User's Guide

Adaptive Long-Term Monitoring at Environmental Restoration Sites

ESTCP Project ER-0629



SEPTEMBER 2009

Karla Harre NAVFAC Engineering Service Center

Tanwir Chaudhry NAVFAC Engineering Service Center

Distribution Statement A: Approved for Public Release, Distribution is Unlimited



Environmental Security Technology Certification Program

	Report Docume		Form Approved OMB No. 0704-0188			
Public reporting burden for the col maintaining the data needed, and c including suggestions for reducing VA 22202-4302. Respondents sho does not display a currently valid (lection of information is estimated to completing and reviewing the collect this burden, to Washington Headqu uld be aware that notwithstanding ar DMB control number.	o average 1 hour per response, inclu ion of information. Send comments arters Services, Directorate for Infor ay other provision of law, no person	ding the time for reviewing inst regarding this burden estimate mation Operations and Reports shall be subject to a penalty for	ructions, searching exis or any other aspect of th s, 1215 Jefferson Davis J failing to comply with	ting data sources, gathering and is collection of information, Highway, Suite 1204, Arlington a collection of information if it	
1. REPORT DATE SEP 2009		3. DATES COVERED 00-00-2009 to 00-00-2009				
4. TITLE AND SUBTITLE			5a. CONTRACT	NUMBER		
Adaptive Long-Te	rm Monitoring at E	nvironmental Resto	ration Sites	5b. GRANT NUM	IBER	
				5c. PROGRAM E	LEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NU	JMBER	
				5e. TASK NUMB	BER	
				5f. WORK UNIT NUMBER		
7. PERFORMING ORGANI Naval Facilities En 23rd Avenue,Port	ZATION NAME(S) AND AE gineering and Expe Hueneme,CA,93043	Center,1000	8. PERFORMING ORGANIZATION REPORT NUMBER			
9. SPONSORING/MONITO	RING AGENCY NAME(S) A	AND ADDRESS(ES)		10. SPONSOR/M	ONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)		
12. DISTRIBUTION/AVAII Approved for publ	LABILITY STATEMENT ic release; distributi	ion unlimited				
13. SUPPLEMENTARY NO	OTES					
14. ABSTRACT						
15. SUBJECT TERMS						
16. SECURITY CLASSIFIC	ATION OF:	17. LIMITATION OF	18. NUMBER	19a. NAME OF		
a. REPORT unclassified	b. ABSTRACT unclassified	0F PAGES 83	RESPONSIBLE PERSON			

Standard Form 298 (Rev. 8-98) Prescribed by ANSI Std Z39-18



Table of Contents

Introduction	6
Environmental Security Technology Certification Program (ESTCP)	7
Credits and Thanks	8
Site Suitability	9
LTMO Process Flow	10
GA Basics	11
SampleOptimizer TM	12
Sample Tracker TM	13
Installation	14
Requirements	14
Running the Installer	14
Running the Software	14
Tips and Tricks	15
Tool Tips	15
Help Menu	15
SampleOptimizer TM	16
Getting Started	16
Overview	16
Electronic Data Deliverable (EDD)	17
Report Dates	18
Non-Detects in SampleOntimizer TM	18
EDD Notes	19
Spatial Data Temporal Data and Spatio-Temporal Data	20
Loading Data	20
Data Density Requirements	21
Application Options	25
Settings	21
Model Type	25
Data Transformation	20
Ontimization Domain	20
How Spatio-Temporal / Temporal Analysis Works	27
Sampling Location Constraints	27
Frequency Alignment Settings	20
Facility ID	29
Objective Functions	29
Location Group Assignments	30
Exection Selection	31
Function Darameters	32
Free Colculation	32
Cost Calculation	33
Combining COC Objectives	37
ModelDwilderTM	38
ModelBuilder ¹	39
Model Fitting, visualization, and Analysis	39
	40
Inverse Distance weighting	40
Kriging	41
Model Visualization	43
Mass Metric	46



Mass Flux	46
Uncertainty Visualization	48
GA Settings	51
SampleOptimizer TM Dashboard	52
Pausing	54
Dashboard Settings	54
Plan Details	55
Plume Maps	59
SampleTracker TM	62
Getting Started	62
Tips and Tricks	62
Overview	62
Electronic Data Deliverable (EDD)	63
Mass Metric and Mass Flux Tracking	63
Non-Detects in SampleTracker [™]	63
EDD Notes	63
Managing Historical Data	65
Loading Data	66
Application Options	67
Loading Historical Data	68
Settings	71
Results	73
Index	79
Appendices	80
Appendix A – Installation prerequisites	80
Appendix B – Additional References	81
Websites	81
Books	81
Journal Papers	81
Documents	81
Appendix C – Troubleshooting	82
Appendix D – Technical Support, Training, Sales, and Consulting Information	83
User Forum	83
Individualized Support	83
Group Training	83
Consulting	83
Contact Information	83
Sales & Consulting	83
Technical Support	83
License Support & Billing	83



Table of Figures

Figure 1. Overview of Optimization and Tracking Process	
Figure 2. GA Flowchart (created by Professor Chunmiao Zheng, U. of Alabama)	11
Figure 3. Running the Software (shown in Vista)	14
Figure 4. Help Menu	15
Figure 5. Example SampleOptimizer [™] Data File	17
Figure 6. SampleOptimizer [™] Initial Screen	21
Figure 7. Data Density Warning Dialog	
Figure 8. Spatial Data Loaded	
Figure 9. Temporal Dataset Loaded (with exclusions)	23
Figure 10. Application Options	24
Figure 11. Spatial Data Settings	
Figure 12. Model Type	
Figure 13. Data Transformation	
Figure 14. Sampling Location Constraints	
Figure 15. Frequency Alignment Settings	
Figure 16. Facility ID	
Figure 17. Monitoring Objectives Settings	
Figure 18. Location Group Assignments	
Figure 19. Disabling a COC	
Figure 20. Example Spatio-Temporal Sampling Data	
Figure 21. Interpolator input data	
Figure 22. Error Objective Calculator Settings	
Figure 23. Total Samples Cost Calculator	
Figure 24. Combination Weights	
Figure 25. ModelBuilder [™] (Inverse Distance Weighting)	
Figure 26. ModelBuilder™ (Ordinary Kriging)	
Figure 27. Inverse Distance Weighting Model Parameters	
Figure 28. Ordinary Kriging Details	41
Figure 29. Variogram Parameters	
Figure 30. Kriging Variogram	
Figure 31. ModelBuilder [™] Settings (Inverse Distance Weighting)	
Figure 32. ModelBuilder [™] Settings (Ordinary Kriging)	
Figure 33. ModelBuilder [™] Progress	
Figure 34. Model Visualization (Ordinary Kriging model)	
Figure 35. Mass Metric	
Figure 36. Mass Flux Settings	
Figure 37. Mass Flux Flow Rate Settings	
Figure 38. Mass Flux Results	
Figure 39. Uncertainty Visualization Settings	
Figure 40. Uncertainty Visualization (Ordinary Kriging)	
Figure 41. GA Settings.	
Figure 42. SampleOptimizer [™] Dashboard	
Figure 43. Optimization Paused	
Figure 44. Dashboard Settings	
Figure 45. Plan Details - Comparison (Spatial)	
Figure 46. Plan Details - Individual (Spatial)	
Figure 47. Plan Details - Comparison (Temporal)	
Figure 48. Plan Details - Individual (Temporal)	
	-



Figure 49.	Save Plan Details	57
Figure 50.	Saved Plan Viewing	57
Figure 51.	GEMS Export	58
Figure 52.	Plume Map Comparison (Spatial)	59
Figure 53.	Plume Map Comparison (Temporal, Ordinary Kriging)	60
Figure 54.	Plume Map Saving	61
Figure 55.	Saved Plume Map Viewing	61
Figure 56.	Example SampleTracker [™] Historical Data File	64
Figure 57.	Example SampleTracker [™] Current Data File	64
Figure 58.	SampleTracker [™] Initial Screen	66
Figure 59.	Application Options	67
Figure 60.	Historical Data Settings	68
Figure 61.	Loading Historical Data	68
Figure 62.	Historical Data Loaded	69
Figure 63.	Loading Current Data	69
Figure 64.	Current Data Loaded	70
Figure 65.	Bounds Type	71
Figure 66.	Bounds Calculation Settings	72
Figure 67.	SampleTracker™ (Ready to Compute Results)	73
Figure 68.	In Bounds Results	74
Figure 69.	Out-Of-Bounds Results	75
Figure 70.	Exporting In-Bounds Table	76
Figure 71.	Exporting Out-Of-Bounds Table	76
Figure 72.	Viewing Graph	77
Figure 73.	Exporting Graph (1 of 2)	77
Figure 74.	Exporting Graph (2 of 2)	78
Figure 75.	Java Error Dialog – Could not find the main class.	82



Introduction

Welcome to SampleOptimizerTM and SampleTrackerTM!

*SampleOptimizer*TM represents the latest evolution in long term monitoring optimization (**LTMO**) software. For the first time, the power of true mathematical optimization has been applied to LTMO in an easy-to-use desktop software tool to reduce sampling redundancy. *SampleOptimizer*TM is the culmination and combination of years of research in mathematical optimization, data analysis, and environmental engineering. Experts in the fields of geology, hydrogeology, computer and environmental engineering, geostatistics, contaminant geochemistry, and remedial system optimization have applied their technical and field expertise to develop the sound approaches implemented in the software.

*SampleTracker*TM is an additional software module that reviews new monitoring data against historical data. The software identifies cases where current data deviates from expectations that are based on the historical dataset.

*ModelBuilder*TM is an additional component within the software that is utilized by the *SampleOptimizer*TM module and, in some cases, by the *SampleTracker*TM module. *ModelBuilder*TM has two sections: one for model fitting, visualization, and analysis, and another for visualizing uncertainty.

For your convenience, we have created a **Quick Start Guide** and accompanying tutorial files which can be accessed from the Start Menu along with the rest of the software executables, or from the Help menu in the software.

We hope that you find the software to be highly effective in your task of optimizing the monitoring of environmental data at your site. We welcome all feature suggestions and comments at our website <u>http://www.sampleoptimizer.com</u>.



Environmental Security Technology Certification Program (ESTCP)

As described on its <u>website</u>, ESTCP is a Department of Defense (DoD) program that promotes innovative, cost-effective environmental technologies through demonstration and validation at DoD sites. ESTCP's goal is to demonstrate and validate promising, innovative technologies that target the most urgent environmental needs of the DoD. These technologies provide a return on investment through cost savings and improved efficiency.

In late 2005, ESTCP funded a project to demonstrate and evaluate SampleOptimizerTM and SampleTrackerTM on three DoD sites. The objective of this project was to demonstrate and validate the use of SampleOptimizerTM and SampleTrackerTM for reducing costs and improving effectiveness of LTM through adaptive assessment while achieving remediation goals.

Substantial benefits are expected from the use of *SampleOptimizer*TM and *SampleTracker*TM, including:

- 1. Realizing significant cost savings by eliminating redundant data;
- 2. Enabling new data to be assessed for significant deviations and other features of interest without substantial labor, a benefit that will become even greater as more facilities move into long-term monitoring modes and/or emerging sensor technologies produce larger volumes of data to be analyzed; and
- 3. Providing a framework for effective implementation of an adaptive approach to monitoring, enabling limited resources to be efficiently directed where the greatest benefits are likely to be incurred.

