ENVIRONMENTAL RESTORATION

INNOVATIVE SAMPLING METHODS AND DATA ANALYSIS FOR REDUCED LONG-TERM MONITORING COSTS

Introduction

Why do we care about long-term groundwater monitoring?

The Navy manages hundreds of sites with contaminant plumes in groundwater that require long-term monitoring. At most of these sites, monitoring is expected to continue for at least a decade, and in some cases, monitoring will continue into the foreseeable future. Across all sites, the future cost for this long-term monitoring is expected to exceed $500 million.

At these sites, a primary purpose of long-term monitoring is to track contaminant concentrations over time. Long-term monitoring data are commonly used to answer three questions (Figure 1):

- Are contaminant concentrations decreasing over time?
- How quickly are concentrations decreasing?
- When will the clean-up goal be attained?

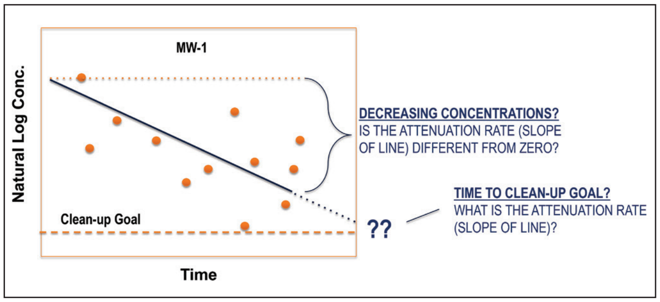

Figure 1. Goals of Long-Term Monitoring

Unfortunately, high levels of event-to-event variability commonly observed in groundwater monitoring results can make it difficult to identify the true long-term concentration trend. Event-to-event variability can cause an overestimation or underestimation of the time to site cleanup or this variability can create misleading patterns that incorrectly suggest increasing contaminant concentrations or a change in remedy effectiveness (Figure 2). Understanding how to manage the effects of event-to-event variability in long-term monitoring records will support more accurate analysis of concentration trends and more cost-effective long-term site management.

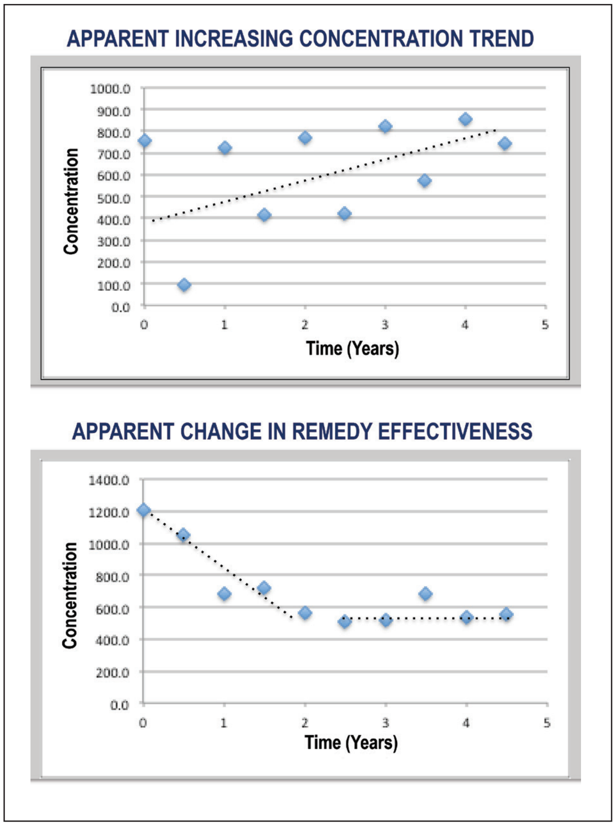

Figure 2. Misleading Concentration Trends Caused by Event-to-Event Variability

Distribution Statement A: Approved for public release; distribution is unlimited.

September 2020

What causes event-to-event variability in long-term monitoring results?

