tracked.

7.1.3 Cost Element: Data Evaluation and Reporting

Following completion of the demonstration, the results and data were reviewed, analyzed, and recorded into a report to document the findings.

7.2 COST DRIVERS

Cost drivers for the specific sampling methods are presented below.

7.2.1 Low Flow Standard

Cost drivers for implementation of low-flow standard include: 1) labor hours, 2) equipment purchase for deeper wells, 3) waste handling and disposal, and 4) equipment rental. A significant cost driver for low flow standard was labor hours, which included taking measurements of water parameters, and waiting for parameter stabilization, which can vary per well and event. Other cost drivers for low flow standard relate to equipment needs for deeper wells. The capital cost is higher for the purchase of submersible pumps in deeper wells, whereas a peristaltic pump may be rented for shallow wells. Regardless of the pump selected, other equipment, including a water quality meter and turbidity meter, will need to be rented for each event.

7.2.2 Low Flow Alternative (Large Volume Purge)

Cost drivers for implementation of low flow alternative (large volume) are the same as those for low flow standard. However, the length of time associated with sampling each well for low flow alternative (large volume) is more predictable than low flow standard, since the purge volume is fixed. Additionally, low flow alternative methods had higher purging rates than low flow standard (i.e., average of 600 mL/min for low flow alternative at the Texas site versus less than 250 mL/min for low flow standard). Additionally, no rental equipment for measuring water quality parameters was required.

7.2.3 Low Flow Alternative (Small Volume Purge)

Cost drivers for implementation of low flow alternative (small volume) are the same as those for low flow alternative (large volume). However, fewer labor hours are required because less time is needed to pump 3 L rather than 18 L.

7.2.4 Passive No Purge (SNAP)

Cost drivers associated with implementation of SNAP samplers include: 1) initial equipment purchase, and 2) replacement SNAP sampler vials per sampling event. This method requires that each well is outfitted with equipment, including: SNAP sampler(s), trigger line, well dock, and SNAP sampler vials. In addition, replacement vials must be purchased for each sampling event.

7.2.5 Active No Purge (HydraSleeve)

Cost drivers associated with implementation of HydraSleeve at the site include: 1) initial equipment purchase, and 2) replacement HydraSleeve purchase per sampling event. This method requires the initial purchase of bottom weights, clips, and installation string/rope for each well. In addition, replacement HydraSleeves must be purchased for each sampling event.

7.3 COST ANALYSIS

The following sections describe the implementation costs at a standard site using cost data acquired during the field demonstration. In particular, these standard implementation costs are based on specific assumptions and are presented for both shallow and deep wells.

As seenin Figures 8 and 9 below, low flow standard is the most expensive groundwater monitoring technology that was analyzed. In assessing the long-term total monitoring cost at a site (10 years, two events/year), the following represents the total cost from least to most expensive for shallow wells: low flow alternative (small volume), active no purge (HydraSleeve), passive no purge (SNAP), low flow alternative (large volume), and low-flow standard.