The version number of the software as used in the ESTCP project is 2.0.



Credits and Thanks

Summit Envirosolutions, Inc. would like to extend our gratitude to those who have been instrumental to the development and testing of this software product:

- **Dr. Barbara Minsker**, professor of Civil and Environmental Engineering at the University of Illinois and former president and founder of HMSI, for organizing the software development effort beginning with her research group.
- The National Center for Supercomputing Applications (NCSA), especially Michael Welge, David Goldberg, Loretta Auvil, Colleen Bushell, Lisa Gatzke, and Tom Redman, for creating D2K, a data mining and modeling toolkit, as well as EMO, a multi-objective analysis toolkit. Both of these tools were used to create the initial versions of SampleOptimizerTM.
- The **U.S. Department of Defense**, for funding the ESTCP demonstration and validation of this product.
- **Dennis Beckmann**, BP Remediation Engineering and Technology, for funding and contributing to a large part of the software development and testing effort.
- **Riverglass, Inc.** for administering the projects to develop and apply the initial version of this software to two BP test sites, and for licensing D2K to be used in this product.
- **Dr. Patrick Reed**, professor at Penn State, for developing the concept of using Genetic Algorithms (GA's) to optimize monitoring network design while he was a member of Dr. Minsker's research group.
- **Dr. Meghna Babbar-Sebens**, former member of Dr. Minsker's research group and former consultant to Riverglass, **Peter Groves**, former Riverglass software developer, and **David Clutter**, former NCSA researcher, for developing the software prototype using D2K.
- **Matthew Zavislak**, former member of Dr. Minsker's research group, former consultant to Riverglass and HMSI, and a lead software engineer at Summit Envirosolutions, for developing the prototype D2K-based design into its current state as a desktop application.
- John Dustman, the founder of Summit, for adding his knowledge of industry, environmental applications, data management, and data visualization into the software.
- **Dr. Abhishek Singh**, former member of Dr. Minsker's research group, for his assistance with GA troubleshooting, verification, and tuning.
- **David Tcheng**, NCSA researcher, for his help in utilizing D2K.
- **Dr. Tim Ellsworth**, professor at the University of Illinois, for his contribution to the development of the Ordinary Kriging algorithms used in the software.
- Lastly, many thanks to all the users who have shared their feature ideas and bug reports.



Site Suitability

Before utilizing *SampleOptimizer*TM to optimize the sampling on your site, it is important to review these baseline requirements and recommendations to make sure your site is suitable for use with this software suite. Your site must have:

- Sampling analytical results (data) which can be
 - o Entered into a spreadsheet, or
 - Exported from a database into a spreadsheet
- A minimum of 15 sampling locations (20 recommended).
- A minimum of 4 sampling events of historical data per sampling location (8 recommended) for temporal or spatio-temporal optimization. Only one event is needed for spatial optimization.
- See the <u>SampleOptimizerTM EDD</u> for more information about the data requirements.

It is recommended that *SampleOptimizer*TM be used at sites where potential exists for negotiation of an alternate sampling plan with the relevant regulatory agency. This is not necessary if only *SampleTracker*TM will be used, or if the site is not under regulatory guidance or control.

In order to use *SampleTracker*[™] on your site it must have:

- One or more sampling locations with at least four samples per sampling location.
- In order to enable mass metric tracking with *SampleTracker*TM via *ModelBuilder*TM, your site must meet the minimum data density requirements required by *SampleOptimizer*TM.
- Please see the <u>SampleTrackerTM EDD</u> for more information about the data requirements.



LTMO Process Flow

The following is a flow chart describing the steps involved in LTMO using *SampleOptimizer*[™] and *SampleTracker*[™].



Figure 1. Overview of Optimization and Tracking Process

Many of the above steps will be covered in detail later in this reference; below we provide details for steps not otherwise mentioned in this document.



GA Basics

(Adapted from "Introduction to Genetic Algorithms")

Genetic algorithms (GA's) are inspired by Darwin's theory about evolution.

The GA begins with a randomly-generated set of solutions (in *SampleOptimizer*TM these are called **plans**), called a population. Each solution is composed of chromosomes; in *SampleOptimizer*TM each chromosome represents a sampling decision for a sampling location. For example, a chromosome could say to "Sample MW-01 Semi-Annually".

Solutions from one population are taken and used to form a new population (also called a new **generation**). This is motivated by a hope that the new population will be better than the old one. Solutions which are selected to form new solutions (offspring) are selected according to their fitness - the more suitable they are the more chances they have to reproduce. After selection, new solutions are created from the selected solutions through operations called **mutation** and **crossover**.

After the fitness of every solution in a generation has been evaluated, *SampleOptimizer*TM shows the population on a scatter plot referred to as the **Pareto front** or **tradeoff curve**. This plot shows how well the solutions perform relative to two or more user-selected measures of fitness (objective functions).

This process is automatically repeated until the appropriate number of generations has been completed.



Figure 2. GA Flowchart (created by Professor Chunmiao Zheng, U. of Alabama)



SampleOptimizer™

- Create / Review Conceptual Site Model (CSM)
 - Before sending a set of data to the optimizer, it is necessary to check the site data. For example, with groundwater optimization, look for wells which are very close to each other and have similar screened intervals and make sure that they have similar hydraulic head readings and COC concentration values. If there are large differences in such metrics between two or more sampling locations that are very close to each other, there could be a problem with the data.
 - If the CSM shows that your site has multiple hydraulic regions (for example, different aquifers), each region should be optimized separately.
 - Verify to the extent possible that the information is accurate (SiteID, coordinates, values, correct aquifer)
 - Make sure the data file being prepared for import is compliant with the applicable EDD (see <u>SampleOptimizerTM EDD</u> and <u>SampleTrackerTM EDD</u>)
- Identify Spatial or Temporal Redundancy
 - After running the Optimizer, you should identify several potential new sampling plans from the tradeoff curve results generated by the *SampleOptimizer*TM. We recommend that you select plans which provide a level of interpolation accuracy (low error) acceptable to your regulator, at a lower annual cost than the current plan.
 - Similarly, after running the *ModelBuilder*TM, you should identify any areas with unusually high uncertainty which may be of potential interest for characterization according to the site's monitoring objectives.
- Develop New Sampling Plan With Regulatory Agency
 - Typically a regulator must approve changes to a site's sampling plan. In order to justify a new reduced sampling plan, you will need to provide evidence (such as the plume images generated by the *SampleOptimizer*TM) that reduction or elimination of sampling at certain locations, possibly in combination with installation of new sampling locations or increased sampling frequency at other locations, will maintain or improve the fulfillment of the monitoring program objectives.



SampleTracker[™]

- Results Do Anomalies of Concern exist?
 - The results of running *SampleTracker*[™] may indicate that a COC was "out-ofbounds" at a location. In such a case, it will be incumbent upon the analyst using the software to decide if an out-of-bounds COC is an "Anomaly of Concern" needing further study and possible corrective action.
 - For example, a COC being found below the anticipated range will most likely not be a concern; a COC being found slightly above the anticipated range may not be a concern either.
 - Some possible reasons for anomalies:
 - Real emergencies requiring a change in remediation strategy, such as a breakdown in a containment or treatment system.
 - Bad data, due to errors during collection, faulty laboratory analysis, mislabeling of bottles, errant sample preparation, etc.
 - Real differences between actuality and expectation that do not require action, but do require readjustment of one's CSM or expectation, possibly related to a real change in geochemistry or hydrology.

• Evaluate & Take Corrective Action As Appropriate

• After identifying one or more anomalies of concern, the analyst may need to notify additional parties who can decide if further action is appropriate, and what action will be needed.

• Update ST Historical Data File

- Lastly, at the end of each sampling event and the analysis thereof, the analyst needs to update the *SampleTracker*[™] historical data file.
- All <u>appropriate</u> samples should be added to the historical data file. For example, samples which were found to be bad data should not be added.
- If there are fewer than 8 historical samples for a given COC at a sampling location, it is recommended that new samples be added to the historical data until the requisite 8 historical samples have been gathered.
- It is important that the analyst understands that adding additional samples to the historical data can cause the upper bound to increase over time if the added samples are of a high concentration.
 - A good rule of thumb to use when deciding whether or not to add a sample to the historical data is if a concentration at a sampling location would almost certainly never be a concern for the foreseeable future, then it can be safely added to the historical data otherwise it may be advisable to leave it out.
- During periodic review, it is advisable to check if changes to the CSM should be reflected in the historical data. For example, a value which was once thought to not be appropriate for the historical data could be added into it, and vice-versa, as appropriate for the "new" groundwater understanding.



Installation

Requirements

<u>Computer requirements</u>: Windows XP or Vista, 1 GB RAM (2 GB preferred), 1.5 GHz singlecore processor (3.0 GHz dual-core preferred). The installer must be run from a user login with at least power-user permissions.

To use *SampleOptimizer*TM 2.0, you must have a valid license file (.inst).

Currently the software is only available for Microsoft Windows XP and Vista. Operation under Windows 7 is likely but has not been fully tested. In addition, you must have Sun Java JRE 6.0 or better installed, as well as Microsoft .NET runtime 3.5 or better. Please refer to <u>Appendix A</u> for more information on downloading those tools if you don't already have them installed.

Running the Installer

Simply double-click on the installer file, **SampleOptimizerSetup.msi** when you have it downloaded and have a license file (which has an .inst extension) ready.

Running the Software

Summit Envirosolutions	
🎉 SampleOptimizer & SampleTracker	
搅 Quick Start Manual	
🔁 Reference Manual	
so SampleOptimizer	
<i>ुँ</i> न: SampleTracker	
1 Back	
	f
Start Search	2

Figure 3. Running the Software (shown in Vista)

Upon running either *SampleOptimizer*TM or *SampleTracker*TM for the first time, you will be prompted to input your license file (ends in .inst). You will not be able to proceed unless your license is validated by Summit's validation server. This procedure sends no personally-identifiable information and is conducted over a secure https connection. Please keep this license file in a safe location, as it will be required to re-install the software.

Please note that per the terms of your retail license you will only be able to install the software on the number of machines for which licenses were purchased. Multiple-boot systems require a license for each OS installation.



Tips and Tricks

Tool Tips

If you see an option in the software and are unsure about it, allow your mouse cursor to hover over it for a few seconds and a "Tool tip" will pop up and give you more information about that option.

Help Menu

At any time, you may access the Help Menu from which the user can open this Reference, the Quick Start, or the product web site.



Figure 4. Help Menu



SampleOptimizer™

Getting Started

Before you begin using *SampleOptimizer*TM, you will need a set of historical data to work with. The format currently supported for data input & output is .csv. <u>CSV</u> is a universal, standard, non-proprietary, and royalty-free format supported by nearly all spreadsheet and database software including *Microsoft's Excel*TM & *Access*TM, and *OpenOffice.org's* free *Calc* and *Base*. It is very simple and CSV files can even be easily created by hand using a text editor such as Microsoft's *Notepad*.

Overview

*SampleOptimizer*TM strives to answer the following question:

How should this site be sampled in the future, in order to best fulfill our goals without spending more than is needed?

In this process, *SampleOptimizer*TM looks at a site's past data and makes decisions based on the assumption that overall trends in the past will continue into the future. For LTMO sites, this is very likely to be a valid assumption.