Strategic Environmental Research and Development Program (SERDP) Project ER-1705 provided an improved understanding of the causes of event-to-event variability in groundwater monitoring results. This project used large monitoring datasets from several sites covering over 20 years of monitoring to identify site factors that contribute to event-to-event variability. The study found that event-to-event variability was NOT related to a variety of well characteristics such as average contaminant concentration in wells, aquifer hydraulic conductivity, depth to groundwater below ground surface, magnitude of water table fluctuation, aquifer heterogeneity, or well screen length (McHugh et al., 2011). The study concluded that the majority of event-to-event variability is attributable to the inherent spatial variability in contaminant concentrations present in most groundwater plumes resulting in random concentration variations over time as groundwater with variable contaminant levels flows through the monitoring well (Figure 3).

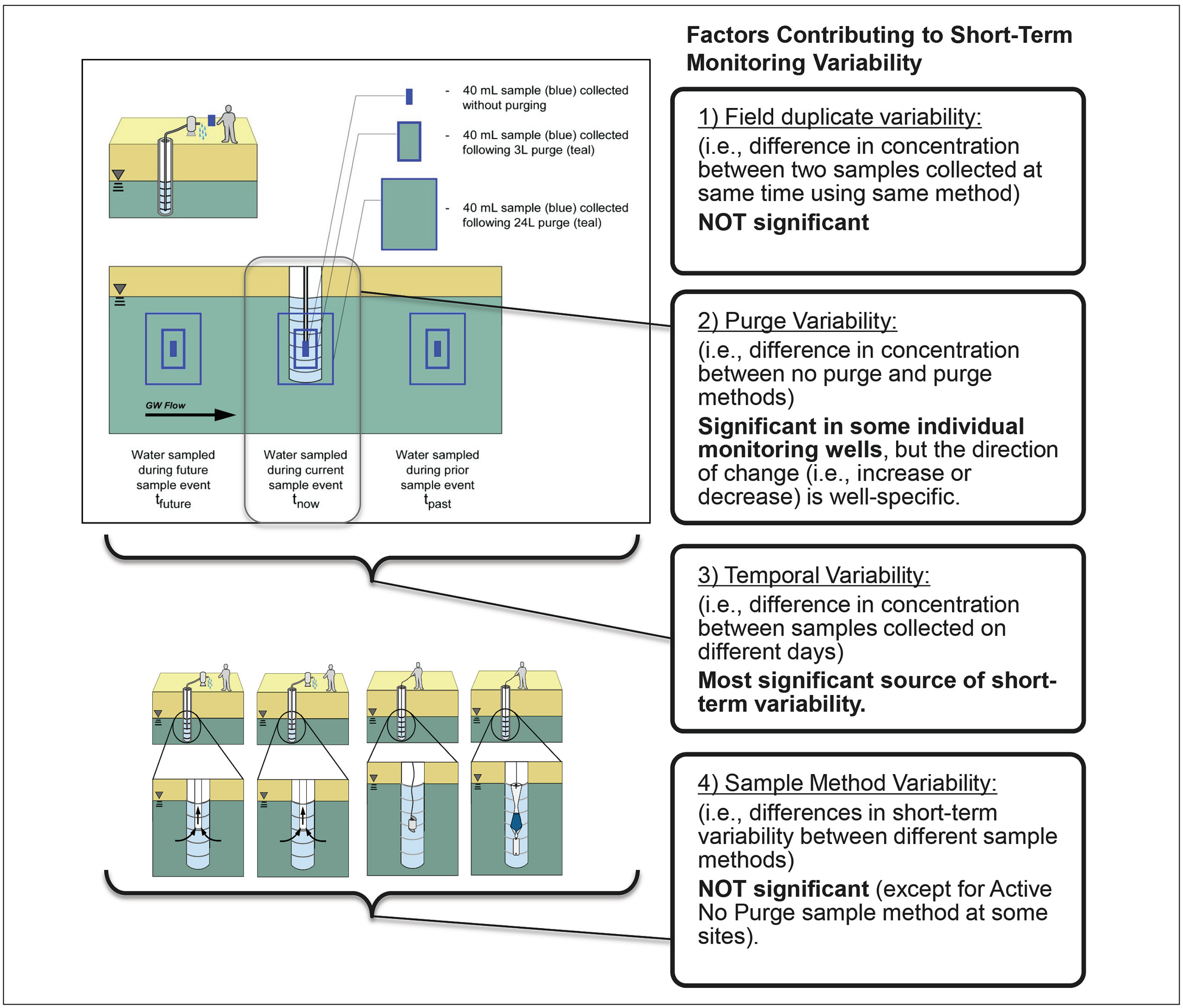

Figure 3. Conceptual Model for Event-to-Event Variability in Groundwater Monitoring Results

Sample methods for long-term monitoring

The groundwater sample method used for long-term monitoring can affect both cost and data quality. Key questions regarding the sample method are:

Are no-purge groundwater sampling methods valid for long-term monitoring?

Yes, a number of different studies have validated various no-purge groundwater sampling methods by comparing results obtained from paired no-purge and purge samples collected from the same monitoring wells (Interstate Technology and Regulatory Council [ITRC], 2006; ITRC, 2007, see Table 1). Today, no-purge sampling is accepted by most state and federal regulatory authorities. It should be noted, however, that not all no-purge samples are suitable for all constituents. The low-density polyethylene passive diffusion samplers are not suitable for metals and other constituents that do not diffuse through polyethylene. In addition, many passive samplers collect a limited volume of water that may not be sufficient for some types of analyses.

Does the sample collection method affect event-to-event variability?