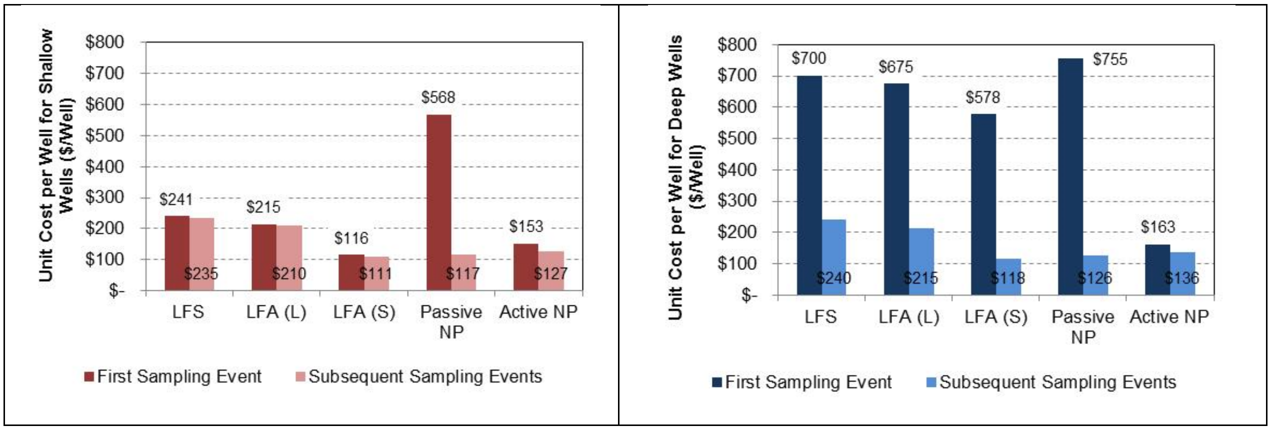


Figure 8 Unit cost per well in shallow (left panel) and deep (right panel) wells
Costs for the first event (darker shade) and subsequent events (lighter shade) are also presented.

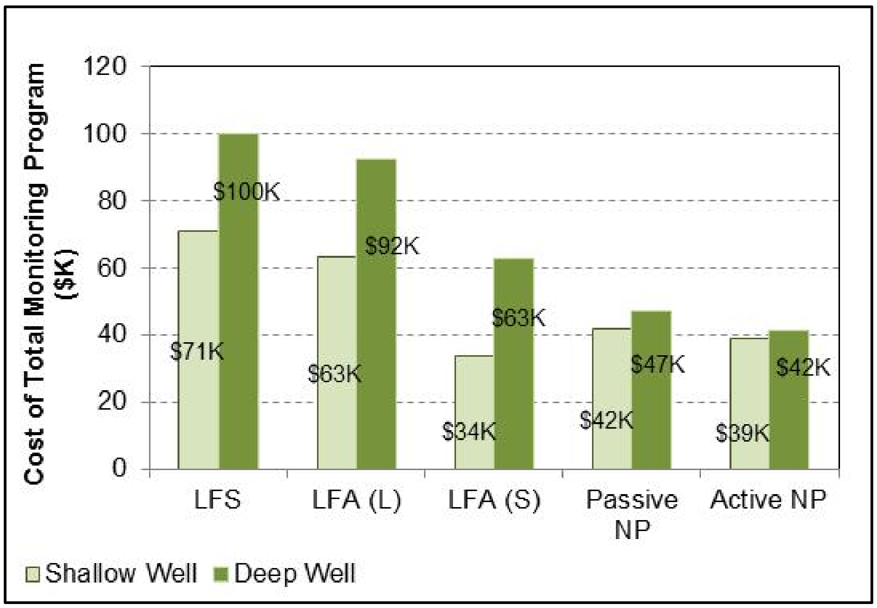


Figure 9 Cost of total monitoring program
10 years, semi-annual sampling, 15 wells, in $K for shallow (lighter shade) and deep (darker shade) wells. LFS = low flow standard, LFA (L) = low flow alternative large volume purge, LFA (S)= low flow alternative small volume purge,
passive no purge = SNAP samplers, active no purge = HydraSleeve.

Additionally, the labor hours required for sampling per well at the field site varied significantly across the sampling methods. Assuming an 8-hour field day, this translated to a varying number of wells that can be sampled in one mobilization, as well as total labor cost per well per mobilization.

As seen in Table 12 below, the labor cost per well in one mobilization associated with applying sampling methods are as follows in increasing order: low-flow alternative, small volume ($90/well)/HydraSleeve ($90/well), SNAP samplers ($90/well), low-flow alternative, large volume ($180/well) and low flow standard ($220/well). Note that these are higher-end estimates of number of wells that can be sampled in one day as they do not include time to travel between sampling wells at large sites, and minimal downtime during the 8-hour day (i.e., 1 hour break).