In order to determine if there is a sufficient amount of information about the site, the site's physical characteristics (e.g., hydrogeology) and regulatory constraints, as well as professional engineering judgment must be used together with the sampling data, plume maps, and other information provided by *SampleOptimizer*TM in order to make an overall determination as to the best course of future action at a site.



Electronic Data Deliverable (EDD)

The EDD for *SampleOptimizer*[™] is designed to import site-specific sampling data in a simple cross-tab format. Here is an example.

Date SiteID EastCoordinate		NorthCoordinate	Benzene	Chlorobenzene	
4/12/1995	BL003	222.5	768	4.10	240.00
4/12/1995	OS004	256	720.75	0.90	20.2
4/12/1995	OS003	441	6	1.20	180.30
4/12/1995	OS005	517.25	449	0.01	
3/13/1996	MWSL001	846	727	2.70	8.20

Figure 5. Example SampleOptimizer[™] Data File

The first four columns (Date, SiteID, EastCoordinate, and NorthCoordinate) must have those exact column names in that exact order. Add or remove columns to the right of NorthCoordinate based on the actual number of COC's. COC column names can be anything you wish, but there cannot be more than one column with the exact same name. Also, for ease of display, it's best to keep COC name lengths to a minimum. For example, use "MTBE (ppb)" rather than "Methyl tertiary butyl ether (parts per billion)".

Important!	Each COC's units must be consistent within that COC. Different
	COC's can have different units, however, since they are analyzed
	individually by the interpolation model. For example, in Figure 5,
	benzene could be in $\mu g/l$, while chlorobenzene could be in mg/L.

The coordinate system must remain constant throughout the site's data, and only one sampling location is allowed per EastCoordinate, NorthCoordinate pair. The program does not impose a minimum separation distance between locations as long as they are not numerically identical.

If your data contains two or more samples for a sampling location at a given time, you must decide which value you choose to input to SampleOptimizerTM. Possible choices include the minimum, maximum, or average of the multiple samples. Here are some general recommendations:

- Average regular and field duplicate data values if neither value is an obvious outlier and the samples are equally within the working range of the analytical method.
- Do not average in quality assurance (QA) samples, usually labeled "Lab Replicate" or "DUP". Often, such values are not comparable and should therefore be ignored.
- If both values are non-detects with different reporting limits, use the result with the lower reporting limit (RL).
- If one value is a non-detect with a "typical" RL and the other a value, average half the RL and the value.
- If one value is a non-detect with a high RL and the other a value, take the value.



Report Dates

Sample dates must be arranged into discrete sampling events and labeled with one report date per sampling event, since the software will internally reference the data in quarters. This convention was chosen because sites are usually sampled relative to the quarters in each year.

Please note that only one report date per quarter is allowed and only one sample per COC per location per quarter is allowed. If the input data do not follow these requirements, the software will display an error message explaining the source of the problem, and the data file must be corrected before proceeding.

For reference, the quarter cutoff dates, which are inclusive of the end points, are as follows:

Q1: January 1 to March 31 Q3: July 1 to September 30 Q2: April 1 to June 30Q4: October 1 to December 31

Non-Detects in SampleOptimizer™

- Zero values are not allowed. Instead, non-detect data with typical RLs should be replaced by a numerical substitute value which should be consistent for each COC. While there is no perfect non-detect substitute value, 1/20th RL has been found to work acceptably in many cases.
- If the typical RLs vary from event to event, using a common value such as 1/20th the median of typical RLs across events may be useful.



EDD Notes

- The maximum number of significant digits supported for coordinates and sample values is 15, in the approximate range of -1.79769E308 to 1.79769E308. Additionally, in extreme cases where the site area is astronomically large from a numerical standpoint, the software will show an error message that optimization will not be possible. This should never happen in practical circumstances.
- The supported date formats are as follows:
 - Leading zeros are OK as in 05/05/2007
 - 4-digit and 2-digit years are both supported
 - o 5/15/2007
 - \circ 5/15/2007 1:00:00 PM (Note: the time will be ignored)
 - \circ 5/15/2007 13:00 (Note: the time will be ignored)
 - o 15-May-07
- If a value for a COC is missing, simply leave the appropriate cell blank as in the example above.
- If you are using *Excel* to create the CSV, all columns & rows outside the valid data should have an "Edit > Clear > All" done to them, or else the exported CSV may contain a potentially large number of blank rows and/or columns which will cause errors during the data import process. You can check for excess columns/rows by opening the CSV in a text editor such as Microsoft Windows' *Notepad*.



Spatial Data, Temporal Data, and Spatio-Temporal Data

There are two categories of data which *SampleOptimizer*TM can work with: spatial, and temporal/spatio-temporal. **Spatial** data is a set of COC concentration values for sampling locations at <u>one reporting date</u>. **Temporal** and **spatio-temporal** data contain COC concentration values for sampling locations at <u>multiple reporting dates</u>.

Depending on the available data for the chosen sampling locations at your site, it may be possible to perform either a spatial analysis or a temporal/spatio-temporal analysis.

If a spatial analysis is desired, a representative sampling event must be chosen from the historical data, stored in a properly formatted .csv file, and then input into the optimizer. Alternatively, you may choose to create an artificial sampling event based on the historical data, perhaps averaging locations' historical values or using the last value at each location, in order to create a representative sample value for each sampling location within the newly created artificial sampling event.

If a spatio-temporal or temporal analysis is desired, a properly formatted .csv file containing at least 4 or more representative sampling events must be input into the optimizer. These may be the actual historical data, or, at the discretion of the user, may be artificial sampling events constructed based on the actual historical data.

For further guidance on the construction of artificial sampling events, please contact one of our optimization consultants.



Loading Data

Upon running *SampleOptimizer*TM, you will be presented with the following screen.

If you already have a saved **.site** file, you can load it with the File menu, drag it into the window area, or simply double-click on it.

To start a new site analysis, either click the Load button and select the data file, or simply drag the data file onto the window area. Alternatively, you can right-click on a .csv data file you have prepared, select "Open With", and then choose *SampleOptimizer*TM.

S Untitled Site	x
jile <u>E</u> dit <u>H</u> elp	
III Sampling Data Settings 🔊 SampleOptimizer™ Dashboard	
Data Import	
Load data	

Figure 6. SampleOptimizer[™] Initial Screen



If your data file is not in the correct format, you will receive an error dialog explaining the condition. You can then fix the problem in a spreadsheet editor program or text editor, and try to import it again.

If data density checking is enabled in **Application Options** (by default it is disabled), when you attempt to load a new data file which has low data density, you will see a warning screen which explains the source of the warning as well as options on how to proceed. Please note that if there are multiple sampling locations and/or COC's with low data density, you will see multiple confirmation dialogs unless you choose either the "Yes to all" or "No to all" options.



Figure 7. Data Density Warning Dialog

After successfully loading a data file or opening a saved site file, you will see the Data Summary screen. In Figure 8, a spatial dataset has been loaded.

🗞 * Curr	ent *							x
<u>F</u> ile <u>E</u> dit	<u>H</u> elp							
	oling Data	A	- Ann					
Jan	ipiing <u>D</u> ata	Settings	5 SampleOptin	hizer 'n Dashboard				
-Data Im	port							
						Load data		
_ Data Su	immary							
COC's: 1	1 Number	of valid san	npling locations: 49	Number of sam	ples: 49	Number of sampling events: 1	Sample date range: Dec 4, 2006 - Dec 4, 2006	
Samplin	g data							
	Data	CitaTD	FactCoordinate	Nexth Coordinate	Banana			
		Sileid	EastCoordinate	NorthCoordinate	berizerie			
1	Dec 4, 2006	MW-001	4102238.938	86960.84475	1.0E-5			<u> </u>
2	Dec 4, 2006	MW-002	4102257.357	87055 00974	1.0E-5			
4	Dec 4, 2006	MW-004	4102261.073	87022,23322	1.0E-5			
5	Dec 4, 2006	MW-007	4102303.014	86784.00202	1.0E-5			
6	Dec 4, 2006	MW-008	4102304.946	86923.65733	1.0E-5			
7	Dec 4, 2006	MW-010	4102315.75	87017.9063	1.0E-5			
8	Dec 4, 2006	MW-011	4102322.899	86839.17967	1.0E-5			=
9	Dec 4, 2006	MW-012	4102325.884	87036.3933	1.0E-5			
10	Dec 4, 2006	MW-013	4102328.976	87191.86734	1.0E-5			
11	Dec 4, 2006	MW-016	4102349.809	86885.27015	1.0E-5			
12	Dec 4, 2006	MW-017	4102350.787	86683.26787	1.0E-5			
13	Dec 4, 2006	MW-018	4102353.914	86641.03048	1.0E-5			
14	Dec 4, 2006	MW-020	4102392.5	86592.34936	1.0E-5			
15	Dec 4, 2006	MW-021	4102393.459	87114.0384	12.5			
16	Dec 4, 2006	MW-024	4102406.5	87056.0313	989.0			
1/	Dec 4, 2006	MW-027	4102428.75	86608 5/272	1.05.5			
10	Dec 4, 2006	MW-020	4102435 001	87018 62239	751.0			
20	Dec 4, 2006	MW-031	4102437 75	86972,7266	1.0E-5			
21	Dec 4, 2006	MW-032	4102440.682	87101.3247	1038.0			
22	Dec 4, 2006	MW-034	4102442.5	86987.5	318.0			
23	Dec 4, 2006	MW-040	4102464.5	86979.8125	253.0			
24	Dec 4, 2006	MW-041	4102465.5	86966.5	92.0			
25	Dec 4, 2006	MW-042	4102466.0	87022.1406	1504.0			
26	Dec 4, 2006	MW-049	4102483.0	86961.1641	1.0E-5			+
	Dee 4, 2000	MW 0C0	4100510.0	00047-0004	0.000			

Figure 8. Spatial Data Loaded



If sampling locations, contaminants, or sampling events have been excluded from the analysis because of lack of sufficient data density, their samples will be grayed out. Figure 9 shows an example of a temporal dataset which has been loaded, but with some disabled samples (locations RW-03, RW-08, and RW-11 have been disabled).