In general, no. The effect of sample collection method on event-to-event variability was evaluated as part of ESTCP Project ER-201209. For this study, three variations of low-flow purge sampling (parameter stability, 0.5 L purge, and 6.0 L purge) and two no-purge methods (e.g., the no-purge device with sealing end caps and the no-purge device with a re-sealing sleeve) were evaluated at two sites and in eight wells at each site. To evaluate the effect of sample method on event-to-event variability, each method was used over six monitoring events and the event-to-event variability was evaluated across these six events. The study confirmed the findings from prior studies showing little or no bias in contaminant concentration between the sample methods. In addition, this new study found that the sample method used had little effect on event-to-event variability in monitoring results (Figure 3). The only exception was that the no-purge device with a re-sealing sleeve was found to yield more variable results for a specific type of monitoring well; specifically, monitoring wells screened more than 10 ft

Table 1. Studies Comparing Purge and No-Purge Sampling Methods

Study	Year(s)	Well Type	Analytes	No-purge Method[a]	Equivalence[b]
SECOR, 1996	1996	Water table only	Petroleum	Bailer	Yes
USGS, 2001	2001	Various, unconsolidated	VOCs	PDB	Yes, for appropriate analytes[c]
Parsons, 2005	2004	Various, unconsolidated	VOC, metals, 1,4-Dioxane	• PDB • RCDM • No-purge device with sealing end caps • RPPS • PSMS • No-purge device with a re-sealing sleeve	Yes (for VOCs) Yes Yes Yes Yes Yes
Parker et al., 2011	2006-2011	Various, bedrock and unconsolidated	VOC, metals, explosives, 1,4-Dioxane	No-purge device with sealing end caps	Yes
Imbrigiotta et al., 2010	2004-2010	Various, unconsolidated	VOC, metals, 1,4-Dioxane	Cellulose diffusion sampler	Yes
Parker et al., 2014	2012-2014	Various, Unconsolidated	VOC, organics	Sorption-based passive sampling	Yes
Britt et al., 2010	2010	Various, bedrock and unconsolidated	VOCs, metals, 1,4-Dioxane	No-purge device with sealing end caps	Yes
Zumbro et al., 2014	2014	Various, unconsolidated	VOC	• No-purge device with sealing end caps • No-purge device with a re-sealing sleeve	Yes Low bias for TCE
McHugh et al., 2015	2015	Various, unconsolidated	VOC	• No-purge device with sealing end caps • No-purge device with a re-sealing sleeve	Yes Low bias and high variability

See Notes and References on page 8.

below the top of the water table. In this type of monitoring well, during some sample events, the no-purge device with a re-sealing sleeve appeared to capture stagnant water from above the well screen (McHugh et al., 2016a).