Table 11 Summary of sampling time in the field and subsequent labor costs

Sampling Method	Approximate Time per Well (hours)	Estimated Number of Wells Sampled in One Field Day	Labor Cost per Well per Mobilization	Labor Cost per Well Ratio Compared to Low Flow Standard
Low flow standard	0.9	8	$220	1.0
Low flow alternative (small volume)	0.4	20	$90	0.4
Low flow alternative (large volume)	0.8	10	$180	0.8
SNAP samplers (passive no purge)	0.4	20	$90	0.4
HydraSleeve (active no purge)	0.4	20	$90	0.4

Notes:
1. Assumes one field mobilization is an 8-hour day, not including travel time to site or travel between sampling wells at site.
2. Labor cost for two field personnel, approx. $170/hour rate total
3. Approximate time per well based on GSI Environmental field program experience and includes time for installation of each sampling method equipment.

8.0 IMPLEMENTATION ISSUES

This study looked at two types of active sampling and two types of passive sampling methods. All sampling methods are mature technologies, with extensive peer-reviewed literature and guidance for the low flow purging method, peer-reviewed literature, ESTCP studies, and an ASTM International (ASTM) Standard for the passive no-purge method, as well as regulatory acceptance for all methods at a variety of sites in recent years. Guidance includes:

- Groundwater sampling protocols are covered in ASTM D4448-01. (ASTM, 2013)
- Guidelines for active sampling, both for the constant volume purge and purge to parameter stability, can be found in documents such as USEPA SOPs. The ASTM Standard that applies to purge sampling is ASTM D6452-99(2012)e1 (ASTM, 2012).
- Guidelines for active and passive no-purge sampling, both for the HydraSleeve and SNAP samplers can be found in documents such as the Interstate Technology and Regulatory Council's (ITRC) 2007 report.
- The ASTM Standard that applies to passive no-purge sampling is D7929-14 (ASTM, 2014).

Thus, all four methods have few end-user concerns, are straightforward to master, and can be easily applied without substantial implementation issues at most sites. Both the no purge sample methods and the alternative (i.e., fixed volume) low flow purge methods were found to be more cost effective than the standard method of low flow purge to parameter stability. The no purge methods result in little to no generation of purge waste and, therefore, may be more strongly favored at sites where management of purge waste is a logistical challenge or is expensive. Sample volume constraints for the no purge methods are the principal implementation concern where certain analyte suites require large water volumes. For those sites, the low flow alternative methods may be more applicable.

Based on the results of our field program, regulatory acceptance of a novel "improved" sampling method will likely not be an issue. However, our project findings do indicate that low flow sampling with a fixed purge volume is less expensive than monitoring purge parameter stability and yield monitoring results of equal quality. There would likely be some regulatory barriers for sites that wanted to switch from purge parameter stability to fixed volume purge. In addition, although no purge sampling methods have been fairly widely accepted, there are still some regulatory barriers for these methods.

The plan for regulatory acceptance of sampling alternatives to low flow sampling with purge parameter stability is as follows:

1) Publication of a journal article presenting the project results;
2) Presentation of the project results at technical conferences; and
3) A comprehensive half day workshop on groundwater sampling variability.

The half-day workshop will include a module on groundwater sampling methods. In addition to presenting the results from the field program, this module will also present results from SERDP

Projects ER-1704 and ER-1705 and other lines of evidence demonstrating that monitoring of purge parameters during low flow sampling does not improve the accuracy or stability of the concentration results. In addition to using this module in the workshop, it may be possible to present it as a webinar to regulatory stakeholders such as the USEPA Groundwater Forum.

9.0 REFERENCES

Air Force Civil Engineer Center, 2006. Long-Term Monitoring Optimization Guide. Air Force Civil Engineer Center, Environmental Restoration Division. November 2006. http://www.afcee.brooks.af.mil/products/rpo/docs/LTM06Guidance1212.pdf

ASTM, 2013. ASTM D4448-01, Standard Guide for Sampling Ground-Water Monitoring Wells, ASTM International, West Conshohocken, PA, 2013, www.astm.org

ASTM International (ASTM), 2012. ASTM D6452-99(2012)e1, Standard Guide for Purging Methods for Wells Used for Groundwater Quality Investigations, ASTM International, West Conshohocken, PA, 2012, www.astm.org.