💰 * Hi	storical *					_		X			
<u>File</u>	lit <u>H</u> elp										
Sa Sa	ampling <u>D</u> ata	Settings (💮 Sample <u>O</u> ptimize	er™ Dashboard							
-Data	- Data Import										
Data											
					Load dat						
Data	Summary										
COC's	: 1 Number of	valid samplin	a locations: 44	Number of samples	440 Number	of sampling events: 24	Sample date range: Dec 11, 1990 - May 30, 2006				
0003	the data	valia samplin	giocadona, Ti	Number of Sumples.	Ho Hamber	or sumpling events, 21	Sample date range, bee 11, 1550 may 56, 2000				
Samp	ling data										
	Date	SiteID	EastCoordinate	NorthCoordinate	Benzene (µg/l)						
37	79 Nov 7, 2005	MW-042	4102466.0	87022.1406	2616.0						
38	30 Nov 7, 2005	MW-049	4102483.0	86961.1641	1.0E-5						
38	31 Nov 7, 2005	MW-060	4102510.0	86947.2031	204.0						
38	32 Nov 7, 2005	MW-062	4102518.227	86798.28865	38.1						
38	33 Nov 7, 2005	MW-064	4102521.5	86942.9922	10.9						
38	34 Nov 7, 2005	MW-069	4102531.368	87314.87412	1.0E-5						
38	35 Nov 7, 2005	MW-072	4102537.376	86776.86545	1.0E-5						
38	36 Nov 7, 2005	MW-073	4102540.0	86941.5859	5.4						
38	37 Nov 7, 2005	MW-078	4102556.519	87316.66137	1.0E-5						
38	38 Nov 7, 2005	MW-082	4102565.007	86955.71873	1.0E-5						
38	39 Nov 7, 2005	MW-086	4102578.647	86907.09457	1.0E-5						
39	90 Nov 7, 2005	MW-087	4102583.557	86923.3831	1.0E-5						
35	91 Nov 7, 2005	MW-088	4102584.703	8/091.7/159	4496.0						
	32 Nov 7, 2005	MW-001	4102566.462	86036 21014	1.0E-5						
	A Nov 7, 2005	MW-091	4102003.440	86052 44062	1.0E-5						
30	Nov 7, 2005	MW-092	4102660 719	86997.05043	1.0E-5						
30	Nov 7, 2005	MW-099	4102667, 186	87280.01315	1.0E-5						
39	97 Nov 7, 2005	MW-100	4102680.0	87126,76421	1.0E-5						
39	Nov 7, 2005	MW-101	4102717.188	87205.00969	1.0E-5						
39	99 May 30, 2006	MW-001	4102238.938	86960.84475	1.0E-5						
40	00 May 30, 2006	MW-002	4102257.557	86997.06925	1.0E-5						
40	01 May 30, 2006	MW-003	4102259.398	87055.90974	1.0E-5						
40	02 May 30, 2006	MW-004	4102261.073	87022.23322	1.0E-5						
40	03 May 30, 2006	MW-007	4102303.014	86784.00202	1.0E-5						
40	04 May 30, 2006	MW-008	4102304.946	86923.65733	1.0E-5			-			
	M 20 2005	MW 010	4100015-35	07017-0002	1.05.5						

Figure 9. Temporal Dataset Loaded (with exclusions)

Data Density Requirements

For temporal/spatio-temporal analysis, *SampleOptimizer*TM will not include sampling locations in the analysis which have fewer than 4 samples, and for spatial and temporal/spatio-temporal analysis will not include sampling events which have fewer than 15 samples.



Application Options

For advanced users, there are some configuration options (in the **Edit** menu under **Application Options**) which are available to be configured during most phases of the software operation.

So Application Options			x
Revert to Defaults	 		
During data import, warn if the data density for a sampling location or event is considered 'low'.			
Append the run title to exported files' names		V	
Enable population cache for SampleOptimizer™; if enabled, control the cache size	10	≤ 1,000	≤ 2000

Figure 10. Application Options.

- Data input density warning (disabled by default)
 - If enabled, will warn the user when 15 to 19 (inclusive) valid, non-excluded samples are available in a sampling event, or for temporal/spatio-temporal analysis if between 4 and 7 (inclusive) valid, non-excluded samples are available for a sampling location. Normally the software includes such events or locations without warning, but the user can choose to be notified and have the opportunity to exclude such events or locations from the analysis.
- Append run title (enabled by default)
 - Appends the run title to the current date and site name when auto-generating file names for image and data exporting.
- SampleOptimizer[™] population cache (disabled by default)
 - This is an experimental (beta) feature which in some cases has been found to reduce the computational time of the optimization process. However, this will increase memory consumption and may cause other tasks on your computer to slow down severely, especially if your installed memory is low and/or the population size is too large. The program may even crash if it runs out of memory. Therefore, use this feature at your own risk. Note: this feature does not affect model or plume image computation.



Settings

Next, in the **Settings** tab, you will see the following screen when performing a spatial optimization:

👸 * Current *					
<u>F</u> ile <u>E</u> dit <u>H</u> elp					
📰 Sampling Data 💕 Settings 🔗	Sample <u>O</u> ptimizer™ Dashboard				
Model Type Inverse Distance Weighting Ordinary Kriging	Data Transformation Use Transform Logarithmic Quantile	Optimization Domain Spatial Temporal Spatio-temporal Sampling Location Constraints Frequency Alignment Settings	Facility ID		
Monitoring Objectives (Set these before optimizing) Image: Function Selection Image: Fu					
-Model Builder: Model Fitting, Visualization, 4	and Analysis linary Kriging Details	Ordinary Kriging Variogram Parameters Visualize vari zation Settings View Mass Metric View Mass Flux	ogram		
-Model Builder: Uncertainty Visualization Genetic Algorithm (GA) Settings	Uncer <u>t</u> ainty Visualizat	tion Settings View Uncertainty Visualization			

Figure 11. Spatial Data Settings

Please note that when working with a temporal/spatio-temporal dataset, **spatial** domain will be disabled, while **temporal** and **spatio-temporal** domains will be enabled.

This is the screen where you will be able to edit the settings for the optimization, and use **ModelBuilder**TM to optimize the parameters for the model. The defaults will usually be a good place to start, and most or all settings will not have to be changed in most cases.

You should use the **Visualize Model** tool to verify that the plume interpolation is acceptable before performing an optimization. If you want to increase the resolution of the images generated, see <u>ModelBuilderTM Settings</u> and increase the setting for **# of vertical slices**.



Model Type

Two types of models are featured in *ModelBuilder*TM: **Inverse Distance Weighting** and **Ordinary Kriging**. Details and configurable parameters for each model are given in the "Model Parameters" dialog. If you need additional information on these models, please consult a geostatistical reference book.

Model Type
Inverse Distance Weighting
Ordinary Kriging

Figure 12. Model Type

In general, while Ordinary Kriging is more computationally-intensive and complex than Inverse Distance Weighting, it is also generally considered to be more statistically sound than Inverse Distance Weighting.

Data Transformation

Inverse Distance Weighting and Ordinary Kriging assume that the input data are normally distributed. <u>Many datasets do not follow this assumption</u>. In such cases, a data transformation should be applied. In *SampleOptimizer*TM, transformations are applied to the data automatically as the data goes into and out of the genetic algorithm code. The two transformations offered by *ModelBuilder*TM are **quantile transformation** and **logarithmic transformation**.

The quantile transformation creates statistics on the dataset and substitutes a sample's quantile ranking in the data for its concentration value. The function used is equivalent to the PERCENTRANK function in *Microsoft Excel*TM.

The logarithmic transformation substitutes the natural log of a concentration for its concentration value.

Г	Data Transformation	
	📝 Use Transform	
	Cogarithmic	
	Quantile	

Figure 13. Data Transformation



Optimization Domain

Also see Spatial Data vs. Temporal Data & Spatio-Temporal Data

The optimization domain is not configurable when working with a spatial dataset. However, when working with temporal / spatio-temporal data, you can choose whether you would like to perform a temporal or spatio-temporal optimization.

Temporal optimization means that sampling locations' frequencies will be optimized, but no sampling locations can be turned off. Spatio-temporal optimization, on the other hand, allows sampling locations to be turned off. Location sampling recommendations are based on past performance. For example, if turning off a certain location nevertheless results in sufficiently accurate plume interpolation, that location will be recommended to be turned off.

How Spatio-Temporal / Temporal Analysis Works

ModelBuilder[™] searches for a parameter set (Inverse Distance Weighting) or theoretical variogram (Ordinary Kriging) for each COC which works well for all historical sampling events. The Optimizer searches for monitoring plans which minimize the maximum errors across all sampling events and COC's. Frequencies are recommended based on their past performance. For example, if reducing sampling from semi-annual to annual at a given location nevertheless results in sufficiently accurate plume interpolation, annual sampling will be recommended for that location.

Since spatio-temporal optimization uses data from multiple sampling events and must cover multiple plume configurations across all the analyzed sampling events, it is typically produces more conservative results than a spatial analysis which only utilizes one sampling event. This effect can be minimized by configuring the objective functions to allow higher error for locations in the plume interior.



Sampling Location Constraints

Additionally, you can configure the **Sampling Location Constraints** for the optimization by clicking **Sampling Location Constraints** in the **Optimization Domain** area of the **Settings** tab.

Below are examples of the dialog box which will pop-up when **Sampling Location Constraints** is clicked, depending on which domain is selected.

🍰 Sampl	ing Location Constra	ints 💶 🗖 🗙	so Sam	pling Location Constraints	s X		🗞 Sampl	ing Location Constraints	
	Revert to Defaults			Revert to Defaults			Revert to Defaults		·
	Maximum Sampling	Minimum Sampling		Maximum Sampling	Minimum Sampling			Maximum Sampling	Minimum Sampling
MW-001	On 👻	Off 👻	MW-001	Quarterly 👻	Every 5 years 👻	Ш	MW-001	Quarterly 👻	Off 👻
MW-002	On 👻	Off 👻	MW-002	Quarterly 👻	Every 5 years 👻	Ш	MW-002	Quarterly 👻	Off 👻
MW-003	On 👻	Off 🗸	MW-003	Quarterly 🗸	Every 5 years 👻	Ш	MW-003	Quarterly 🗸	Off 🗸
MW-004	On 👻	Off 🗸	MW-004	Quarterly 🗸	Every 5 years 👻	Ш	MW-004	Quarterly 🗸	Off 🗸
MW-007	On 👻	Off 🗸	MW-007	Quarterly 🗸	Every 5 years 👻		MW-007	Quarterly 🗸	Off 🗸
MW-008	On 👻	Off 🗸	MW-008	Quarterly 🗸	Every 5 years 👻		MW-008	Quarterly 🗸	Off 🗸
MW-010	On 👻	Off 🗸	MW-010	Quarterly 🗸	Every 5 years 👻	L.	MW-010	Quarterly 🗸	Off 🗸
MW-011	On 👻	Off 👻	MW-011	Quarterly 🗸	Every 5 years 👻	Ш	MW-011	Quarterly 🗸	Off 🗸
MW-012	On 👻	Off 👻	MW-012	Quarterly 👻	Every 5 years 👻	Ш	MW-012	Quarterly 👻	Off 👻
MW-013	On 👻	Off 👻	MW-013	Quarterly 👻	Every 5 years 👻	Ш	MW-013	Quarterly 👻	Off 🔹
MW-016	On 👻	Off ▼ E	MW-016	Quarterly 👻	Every 5 years 👻	Ш	MW-016	Quarterly 👻	Off 👻
MW-017	On 👻	Off 👻	MW-017	Quarterly 👻	Every 5 years 👻 🗏		MW-017	Quarterly 👻	Off ▼ ≡
MW-018	On 👻	Off 👻	MW-018	Quarterly 👻	Every 5 years 👻		MW-018	Quarterly 👻	Off 👻
MW-020	On 👻	Off 👻	MW-020	Quarterly 👻	Every 5 years 👻		MW-020	Quarterly 👻	Off 👻
MW-021	On 👻	Off 👻	MW-021	Quarterly 👻	Every 5 years 👻	Ш	MW-021	Quarterly 🗸	Off 👻
MW-024	On 👻	Off 👻	MW-024	Quarterly 👻	Every 5 years 👻		MW-024	Quarterly 👻	Off 👻
MW-027	On 👻	Off 👻	MW-027	Quarterly 👻	Every 5 years 👻		MW-027	Quarterly 👻	Off 🔹
MW-028	On 👻	Off 👻	MW-028	Quarterly 👻	Every 5 years 👻		MW-028	Quarterly 👻	Off 👻
MW-031	On 👻	Off 👻	MW-031	Quarterly 👻	Every 5 years 👻		MW-031	Quarterly 👻	Off 👻



The minimum frequency is the least frequent sampling rate which will be allowed in the optimization, and the maximum frequency is the most frequent sampling rate which will be allowed in the optimization. For example, minimum frequency is useful when sampling locations have been designated as sentinels by a regulator, and maximum frequency is useful to include human assessment that a sampling location should definitively have its sampling frequency capped at a certain rate.