Based on the study results indicating that most sample methods have little impact on data quality, the ER-201209 project report recommended the selection of sample method based primarily on considerations of cost and implementability. No-purge sample methods are often less expensive than low-flow purge to parameter stability (Figure 4).

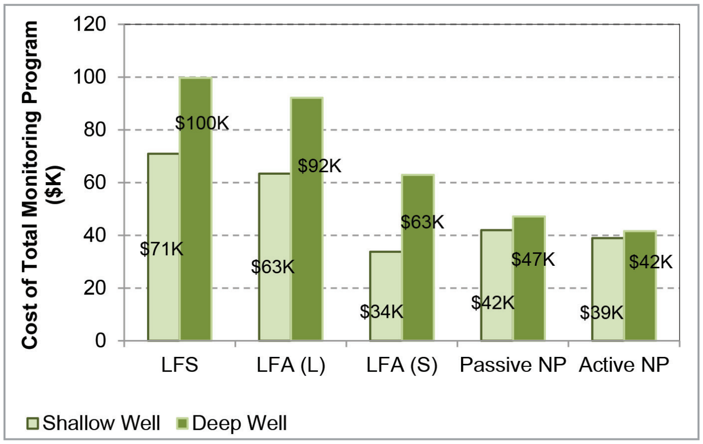

Figure 4. Cost of Total Monitoring Program for Shallow and Deep Wells LFS = Low Flow Standard, LFA (L) = Low Flow Alternative, Large Volume Purge, LFA(S) = Low Flow Alternative, Small Volume Purge, Passive NP = Passive No-Purge (no-purge device with sealing end caps), Active NP = Active No-Purge (no-purge device with a re-sealing sleeve).

Tools for evaluation of monitoring results and monitoring optimization

For a cost effective-long-term monitoring program, it is important to understand how much monitoring data are needed to satisfy the long-term monitoring goals.

How much data are needed to accurately evaluate remedy effectiveness?

Event-to-event variability in groundwater monitoring results increases the uncertainty regarding whether contaminant concentrations are decreasing over time. In addition, this variability reduces the accuracy of the long-term attenuation rate. As part of ESTCP Project ER-201209, monitoring data from 20 real long-term monitoring sites were used in a statistical analysis to determine data needs.

How much monitoring data are needed to determine a site's long-term source attenuation rate with a defined level of accuracy or confidence (Table 2)?

Key points from this evaluation are:

- At many sites, it takes longer than expected to obtain accurate concentration trends. It is important for project managers to recognize that apparent trends characterized using too little data can be misleading and may result in inappropriate management decisions.
- When evaluating natural attenuation, there are often situations where the project manager can be confident that contaminant concentrations are decreasing but highly uncertain as to when numerical cleanup goals will be attained.
- For sites with slow attenuation rates, it may be difficult to prove with statistical confidence that contaminant concentrations are decreasing.

Time versus money: what is the trade-off between monitoring frequency and monitoring duration?

Less frequent groundwater monitoring (e.g., semi-annual monitoring instead of quarterly monitoring) is inherently more cost effective (i.e., less monitoring reduces the annual monitoring costs). However, reduced monitoring will also increase the amount of time required to characterize the

Table 2. Monitoring Data Required to Determine the Long-Term Attenuation Rate

Accuracy/Confidence Goal	Years of Quarterly Monitoring Required		
	Best Site	Median Site	Worst Site
Medium Confidence: Statistically-significant decreasing concentration trend (p<0.1) for 80% of monitoring wells.	2.8 years	7.3 years	30 years
Medium Accuracy: Determine the long-term attenuation rate with an accuracy (i.e., 95% confidence interval) of +/- 50% or +/- 0.1 yr-1 (whichever is larger) for 80% of monitoring wells.	4 years	7.4 years	14.5 years