ASTM, 2014. ASTM D7929-14, Standard Guide for Selection of Passive Techniques for Sampling Groundwater Monitoring Wells, ASTM International, West Conshohocken, PA, www.astm.org.

Britt, S.L, J.M. Martin-Hayden, and M.A. Plummer, 2014. Final Report: An Assessment of Aquifer/Well Flow Dynamics – Identification of Parameters key to Passive Sampling and Application of Downhole Sensor Technologies. SERDP Project No. ER-1704. Issued December 2014.

GeoInsight, 2010. Procedures for Sampling with HydraSleeve. http://www.blm.gov/pgdata/etc/medialib/blm/wy/field-offices/pinedale/papadocs/sampleplan.Par.29949.File.dat/appc.pdf. Accessed January 26, 2015.

Interstate Technology and Regulatory Council (ITRC), 2004. Technical and Regulatory Guidance for Using Polyethylene Diffusion Bag Samplers to Monitor Volatile Organic Compounds in Groundwater. Available from http://www.itrcweb.org/Guidance/ListDocuments?TopicID=17&SubTopicID=27. Accessed December 20, 2012.

ITRC, 2007. Protocol for Use of Five Passive Samplers to Sample for a Variety of Contaminants in Groundwater. Available from http://www.itrcweb.org/Guidance/ListDocuments?TopicID=17&SubTopicID=27. Accessed December 20, 2012.

McHugh, T.E., C.J Newell, D. Adamson, K. Hamel, L. Molofsky, and L. Beckley, 2013. Final Report: Improved Understanding of Sources of Variability in Groundwater Sampling for Long-term Monitoring Programs. SERDP Project No. ER-1705. Issued February 2013.

McHugh, T.E., L.M. Beckley, C.Y. Liu, and C.J. Newell, 2011. Factors Influencing Variability in Groundwater Monitoring Data Sets, Groundwater Monitoring and Remediation, Vol. 31, No. 2.: 92-101. Spring.

National Research Council (NRC), 2013. Alternatives for Managing the Nation's Complex Contaminated Groundwater Sites. National Academies Press: Washington, DC.

U.S. Environmental Protection Agency, 2004. Guidance for Monitoring at Hazardous Waste Sites – Framework for Monitoring Plan Development and Implementation. USEPA Office of Solid Waste and Emergency Response (OSWER) Directive No. 9355.4-28. January.

APPENDIX A

POINTS OF CONTACT

Point of Contact	Organization	Phone Fax E-Mail	Role In Project
Thomas E. McHugh	GSI Environmental, Inc. 2211 Norfolk, Suite 1000, Houston, TX 77098-4054	Phone: 713-522-6300 Fax: 713-522-8010 Email: temchugh@gsi-net.com	PI
Poonam R. Kulkarni	GSI Environmental, Inc. 2211 Norfolk, Suite 1000, Houston, TX 77098-4054	Phone: 713-522-6300 Fax: 713-522-8010 Email: prk@gsi-net.com	Co-PI
Charles J. Newell	GSI Environmental, Inc. 2211 Norfolk, Suite 1000, Houston, TX 77098-4054	Phone: 713-522-6300 Fax: 713-522-8010 Email: cjnewell@gsi-net.com	Co-PI
Sanford L. Britt	ProHydro, Inc. 1011 Fairport Road Fairport, NY 14450	Phone: 585-385-0023 Email: Sandy.Britt@ProHydroInc.com	Co-PI
Andrea Leeson	SERDP & ESTCP 4800 Mark Center Drive, Suite 17D08 Alexandria, VA 22350-3605	Phone: 571-372-6398 Email: andrea.leeson.civ@mail.mil	Environmental Restoration Program Manager