Frequency Alignment Settings

For best optimization results when using temporal or spatio-temporal optimization, the frequency alignment settings must be properly set. The description in the dialog box below explains how to correctly set this parameter.



Figure 15. Frequency Alignment Settings

Facility ID

The **Facility ID** dialog allows you to store information about the dataset and the optimization process. This information is stored in the site file.

In future updates to the program, the user will be able to include this information in generated reports.



Figure 16. Facility ID



Objective Functions

Objective functions are methods of evaluating a potential sampling plan. In *SampleOptimizer*TM, there are two types of objective functions: accuracy (error) objectives, and cost objectives. By default, the x-axis displays the accuracy objective, and the y-axis displays the cost objective.

The cost objective is the way for the user to specify the optimization objective of minimizing monitoring costs through minimizing the number of sampling locations and/or the sampling frequency. The default cost objective is basic, and lets the user specify a cost per sample for purposes of displaying the overall cost of a sampling plan.

The accuracy (error) objective is the way for the user to specify the optimization objective of minimizing the loss of information which can occur when fewer locations are sampled and/or when sampling frequencies are reduced. Accuracy is based on the similarity between an interpolated value for a sampling location, versus the actual value at that location.

Important! We highly recommended that you configure the settings for the Error Calculator according to the COC's and monitoring objectives of your site. Please see the section Function Parameters for additional information.

Monitoring Objectives (Set these befo	re optimizing)
	Function Selection Location Group Assignments $\sqrt{\frac{x}{y}}$ Function Parameters
	Combine COC Objectives? Maximum Error Across COC's Additive Error Across COC's

Figure 17. Monitoring Objectives Settings



Location Group Assignments

For additional flexibility and power, SampleOptimizerTM allows the user to assign sampling locations to either of two location groups: exterior and interior locations. This categorization enables the user to choose different objective functions and/or objective function parameters for interior and exterior.

You may find this feature to be useful for allowing less interpolation accuracy inside a plume interior than in the locations considered to be external to the plume. For example, you could use the **Cutoff Error Calculator** (discussed later) and set the percentage error allowed for external locations to be greater than that allowed for internal locations.

🍰 Location Grou	ıp Assignments 💶 💷 💌
Re	evert to Defaults
MW-001	Interior 💌
MW-002	Interior 💌
MW-003	Interior 🔻
MW-004	Interior 👻
MW-007	Interior 💌
MW-008	Interior 💌
MW-010	Interior 💌
MW-011	Interior 👻
MW-012	Interior 👻
MW-013	Interior 💌
MW-016	Interior 💌
MW-017	Interior 🔻
MW-018	Interior 🔻
MW-020	Interior 💌
MW-021	Interior 💌
MW-024	Interior 💌
MW-027	Interior 👻
MW-028	Interior 👻
MW-031	Interior 💌
MW-032	Interior 💌
MW-033	Interior 💌
MW-034	Interior 💌
MW-040	Interior 👻

Figure 18. Location Group Assignments



Function Selection

Clicking **Function Selection** opens a window where you can change the objective functions and disable or enable analysis of each COC. By default, all COCs with sufficient data density will be analyzed by the Optimizer. If there is more than one COC with sufficient data density and you wish to exclude one of them from analysis, click the check box below the COC name (see Figure 19 below, showing Chlorobenzene disabled). To change the objective function for a COC or for the cost objective, select an alternate function from the drop-down menu.

So Function Selection (Confirmed changes will be applied immediately)							
	Cost	Benzene (µg/l)	Chlorobenzene (µg/l)				
Include?	Always included						
Interior Function	Total Samples Cost Calculator 👻	Cutoff Error Calculator 🛛 🗸	Cutoff Error Calculator 🚽				
Exterior Function	Exterior Function Total Samples Cost Calculator - Cutoff Error Calculator - Cutoff Error Calculator -						

Figure 19. Disabling a COC

If your site requires a new objective which is not in the drop-down list, please contact our Technical Support department for more information on ordering custom objective function(s).

Function Parameters

For each combination of COC and location group (interior and exterior), there are different options for calculation of cost and error. These include "Cutoff Error Calculator" and "Percentage Error Calculator" for error, and "Total Samples Cost" for cost. To set these function values, click the **Function Parameters** button.



Error Calculation

To evaluate a potential sampling plan with fewer samples than the baseline maximum sampling plan, the optimizer uses interpolation algorithms (either Inverse Distance Weighting or Ordinary Kriging) to predict the sample values at the locations with reduced frequency, using the values at the remaining locations, and compares the interpolated values at the locations of the removed samples to the actual values.

For example, let's say we have the following data where water samples are taken semi-annually at 15 wells in the baseline sampling plan and analyzed for Benzene and MTBE:

Date	SiteID	Benzene	MTBE]	Date	SiteID	Benzene	MTBE
3/1/2006	W01	26.3	12		8/1/2006	W01	1.00E-03	1.00E-03
3/1/2006	W02	51.1	1.00E-03		8/1/2006	W02	1.00E-03	1.00E-03
3/1/2006	W03	14600	440		8/1/2006	W03	16000	560
3/1/2006	W04	41	15		8/1/2006	W04	0.96	1.00E-03
3/1/2006	W05	1.00E-03	1.00E-03		8/1/2006	W05	1.00E-03	1.00E-03
3/1/2006	W06	1.00E-03	1.00E-03		8/1/2006	W06	1.00E-03	1.00E-03
3/1/2006	W07	1.00E-03	1.00E-03		8/1/2006	W07	1.00E-03	1.00E-03
3/1/2006	W08	1.00E-03	1.00E-03		8/1/2006	W08	1.00E-03	1.00E-03
3/1/2006	W09	1.00E-03	1.00E-03		8/1/2006	W09	1.00E-03	1.00E-03
3/1/2006	W10	1.00E-03	1.00E-03		8/1/2006	W10	1.00E-03	1.00E-03
3/1/2006	W11	1.00E-03	1.00E-03		8/1/2006	W11	1.00E-03	1.00E-03
3/1/2006	W12	1.00E-03	1.00E-03		8/1/2006	W12	1.00E-03	1.00E-03
3/1/2006	W13	1.00E-03	1.00E-03		8/1/2006	W13	1.00E-03	1.00E-03
3/1/2006	W14	1.00E-03	1.00E-03		8/1/2006	W14	1.00E-03	1.00E-03
3/1/2006	W15	1.00E-03	1.00E-03		8/1/2006	W15	1.00E-03	1.00E-03
Date	SiteID	Benzene	MTBE		Date	SiteID	Benzene	MTBE
3/6/2007	W01	1.00E-03	1.00E-03		8/9/2007	W01	1.00E-03	1.00E-03
3/6/2007 3/6/2007	W01 W02	1.00E-03 1.00E-03	1.00E-03 1.00E-03		8/9/2007 8/9/2007	W01 W02	1.00E-03 1.00E-03	1.00E-03 1.00E-03
3/6/2007 3/6/2007 3/6/2007	W01 W02 W03	1.00E-03 1.00E-03 9280	1.00E-03 1.00E-03 160		8/9/2007 8/9/2007 8/9/2007	W01 W02 W03	1.00E-03 1.00E-03 2500	1.00E-03 1.00E-03 230
3/6/2007 3/6/2007 3/6/2007 3/6/2007	W01 W02 W03 W04	1.00E-03 1.00E-03 9280 1380	1.00E-03 1.00E-03 160 90		8/9/2007 8/9/2007 8/9/2007 8/9/2007	W01 W02 W03 W04	1.00E-03 1.00E-03 2500 110	1.00E-03 1.00E-03 230 50
3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007	W01 W02 W03 W04 W05	1.00E-03 1.00E-03 9280 1380 1.00E-03	1.00E-03 1.00E-03 160 90 1.00E-03		8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007	W01 W02 W03 W04 W05	1.00E-03 1.00E-03 2500 110 1.00E-03	1.00E-03 1.00E-03 230 50 1.00E-03
3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007	W01 W02 W03 W04 W05 W06	1.00E-03 1.00E-03 9280 1380 1.00E-03 1.00E-03	1.00E-03 1.00E-03 160 90 1.00E-03 1.00E-03		8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007	W01 W02 W03 W04 W05 W06	1.00E-03 1.00E-03 2500 110 1.00E-03 1.00E-03	1.00E-03 1.00E-03 230 50 1.00E-03 1.00E-03
3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007	W01 W02 W03 W04 W05 W06 W07	1.00E-03 1.00E-03 9280 1380 1.00E-03 1.00E-03 1.00E-03	1.00E-03 1.00E-03 160 90 1.00E-03 1.00E-03 1.00E-03		8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007	W01 W02 W03 W04 W05 W06 W07	1.00E-03 1.00E-03 2500 110 1.00E-03 1.00E-03 1.00E-03	1.00E-03 1.00E-03 230 50 1.00E-03 1.00E-03 1.00E-03
3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007	W01 W02 W03 W04 W05 W06 W07 W08	1.00E-03 1.00E-03 9280 1380 1.00E-03 1.00E-03 1.00E-03	1.00E-03 1.00E-03 160 90 1.00E-03 1.00E-03 1.00E-03 1.00E-03		8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007	W01 W02 W03 W04 W05 W06 W07 W08	1.00E-03 1.00E-03 2500 110 1.00E-03 1.00E-03 1.00E-03 1.00E-03	1.00E-03 1.00E-03 230 50 1.00E-03 1.00E-03 1.00E-03 1.00E-03
3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007	W01 W02 W03 W04 W05 W06 W07 W08 W09	1.00E-03 1.00E-03 9280 1380 1.00E-03 1.00E-03 1.00E-03 1.00E-03	1.00E-03 1.00E-03 160 90 1.00E-03 1.00E-03 1.00E-03 1.00E-03		8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007	W01 W02 W03 W04 W05 W06 W07 W08 W09	1.00E-03 1.00E-03 2500 110 1.00E-03 1.00E-03 1.00E-03 1.00E-03	1.00E-03 1.00E-03 230 50 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03
3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007	W01 W02 W03 W04 W05 W06 W07 W08 W09 W10	1.00E-03 1.00E-03 9280 1380 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03	1.00E-03 1.00E-03 160 90 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03		8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007	W01 W02 W03 W04 W05 W06 W07 W08 W09 W10	1.00E-03 1.00E-03 2500 110 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03	1.00E-03 1.00E-03 230 50 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03
3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007	W01 W02 W03 W04 W05 W06 W07 W08 W09 W10 W11	1.00E-03 1.00E-03 9280 1380 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03	1.00E-03 1.00E-03 160 90 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03		8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007	W01 W02 W03 W04 W05 W06 W07 W08 W09 W10 W11	1.00E-03 1.00E-03 2500 110 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03	1.00E-03 1.00E-03 230 50 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03
3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007	W01 W02 W03 W04 W05 W06 W07 W08 W09 W10 W11 W12	1.00E-03 1.00E-03 9280 1380 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03	1.00E-03 1.00E-03 160 90 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03		8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007	W01 W02 W03 W04 W05 W06 W07 W08 W09 W10 W11 W12	1.00E-03 1.00E-03 2500 110 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03	1.00E-03 1.00E-03 50 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03
3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007	W01 W02 W03 W04 W05 W06 W07 W08 W09 W10 W11 W12 W13	1.00E-03 1.00E-03 9280 1380 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 6.9	1.00E-03 1.00E-03 160 90 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 20		8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007	W01 W02 W03 W04 W05 W06 W07 W08 W09 W10 W11 W12 W13	1.00E-03 1.00E-03 2500 110 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 4.8	1.00E-03 1.00E-03 50 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03
3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007	W01 W02 W03 W04 W05 W06 W07 W08 W09 W10 W11 W12 W13 W14	1.00E-03 1.00E-03 9280 1380 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03	1.00E-03 1.00E-03 160 90 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03		8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007 8/9/2007	W01 W02 W03 W04 W05 W06 W07 W08 W09 W10 W11 W12 W13 W14	1.00E-03 1.00E-03 2500 110 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 4.8 1.00E-03	1.00E-03 1.00E-03 230 50 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03

Figure 20. Example Spatio-Temporal Sampling Data Shaded cells show the 12 samples removed by "Plan A".