Table 3. Trade-Off between Monitoring Frequency and Monitoring Time

Monitoring Frequency	Relative Time Required to Characterize Long-Term Trend	Relative Cost to Characterize Long-Term Trend
Weekly	60% less time	Five times the cost
Monthly	33% less time	Twice the cost
Quarterly	**Baseline Time**	**Baseline Cost**
Semi-Annual	25% more time	40% lower cost
Annual	56% more time	60% lower cost
Every 2 Years	Twice the time	75% lower cost
Every 5 Years	Three times as long	85% lower cost

Note: Relative cost is the same as the relative total number of monitoring events required (i.e., based on the assumption that cost is proportional to number of monitoring events).

long-term concentration trend. Although the absolute amount of time required to characterize the long-term attenuation rate depends on both the event-to-event variability and the true contaminant attenuation rate, the trade-off between monitoring frequency and monitoring time is independent of these parameters. The relative trade-off between monitoring frequency and time required to characterize the long-term trend can be quantified through a mathematical analysis of statistical calculations used to evaluate the long-term trend (McHugh et al., 2016b). The relationship between monitoring frequency and monitoring duration is summarized in Table 3.

For example, a site that required four years of quarterly monitoring to characterize the long-term attenuation rate would require five years (= 4 x 1.25) of semi-annual monitoring to characterize the long-term trend with the same level of accuracy. Four years of quarterly monitoring is 16 total monitoring events while five years of semi-annual monitoring is 10 total monitoring events. Therefore, the relative cost of the semi-annual monitoring program would be 60% (10/16) of the cost of the quarterly monitoring program. This is equivalent to saying that the total semi-annual monitoring cost is 40% lower than the total quarterly monitoring cost (Table 3). This trade-off is the same for all long-term monitoring sites. A project manager can use the trade-off between monitoring frequency and monitoring time to select an optimal monitoring frequency.

One important finding from this analysis is that more monitoring data significantly increase the accuracy of the estimated long-term trend. For example, an analysis using eight years of quarterly results (32 samples) will be twice as accurate as an analysis using five years of quarterly data (20 samples). Because of this, when evaluating concentration trends, the project manager should include all available monitoring data. Older monitoring results should be excluded only where there is a specific reason to believe that these older results are not representative of current trends (e.g., monitoring data from before implementation of a new site remedy).

Rules of thumb for monitoring optimization

For long-term monitoring sites, monitoring optimization is an important tool for controlling costs while ensuring that the monitoring program will satisfy the long-term monitoring goals. The following simple rules of thumb can be used to quickly evaluate the likely impact of a more formal monitoring optimization analysis.

Should I change my chemical monitoring parameters?

Generally, no. A consistent set of parameters over time makes it easier to track remedy effectiveness. Reducing the monitoring parameter list rarely results in meaningful cost savings unless an entire chemical class is eliminated (e.g., all semivolatile organic compounds [SVOCs] are dropped, eliminating the need to run the United States Environmental Protection Agency [USEPA] Method 8270). Consider adding parameters if needed to evaluate risk (e.g., emerging contaminants) or monitoring remedy effectiveness (e.g., geochemical parameters).

Should I change my monitoring network?

Sometimes. Some monitoring wells originally installed for plume delineation may not be important for long-term monitoring. When deciding to add or drop wells, consider whether the existing network provides good spatial coverage of the plume. Consider removing wells that are clearly outside the plume (and not serving as sentinel wells) and wells that are providing redundant monitoring results (i.e., a well located close to another well and yielding similar monitoring results). Consider adding wells, if needed, to monitor plume stability or possible migration towards critical receptors.

How frequently should I monitor?

Use semi-annual monitoring as the default starting point and adjust based on site factors. Since the early 2000s, most long-term monitoring sites have moved from quarterly monitoring to semi-annual or less frequent monitoring. For example, by 2015, over 75% of monitoring wells in California were monitored semi-annually or less frequently (Figure 5). Site factors that may support more frequent monitoring include higher risk of a receptor impact, an expanding plume, or operation of a new or high cost site remedy. Site factors that may support less frequent monitoring include low risk of receptor impact, a stable or shrinking plume, or a passive or low cost remedy that has been in place for several years.