If we wanted to evaluate a potential plan (let's call it "Plan A") which turned off W06 and W07, and reduced W08 and W09 to annual sampling, the optimizer would perform interpolations to estimate values for the removed samples based on the remaining samples in each sampling period.

Date	SiteID	Benzene	MTBE
3/1/2006	W01	26	12
3/1/2006	W02	51	1.00E-03
3/1/2006	W03	14600	440
3/1/2006	W04	41	15
3/1/2006	W05	1.00E-03	1.00E-03
3/1/2006	W08	1.00E-03	1.00E-03
3/1/2006	W09	1.00E-03	1.00E-03
3/1/2006	W10	1.00E-03	1.00E-03
3/1/2006	W11	1.00E-03	1.00E-03
3/1/2006	W12	1.00E-03	1.00E-03
3/1/2006	W13	1.00E-03	1.00E-03
3/1/2006	W14	1.00E-03	1.00E-03
3/1/2006	W15	1.00E-03	1.00E-03
Date	SiteID	Benzene	MTBE
Date 3/6/2007	SiteID W01	Benzene 1.00E-03	MTBE 1.00E-03
Date 3/6/2007 3/6/2007	SiteID W01 W02	Benzene 1.00E-03 1.00E-03	MTBE 1.00E-03 1.00E-03
Date 3/6/2007 3/6/2007 3/6/2007	SiteID W01 W02 W03	Benzene 1.00E-03 1.00E-03 9280	MTBE 1.00E-03 1.00E-03 160
Date 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007	SiteID W01 W02 W03 W04	Benzene 1.00E-03 1.00E-03 9280 1380	MTBE 1.00E-03 1.00E-03 160 90
Date 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007	SiteID W01 W02 W03 W04	Benzene 1.00E-03 1.00E-03 9280 1380 1.00E-03	MTBE 1.00E-03 1.00E-03 160 90 1.00E-03
Date 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007	SiteID W01 W02 W03 W04 W05 W08	Benzene 1.00E-03 1.00E-03 9280 1380 1.00E-03 1.00E-03	MTBE 1.00E-03 1.00E-03 160 90 1.00E-03 1.00E-03
Date 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007	SiteID W01 W02 W03 W04 W05 W08 W09	Benzene 1.00E-03 1.00E-03 9280 1380 1.00E-03 1.00E-03 1.00E-03 1.00E-03	MTBE 1.00E-03 1.00E-03 160 90 1.00E-03 1.00E-03 1.00E-03
Date 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007	SiteID W01 W02 W03 W04 W05 W08 W09 W10	Benzene 1.00E-03 1.00E-03 9280 1380 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03	MTBE 1.00E-03 1.00E-03 160 90 1.00E-03 1.00E-03 1.00E-03 1.00E-03
Date 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007	SiteID W01 W02 W03 W04 W05 W08 W09 W10 W11	Benzene 1.00E-03 1.00E-03 9280 1380 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03	MTBE 1.00E-03 1.00E-03 160 90 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03
Date 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007 3/6/2007	SiteID W01 W02 W03 W04 W05 W08 W09 W10 W11 W12	Benzene 1.00E-03 1.00E-03 9280 1380 1.00E-03	MTBE 1.00E-03 1.00E-03 160 90 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03
Date 3/6/2007	SiteID W01 W02 W03 W04 W05 W08 W09 W10 W11 W12 W13	Benzene 1.00E-03 1.00E-03 9280 1380 1.00E-03 1.00E-03	MTBE 1.00E-03 1.00E-03 160 90 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 20
Date 3/6/2007	SiteID W01 W02 W03 W04 W05 W08 W09 W10 W11 W12 W13 W14	Benzene 1.00E-03 1.00E-03 9280 1380 1.00E-03	MTBE 1.00E-03 1.00E-03 160 90 1.00E-03 1.00E-03 1.00E-03 1.00E-03 1.00E-03 20 1.00E-03

Date	SiteID	Benzene	MTBE
8/1/2006	W01	1.00E-03	1.00E-03
8/1/2006	W02	1.00E-03	1.00E-03
8/1/2006	W03	16000	560
8/1/2006	W04	1	1.00E-03
8/1/2006	W05	1.00E-03	1.00E-03
8/1/2006	W10	1.00E-03	1.00E-03
8/1/2006	W11	1.00E-03	1.00E-03
8/1/2006	W12	1.00E-03	1.00E-03
8/1/2006	W13	1.00E-03	1.00E-03
8/1/2006	W14	1.00E-03	1.00E-03
8/1/2006	W15	1.00E-03	1.00E-03

Date	SiteID	Benzene	MTBE
8/9/2007	W01	1.00E-03	1.00E-03
8/9/2007	W02	1.00E-03	1.00E-03
8/9/2007	W03	2500	230
8/9/2007	W04	110	50
8/9/2007	W05	1.00E-03	1.00E-03
8/9/2007	W10	1.00E-03	1.00E-03
8/9/2007	W11	1.00E-03	1.00E-03
8/9/2007	W12	1.00E-03	1.00E-03
8/9/2007	W13	4	1.00E-03
8/9/2007	W14	1.00E-03	1.00E-03
8/9/2007	W15	1.00E-03	1.00E-03

Figure 21. Interpolator input data

Some samples have been removed compared to Figure 20.

The interpolator would use the data in Figure 21 to predict what the values for the following 12 missing samples would have been:

- 1. Use the 3/1/2006 data in Figure 21 to predict the values for W06 & W07 on 3/1/2006.
- 2. Use the 8/1/2006 data in Figure 21 to predict the values for W06, W07, W08, and W09 on 8/1/2006.
- 3. Use the 3/6/2007 data in Figure 21 to predict the values for W06 & W07 in 3/6/2007.
- 4. Use the 8/9/2007 data in Figure 21 to predict the values for W06, W07, W08, and W09 on 8/9/2007.



The objective calculator then compares the above 12 interpolated values to the actual values and assigns an overall error value for "Plan A" which is the maximum error for any of the two COC's at any of the 12 removed points in any of the four events, based on the spatial interpolations in each event as described above. The way that each error value is calculated depends on the user's choice in **Function Selection**.

This overall error value is then used by the GA to compare "Plan A" to the other plans. The GA seeks to preserve diversity of the population while also selecting plans which are clearly advantageous compared to other plans. For example, if there are two plans which have the same cost but one is more accurate than the other, the more accurate one will be chosen. If there is one plan with less cost and less accuracy and another with more cost and more accuracy, the GA will seek to keep both, to the extent that there is room in the population.

The overall purpose of Error Calculators is to come up with an objective function value that represents the overall similarity of a new sampling plan to the baseline Max Sampling plan. This can be thought of as an accuracy metric, with accuracy increasing as error decreases.

In most cases we recommend that the Cutoff Error Calculator be used. When accuracy inside a plume interior can be less important than the accuracy outside it, we recommend utilizing the **Location Group Assignments** feature and setting a different percentage error allowed for interior vs. exterior points.

The user must be careful when using the Percentage Error Calculator when very small non-detect values are in the dataset, because a very small error can yield a very high percentage error of a non-detect. For example, an error of just 0.001 for a non-detect value of 0.0001 is a 1000% error, which will almost certainly be much higher than any error for any detected sample values.

Figure 22 shows the settings screen for the Cutoff Error Calculator which is displayed after clicking on **Function Parameters**. Note that we recommend that o, p, q values be chosen such that $o = p^*q$, to insure that there is continuity between errors calculated above and below the cutoff concentration level.
SampleOptimizer[™] & SampleTracker[™] 2.0 Reference Manual





Figure 22. Error Objective Calculator Settings

The **Cutoff Error Calculator** is designed so that error is calculated in a manner which allows more significant percentage deviation, but less significant absolute deviation, between interpolated and actual values in areas of low concentration than in areas of high concentration. This is accomplished as follows:

- The user defines a cutoff concentration (p) for the actual data values that differentiates between low concentrations versus high concentrations, and also defines a value for Acceptable absolute error (o).
- When a low concentration data point is removed (i.e., below the cutoff), error is calculated as the absolute value of the actual value minus the interpolated value, divided by the acceptable absolute error. For example, if the actual value is 5 μ g/l (i.e., below the cutoff concentration of 10 μ g/l) and acceptable absolute error is 1.0, and the difference between the actual and interpolated value is 5 μ g/l, then the error would be |5| / 1 = 5.
- When a high concentration data point is removed (i.e, above the cutoff), error is calculated as the absolute value of the actual value minus the interpolated value, divided by a percentage (q) of the actual value, where q is specified by the user. For example, if the actual value is 100 μ g/l (i.e., above the cutoff concentration of 10 μ g/l) and the percentage input by the user is 10%, and the difference between the actual and interpolated value is 5 μ g/l, then the error would be 5 / (0.10 * 100) = 0.5.



In these examples, the difference between the actual value and the interpolated value was $5 \mu g/l$ in both cases, but in the first case the calculated error is 5.0 whereas in the second case it is only 0.5. This illustrates how the Cutoff Error Calculator appropriately reduces the significance of deviations between actual and interpolated values in lower concentration areas of the plume.

If your dataset does not have any very small values relative to the rest of the data (for example, non-detect values), you can use the Percentage Error Calculator. This calculator is a more basic version of the Cutoff Error Calculator that does not have a cutoff or acceptable absolute error value. You can also use the Percentage Error Calculator for external locations which do not have non-detects.

Cost Calculation

Figure 23 shows the configuration window for the Total Samples Cost Calculator. This is used to calculate per event cost (spatial analysis) or per year cost (spatio-temporal or temporal analysis). The default is \$100 per sample, but if you specify a more accurate number for your site, the tradeoff curve will be more useful for deciding the best cost/accuracy tradeoff for your site.