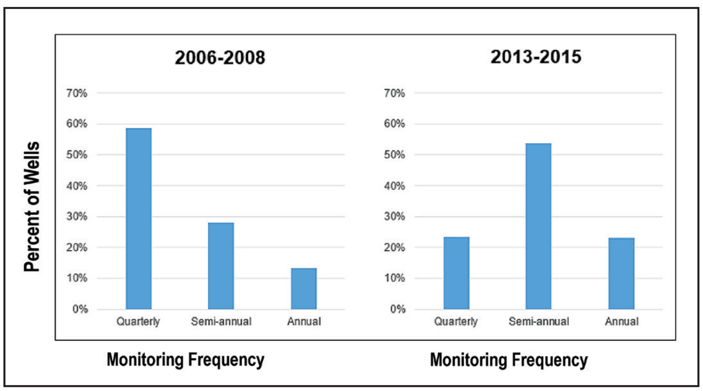

Figure 5. Changes in Monitoring Frequency for Monitoring Wells in California

Should I sample all wells with the same frequency?

Sometimes. For sites with smaller monitoring networks (i.e., with tens of wells), there is likely to be little cost savings associated with sampling some wells at a lower frequency while maintaining a higher frequency at other wells because, at these smaller sites, mobilization and other fixed costs are a more significant part of overall site monitoring costs. At sites with larger monitoring networks (i.e., with hundreds of wells), site-specific considerations may justify different monitoring frequencies for different wells.

Should I change my groundwater sampling method?

Sometimes. For most sites, no-purge / passive sampling methods are less expensive than conventional low-flow purge to parameter stability. However, switching from one sample method to another may involve some upfront costs for new equipment or regulatory approval. Avoid frequently changing the sample collection method and consider the pay-back time for upfront costs.

Should I consider regulatory site closure?

Yes. As part of monitoring optimization, consider whether regulatory site closure is appropriate. For sites where the existing monitoring record demonstrates a stable or shrinking plume and protected receptors, consider whether the applicable regulatory framework supports closure of low risk sites even if numerical objectives have not been attained.

Where can I find the Monitoring Optimization and Trend Analysis Toolkit?

The Monitoring Optimization and Trend Analysis Toolkit is an Excel-based spreadsheet tool developed as part of ESTCP Project ER-201209. This tool helps the site manager analyze existing results from a long-term monitoring program to evaluate:

Monitoring Variability

Question 1: Based on historical monitoring results, when will this site meet the groundwater clean-up goal?

Question 2: Do any wells appear to be attenuating more slowly than the source as a whole?

Monitoring Optimization

Question 1: How much monitoring data are needed to determine a site's long-term source attenuation rate with a defined level of accuracy or confidence?

Question 2: What are the trade-offs between monitoring frequency and time required for trend identification?

The Monitoring Optimization and Trend Analysis Toolkit can be found at the link below:
https://www.serdp-estcp.org/Program-Areas/Environmental-Restoration/Contaminated-Groundwater/Monitoring/ER-201209/(language)/eng-US