30 Function Param	neters					
Cost Benzene (µg	Ø					
Interior Function	Total Samples Cost Calculator					
Interior Function Exterior Function	Total Samples This objective function minimizes the number of samples taken, as a surrogate for minimizing overall sampling cost: $Minimize \sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij}$ Where: $x_{ij} = 1$ if location <i>i</i> was sampled at time period <i>j</i> ; 0 otherwise <i>m</i> is the number of sampling locations <i>n</i> is the number of time periods in one year For spatial optimization, n = 1 and there is only one value for <i>j</i> .	Cost Calculator Revert to Defaults				
		Cost Per Sample (\$USD) 0.0 < 100.0				

Figure 23. Total Samples Cost Calculator



Combining COC Objectives

If the dataset you have loaded has more than one valid, enabled COC, by default the software enables COC objective combination. This can help make the optimization results easier for the user to analyze successfully, by reducing the number of objectives to analyze. Instead of looking at two tradeoff curves for a two-COC analysis, the user just has to look at one curve.

When choosing to combine COC objectives while also using the **Cutoff Error Calculator**, it is imperative that the Error Objective Calculator Settings are chosen consistently between COC's in order to insure that one COC does not accidentally have more preference than others. For example, if analyzing samples for Benzene and Toluene, if you set the cutoff concentration for Benzene to 5 ppb, its USEPA MCL, you should also set the Toluene cutoff concentration to its USEPA MCL: 1,000 ppb.

Some care should be taken in deciding how to combine the objectives, as well. The easiest and safest method is to choose the "Maximum Error Across COC's" method, and this is the default. With this approach the error for more than one COC is simply the highest error across all COC's.

Alternatively, you can choose to use the "Additive Error Across COC's" method, and optionally specify the combination weights used (see Figure 24). This method can be useful to specify the relative importance of accuracy for each COC.

\hat{s}_0 Combination We	ights	
Rev	ert to	Defaults
Benzene (µg/l)	0.0	< 0.5 < 1.0
Chlorobenzene (µg/l)	0.0	< 0.5 < 1.0

Figure 24. Combination Weights



ModelBuilder™

*ModelBuilder*TM is used to perform the geo-statistical functions needed during use of *SampleOptimizer*TM. It has two sections: one for model fitting, visualization, and analysis, and another for visualizing uncertainty. It has been designed to be powerful, yet easy to use effectively.

Model Fitting, Visualization, and Analysis

This portion represents the bulk of *ModelBuilder*TM's capabilities. Perhaps most importantly, you can fit model parameters to your data and visualize the results. Both automated and manual model parameter fitting are supported for Ordinary Kriging as well as Inverse Distance Weighting.

Important! It is extremely important to visualize and review the interpolation model used by ModelBuilder[™] before attempting to analyze the Optimization Results. It is best to review and approve the model before running the Optimizer.

Model Builder: Model Fitting, Visualization, and Analysis					
Inverse Distance Weighting Parameters					
Model Visualization Settings Mass Flux Settings					
Visualize Model View Mass Metric View Mass Flux					

Figure 25. ModelBuilderTM (Inverse Distance Weighting)

Model Builder: Model Fitting, Visualization, and Analysis
Ordinary Kriging Details Or <u>d</u> inary Kriging Variogram Parameters
Model Visualization Settings Ass Flux Settings
Visualize Model View Mass Metric View Mass Flux

Figure 26. ModelBuilderTM (Ordinary Kriging)

When you load your data, *SampleOptimizer*TM chooses Ordinary Kriging for the geo-statistical model by default, and *ModelBuilder*TM automatically finds a set of model parameters which should achieve a good visual representation of the data (plume image).



Manual Model Fitting

Inverse Distance Weighting

To manually fit an Inverse Distance Weighting model, adjust the **Distance Weighting Power** in the model parameters and click on **Visualize Model** after closing this input screen to create and view an interpolated image of the COC distribution in a 2-D plan view.



Figure 27. Inverse Distance Weighting Model Parameters



Kriging

Kriging is a more complex model than Inverse Distance Weighting. Its user-configurable parameters are set in **Variogram Parameters**, while additional details about the algorithm can be seen in **Ordinary Kriging Details**.

nordinary Kriging Details	
Ordinary	/ Kriging
Ordinary Kriging creates an experimental variogram which needs a theoretical variogram to be fit to it. When creating the experimental variogram, data points are grouped by distance into lags.	There are no configurable parameters for Ordinary Kriging.
There are three configurable settings for Kriging which can be accessed in Model Builder Settings .	
 The number of lags is determined by "# of Kriging h, γ(h) pairs per experimental variogram lag". Outlier removal is controlled by "Remove Kriging Outlier Sample(s)". The algorithm uses all points as neighbors unless the user selects "Limit Kriging experimental variogram range". This setting can also be helpful sometimes for variogram fitting when the experimental variogram varies wildly outside a certain range. 	

Figure 28. Ordinary Kriging Details

🕉 Ordinary Kriging Variogram Parameters		
The variogram value $\gamma(k)$ is:		Revert to Defaults
$\gamma(h) = \begin{cases} N_0, if \ h=0\\ 0, if \ h\neq 0 \end{cases} + \begin{cases} S_0 \cdot \left(\frac{1.5 \cdot h}{S_1} - 0.5 \cdot \left(\frac{h}{S_1}\right)^3\right), if \ h < S_1\\ S_0, if \ h > S_1 \end{cases} + L_0 \cdot h$	N ₀ (Nugget)	0.0 < 1.0E-12
where h is the lag distance.	S ₀ (Spherical Function Coefficient)	0.0 ≤ 0.3707346342481478
 There are 3 terms which make up the theoretical variogram used in this program: Nugget Term (N₀) Spherical Term (S₀, S₁) Linear Term (L₀) 	S ₁ (Spherical Function Range)	0.0 ≤ 371.9844263593239
The spherical and linear terms can be turned off by setting their naught terms $(S_0, \text{ or } L_0)$ to 0.	L ₀ (Linear Function Slope)	0.0 ≤ 0.0

Figure 29. Variogram Parameters





Figure 30. Kriging Variogram

The overall process for manually fitting a Kriging model is similar to that for Inverse Distance Weighting: you should adjust the parameters until the interpolated plumes are acceptable.

It is highly recommended that you verify the appropriateness of the model parameters by looking at some medium to high resolution plume images of each COC plume.

- Increase the number of vertical slices in the ModelBuilderTM settings to at least 200 (possibly more depending on your site) and then click on **Visualize Model**. (note this value for # vertical slices only pertains to plume visualization in *ModelBuilder*TM a separate value for # vertical slices for visualizing plumes for resulting plans selected from the tradeoff curve in *SampleOptimizer*TM is input in the **Dashboard Settings** window)
- Turn on sampling location labels and look at the sample values vs. the values interpolated by the model, and verify that they look correct.
- A good first step for model validation is to adjust the color range to different maximum values to see the area(s) of the plume which are above certain values, and compare them to the reported sample values.



Model Visualization

You can visualize your data just by clicking on **Visualize Model**, but you can also adjust the **Model Visualization Settings**.

For both Inverse Distance Weighting and Ordinary Kriging, the resolution, or # of vertical slices can be chosen. Note that the number of horizontal slices is chosen by the software based on the chosen # of vertical slices in order for every cell to be a square. The # of border slices determines the amount of blank space around the edge of the visualization. Also, the run title can be specified, and this can optionally be appended (see Application Options) to the autogenerated filenames for exported data and images.

For Ordinary Kriging, additional options are available but these are intended for advanced users who have some knowledge of how Ordinary Kriging models operate.

- Remove outlier sample(s) (disabled by default)
 - If enabled, the algorithm will iteratively remove the highest concentration values until no lag changes by more than the fractional value chosen.
- Number of pairs per lag
 - This parameter influences how many lags will be created in experimental variograms. The number of lags is equal to the total number of h, γ (h) pairs which are created by the data divided by the number of pairs per lag.
- Use means to create lags (enabled by default)
 - If selected, when creating lags, the mean of the experimental variogram points within a lag will be used as the substitute data point for that lag. If unselected, the median of the experimental variogram points will be used instead. NOTE: data which contain a very large number of one value (such as a small non-detect value) should always use the mean to help insure that an experimental variogram can be successfully created.
- Limit variogram range (Disabled by default)
 - If enabled, the experimental variogram range will be limited, and the value entered becomes the distance range fractional cutoff used when creating experimental variograms. This fractional number times the distance between the two monitoring points furthest away from each other is used as the largest distance which will be included in experimental variograms.





Figure 31. ModelBuilderTM Settings (Inverse Distance Weighting)

\widehat{so} Model Visualization Settings			- 0	x
Revert to Defaults				
# of vertical slices for image	50	≤ 200	≤	1000
# of border slices	2	≤ 10	≤	500
Remove Kriging Outlier Sample(s) (Advanced users only)	0.0	< 0.5	<	1.0
# of Kriging h, $\gamma \#$ of (h) pairs per experimental variogram lag (Advanced users only)	1	≤ 100	≤	1000
Use means to create lags (Advanced users only)			/	
Limit Kriging experimental variogram range (Advanced users only)	1.0	≤ 1.0	≤	1.0
Run Title		2009-	01-01	

Figure 32. ModelBuilderTM Settings (Ordinary Kriging)

While *ModelBuilder*TM is working you will see a window similar to that shown in Figure 33.



Figure 33. ModelBuilderTM Progress

Note that displaying the image(s) may take a while to complete. Figure 34 shows an example completed model visualization using Ordinary Kriging.





Figure 34. Model Visualization (Ordinary Kriging model)



Mass Metric

This feature is intended for use in tracking data over time (as is done with *SampleTracker*TM), in order to get a rough estimate of a mass metric which could be tracked in time so that relative mass of contaminants in different sampling events can be compared. However, this functionality is based on spatial interpolations and therefore is included in the *SampleOptimizer*TM menus of the software. To use it, click on Mass Metric to bring up a window similar to that shown in Figure 35. Note that if model visualizations have not yet been generated, they will automatically be generated so that the Mass Metric data can be shown.

Important! Mass Metric is not designed to authoritatively estimate the absolute amount of mass at a site. Rather, it calculates mass per unit volume of aquifer. Therefore, it is useful for comparisons of relative mass between sampling events, but it is not intended to estimate absolute mass within the plume.

츐 Mass M	letric	_ D _ X	Υ	🗞 Mass N	letric	_ D X		
Export			<u>E</u> xport					
Date	SiteID	Benzene		Date	SiteID	Benzene		
Q3 1994	Mass Metric	26,431,605.6648		Q4 2006	Mass Metric	33,025,016.7666		
Q2 1995	Mass Metric	26,528,211.7646	N					
Q3 1995	Mass Metric	20,362,376.8703						
Q4 1995	Mass Metric	10,470,882.562						
Q1 1996	Mass Metric	19,025,106.1845						
Q3 1999	Mass Metric	4,493,458.0582						
Q2 2000	Mass Metric	27,243,824.7731						
Q1 2001	Mass Metric	27,258,132.9315						
Q2 2001	Mass Metric	17,344,840.215						
Q1 2002	Mass Metric	26,356,030.5962						
Q1 2003	Mass Metric	34,673,663.9404						
Q2 2005	Mass Metric	16,233,846.5192						
Q4 2005	Mass Metric	6,474,132.7449						
Q2 2006	Mass Metric	24,800,730.3996						

Figure 35. Mass Metric

To estimate the mass metric for a COC at a given time period at a site, the interpolated concentration value at every cell in the plume map for that COC at that time period is summed, and each cell is approximated to represent the same volume (i.e., incorporating vertical extent and porosity) as every other cell.