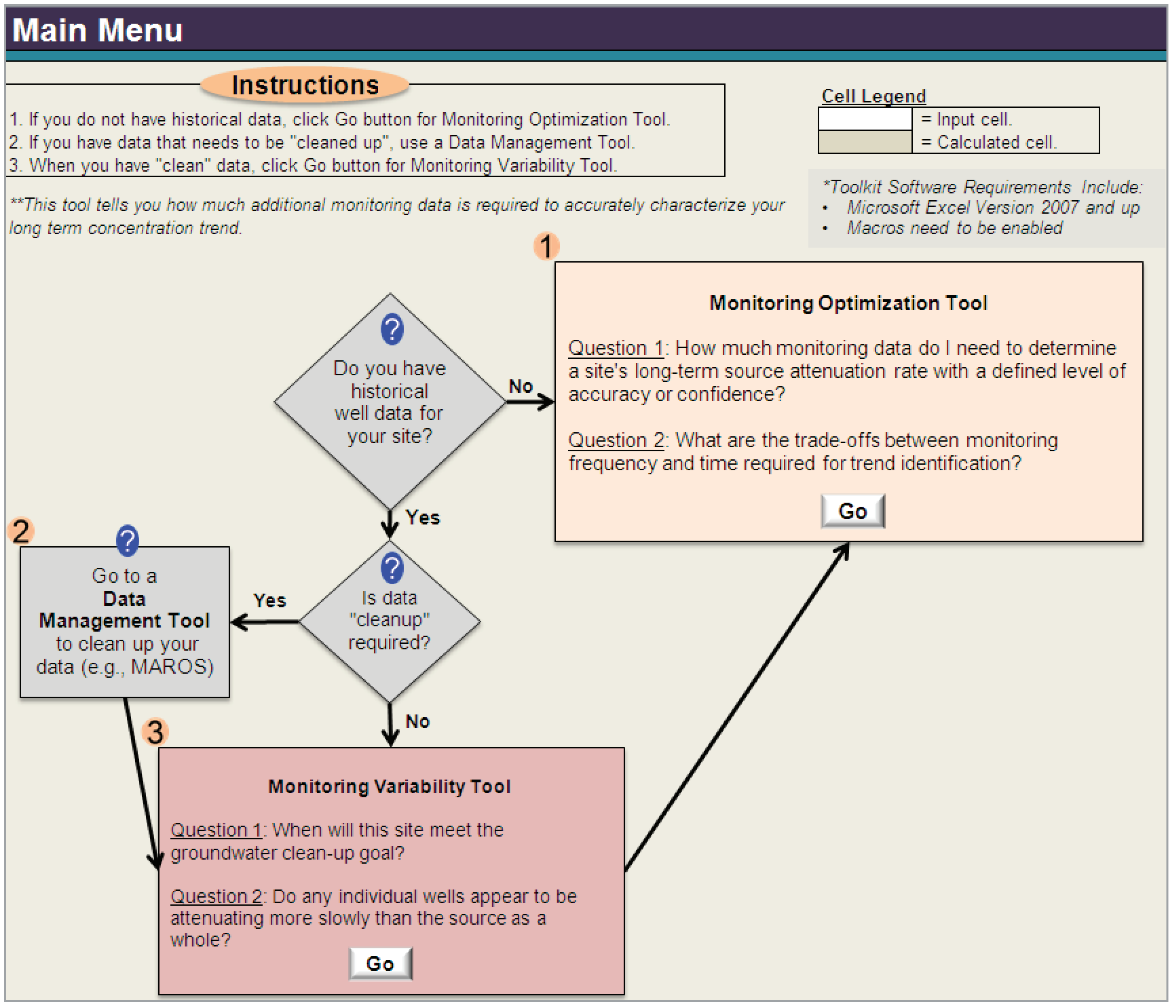

Figure 6. Screenshot from Monitoring Optimization and Trend Analysis Toolkit

Resources

Battelle. 2010. Department of the Navy Guidance for Planning and Optimizing Monitoring Strategies. UG-2081-ENV Rev. 1, November.

Environmental Protection Agency (USEPA). 2000. Environmental Technology Verification Report, Groundwater Sampling Technologies. EPA/600/R-00/091.

Interstate Technology and Regulatory Council (ITRC). 2006. Technology Overview of Passive Sampler Technologies. DSP-4.

Interstate Technology and Regulatory Council (ITRC). 2007. Protocol for Use of Five Passive Samplers to Sample for a Variety of Contaminants in Groundwater. DSP-5.

McHugh, T., L. Beckley, C. Liu, and C. Newell. 2011. "Factors Influencing Variability in Groundwater Monitoring Data Sets", *Ground Water Monitoring and Remediation*, Volume 31, Issue 2, pages 92–101.

McHugh, T., P. Kulkarni, L. Beckley, C. Newell, and M. Zumbro. 2016a. "Negative Bias and Increased Variability in VOC Concentrations Using the HydraSleeve in Monitoring Wells", *Groundwater Monitoring and Remediation*, Volume 36, Issue 1, Pages 79–87.

McHugh, T., P. Kulkarni, and C. Newell. 2016b. "Time vs. Money: A Quantitative Evaluation of Monitoring Frequency vs. Monitoring Duration", *Groundwater*, Volume 54, Issue 5, Pages 692–698.

McHugh, T. 2016c. ESTCP Project ER-201209: Methods for Minimization and Management of Variability in Long-Term Groundwater Monitoring Results. Strategic Environmental Research and Development Program (SERDP)/Environmental Security Technology Certification Program (ESTCP).

United States Army Corps of Engineers (USACE) and Air Force Center for Environmental Excellence (AFCEE). 2005. Results Report for the Demonstration of No-Purge Groundwater Sampling Devices at Former McClellan Air Force Base, California.

Notes for Sampling Studies

a) PDB: polyethylene diffusion bag; RCDM: Regenerated cellulose diffusion membrane sampler; RPPS: rigid porous polyethylene sampler; PSMS: polysulfone membrane sampler.

b) Equivalence is defined as no statistically-significant difference between the purge and no-purge method. For VOCs, if concentrations were higher with the no-purge method, this is also considered conservatively equivalent.

c) Water-filled PDB samplers are not appropriate for all compounds. The samplers are not suitable for inorganic ions and have a limited applicability for non-VOCs and for some VOCs. For example, although methyl-tert-butyl ether and acetone and most semivolatile compounds are transmitted through the polyethylene bag, laboratory tests have shown that the resulting concentrations were lower than in ambient water (USGS, 2001).

References for Sampling Studies

SECOR International, Inc. 1996. Final Report: The California Groundwater Purging Study for Petroleum Hydrocarbons. Prepared for the Western States Petroleum Association, October 28.

United States Geological Service (USGS). 2001. *User's Guide for Polyethylene-Based Passive Diffusion Bag Samplers to Obtain Volatile Organic Compound Concentrations in Wells.* Water Resources Investigation Report 01-4060, USGS, Columbia, SC.

Parsons. 2005. Demonstration of No-Purge Groundwater Sampling Devices, McClellan AFB, Sacramento, CA.

Parker, L., N. Mulherin, G. Gooch, T. Hall, C. Scott, K. Gagnon, J. Clausen, W. Major, R. Willey, T. Imbrigiotta, J. Gibs, and D. Gronstal. 2011. ESTCP Project ER-0630: Demonstration/Validation of the Snap Sampler Cost and Performance Final Report.

Imbrigiotta, T., and J. Trotsky. 2010. Demonstration and Validation of a Regenerated Cellulose Dialysis Membrane Diffusion Sampler for Monitoring Ground-Water Quality and Remediation Progress at DoD Sites: Metals, BTEX, and CVOCs, ER-313.

Parker, L., R. Willey, T. McHale, W. Major, T. Hall, R. Bailey, K. Gagnon, and G. Gooch. 2014. ESTCP ER-200921: Demonstration of the AGI Universal Samplers (F.K.A. the GORE® Modules) for Passive Sampling of Groundwater.

Britt, S., B. Parker, and J. Cherry. 2010. "A Downhole Passive Sampling System to Avoid Bias and Error from Groundwater Sample Handling," *Environmental Science & Technology*, 44(13), 4917-4923.

Zumbro, M. 2014. Performance Comparison of No-Purge Samplers for Long-Term Monitoring of a Chlorinated Solvent Plume, Battelle Recalcitrant Compounds Conference, Monterey California, May.

McHugh, T., P. Kulkarni, L. Beckley, C. Newell, and M. Zumbro. 2015. Negative Bias and Increased Variability in VOC Concentrations Using the HydraSleeve in Monitoring Well. Accepted for publication in Groundwater Monitoring and Remediation, September.

ESTS N39430-16-D-1802