To export the mass metric data (which can then be imported into *SampleTracker*TM if desired), select "Export Table to .csv" from the "Export" menu. Then, add the data in this new file to your site's historical and current data files as appropriate.

Mass Flux

This feature is also intended for use with SampleTracker[™], in order to get a rough estimate of the mass which might be flowing through a cross-section of a site. To use it, one must first define a cross-section for the software to analyze. Click **Mass Flux Settings** to bring up a window similar to the one shown in Figure 36.



Important! Mass Flux is not designed to authoritatively estimate the absolute amount of mass flux at a site, because plume thickness and porosity are not accounted for. Similar to Mass Metric, it is useful for comparisons of relative mass flux across a boundary between sampling events, but it is not intended to estimate <u>absolute</u> mass flux in each sampling event.

$\widehat{\mathcal{S}_0}$ Mass Flux Settings (All confirmed changes will be imm	ediately applied)	x
5	Benzene (µg/l)	
	6	
Line Management	4.5 -	Ň
Line Flow Rate East Boundary One Rate North Boundary One Rate	4	
Color Auto 1 range/site Reset	3.5 — 0 0	
Benzene (µg/l) Maximum 5.0 Minimum 0.0		
 Legends, Symbols, & Labels 		
Add Unsampled Locations Toraw Sampling Location Labels	2 0	
Cocation IDS Values Interpolations Font Size Use Default 8 - Save		
Model Visualization Save	1 -	
	0.5 — a	
	0	

Figure 36. Mass Flux Settings

The user can create a boundary consisting of several line segments, each of which can have a unique groundwater flow rate. To create a new line, first click in a plume image to bring up the crosshairs. Then, click again on the place where you would like the cross-section to begin, and click a third time on the place where you would like the cross-section to end. If you would like to cancel the creation of a line when it is already in progress, press the right mouse button to cancel.



After you have created a line, right click on it in the Line Management table and choose "Adjust Flow Rate Settings" to bring up a window similar to the one shown in Figure 37.

Important! Make sure that the flow rates you enter are in the same volumetric units as the site concentration data.

🕉 Flow rate settings 💌						
Line Name Flow rate						
East Boundary	100					
North Boundary	300					

Figure 37. Mass Flux Flow Rate Settings

Mass flux uses similar mathematical and physical principles as Mass Metric. The interpolated value at every location in the cross section is multiplied by the flow rate to get an estimate of the mass flux at each point in the cross section, and the flux at each point is then added together to create the estimate. To view the results, click **View Mass Flux** in the main *ModelBuilder*TM section of the Settings tab. A brief pause may be required to perform the necessary calculations; afterwards, a window similar to the one in Figure 38 will appear.

Free aut						
Export						
East Boundary North Boundary						
Date Benzene (µg/l)						
Q4 2006 89.9892						

Figure 38. Mass Flux Results

Uncertainty Visualization

By clicking on **View Uncertainty Visualization** in *ModelBuilder*TM, you can visualize a metric of uncertainty for your site. This metric shows the user an image which is a combination of variability in the model as well as the data. It can be used along with site-specific knowledge to assist in the siting of new sampling locations.

The uncertainty calculation process uses multiple iterations of a cross validation and visualization generation process, and creates the uncertainty map based on the root-mean-squared (RMS) value of each pixel in the generated visualizations.

Figure 39 shows the Uncertainty Visualization Settings screen. On the left are some more details about how the maps are calculated, and on the right are a set of parameters which can be set by the user to control how the maps are calculated. Hovering the cursor over any of these parameters will bring up a pop-up description.



So Uncertainty Visualization Settings	and I have I had send the I have a local			
Uncertain	ty Settings			
The uncertainty many are calculated with every validation iteration	Revert to Defaults			
loops (either leave-one-out or n-fold) until the overall calculated	# of vertical slices for image	50	≤ 100	≤ 1000
uncertainty of the site map has stabilized to within the stabilization	# of border slices	2	≤ 10	≤ 500
fraction amount.	# of Random Parameters (Advanced users only)	10	≤ 100	
	Remove Kriging Outlier Sample(s) (Advanced users only)	0.0	< 0.5	< 1.0
The maps can either be based on the Standard Deviation, or the	# of Kriging h, γ # of (h) pairs per experimental variogram lag (Advanced users only)	1	≤ 100	≤ 1000
	Use means to create lags (Advanced users only)			
Standard Deviation = $\sqrt{\sum (x_i - \bar{x})^2}$	Limit Kriging experimental variogram range (Advanced users only)	1.0	≤ 1.0	≤ 1.0
	Use n-fold cross-validation (Advanced users only)	2	≤ 10	≤ 49
	Reserve data fraction (Advanced users only)	0.0	< 0.1	< 1.0
$Coefficient of Variation = \frac{Standard Deviation}{2}$	Stabilization fraction (Advanced users only)	0.0	< 0.01	< 1.0
\bar{x}	Optimize for lowest average error (Advanced users only)			
	Calculate Coefficient of Variation or Standard Deviation (Advanced users only)		$\mathbf{\nabla}$	
	Specify the random seed (Advanced users only)	v 0	< 347,182	
	Run Title		2009-04-23]

Figure 39. Uncertainty Visualization Settings

Figure 40 shows an example of uncertainty visualization for a spatial dataset using an Ordinary Kriging model. Note that by default, the coloring range goes from the minimum uncertainty to the 75th percentile of uncertainty. As with plume visualization, this coloration can be adjusted by utilizing the "Color" control in the window.

The uncertainty visualization shows areas of the plume map with the highest root-mean-square error in the estimated concentrations. Note that large differences in concentrations in two nearby wells can lead to high cross-validation errors that may not be a major concern if the plume is thought to be stable and well characterized. If there are locations with high errors where accurate concentration estimates are particularly important (e.g., close to down-gradient exposure locations), these locations may be candidates for additional sampling activities.





Figure 40. Uncertainty Visualization (Ordinary Kriging)

The Uncertainty calculation process for Ordinary Kriging may take from 15 minutes up to several hours to complete depending on the number of sampling locations, number of COC's, number of time periods (spatio-temporal or temporal optimization only), and the speed of your computer's processor. Inverse Distance Weighting is usually much faster.

Please note that the time listed to completion is the *worst-case* completion time for the calculations to complete. Calculation continues until the change in standard deviation of the uncertainty maps is within the **Stabilization %** for two consecutive iterations, which usually occurs before the maximum number of parameter optimization iterations is completed.

Lastly, keep in mind that when the number of sampling locations is large (over 40 or so), Uncertainty Visualization can take an extremely long amount of time to complete. In such situations, we recommend checking the option for "use n-fold cross validation" with n = 3. This is a less rigorous approach but usually yields appropriate results.



GA Settings

One of the powerful features of the software is that it automatically configures the GA settings to values which should give good performance in many cases. If you wish to manually adjust change the settings, you can do so under **GA Settings** (see Figure 41).

So GA Settings	1	-				x
Revert to	o Defa	ults]			
Population Size (Advanced users only)		8	≤	186	≤	8986
Maximum # of Generations (Advanced users only)		1	≤	264	≤	1000
Tournament Size (Advanced users only)				2 🗸		
Crossover Probability (Advanced users only)		0.0	≤	0.5	≤	1.0
Mutation Probability (Advanced users only)		0.0	≤	0.00757575757575757576	≤	1.0
Compare genes in elitism (Advanced users only)						
Tournament replacement (Advanced users only)						
Specify a random seed	V	0	<	347,182		
GA Run Title				2009-04-24		

Figure 41. GA Settings.

- Population size
 - The number of candidate solutions that will be held in memory at any given point in the optimization process, as well as the number of initial randomly created solutions from which the GA evolutionary process begins. This must be an even number. Additionally, it must be a multiple of the tournament size if tournament replacement is off.
- Maximum # of generations
 - Maximum number of permutations of the overall GA evolutionary process that will be performed.
- Tournament size
 - The number of individuals that will compete for mating at any given reproduction instance.
- Crossover probability
 - The probability that the Uniform Crossover process will be performed to combine two parent plans into two new child plans. This should be kept at 0.5.
- Mutation probability
 - The probability that each gene will be modified in a random fashion.
- Elitism gene comparison
 - If enabled, the GA will take into account the differences between plans' genes when computing the crowding distances. While this increases the diversity of the population and ultimately can result in better solutions, unfortunately, it takes exponential computational time as the population size increases. Therefore, this option is only practical for populations of less than a few hundred.
- Tournament replacement
 - If enabled, return individuals to the competition pool after they have competed in a tournament.



SampleOptimizer[™] Dashboard

Now that the settings have been configured, you are ready to start the main optimizer.



Figure 42. SampleOptimizer[™] Dashboard

To begin, click on Start Optimizer.

While the optimizer is running, you can click on pause to momentarily halt the progress of the optimization. While you can click on pause at any time while the optimization is running, the optimizer will continue to run until it reaches a checkpoint. Checkpoints occur after every **generation** of the GA.



Figure 43 shows the optimizer paused at a checkpoint. The graph, called a *Pareto Front*, is a visual representation of the current sampling plans under consideration by the optimizer. Each plan is represented by a dot on the graph. By default, the y-axis is the average sampling cost (per event for spatial optimization, and per year for temporal / spatio-temporal optimization), while the x-axis is the modeling error. The axes are user-configurable with a drop-down menu accessible by right clicking on the axis title and using the resulting drop-down menu.

Important! Error, as used by the "Cutoff" and "Percentage" error objective calculators, does not have a direct physical unit because it is a ratio of the actual error to the acceptable error. When using these Objective Calculators, errors of less than 1 are in compliance with the error objective function.

n * Current *									
Eile Edit Help									
Sampling Data 🚳 Settings 🔗 S	ample <u>O</u> ptimizer™ Dashboard								
	Continue Optimizer	Pause Optimizer Stop Optimizer							
Generation: 23 of 98 Number of Nondominated Plans: 59									
0 Sampling Cost Per Event \$4,900		:							
	Ren	enzene (unt) Error Batio							
🛛 🖉 Plan Det	ails Compare Pl <u>u</u> mes	Dashboard Settings Save Graph							

Figure 43. Optimization Paused

The area of the graph highlighted in green encompasses plans which have an error objective less than 1.



Pausing

When the Optimizer is in the paused state (such as in Figure 43), you may interact with the intermediate results.

- Click inside the Pareto Front graph area to bring up the crosshairs.
- Click and drag a selection box around a plan or plans you wish to compare to the maximum sampling plan.
- When the mouse cursor is hovered over the Pareto front a few seconds, information will pop up about that location on the graph, and any plan(s) which may be located therein.
- Hold down the Control key while dragging to add to the current selection.
- Hold down the Alt key to subtract from the current selection.
- Hold down the Shift key to zoom in.
- Right click in the graph to reset the zoom.

There are two options when comparing plans. You can analyze the details of the plan (click **Plan Details**), or look at the plume maps built using them (click **Plumes**).

Dashboard Settings

There are several configurable options which affect the **Plan Details** and **Compare Plumes** functions; these are accessed through the Dashboard Settings button.



Figure 44. Dashboard Settings

- Show plan details
 - Remove duplicate Plans (enabled by default)
 - By default the software only shows unique plans; if disabled, duplicate plans will be shown in the Pareto front and in the plan details window.
 - Display crowding information (disabled by default)
 - This is extra information typically of use only to a GA expert.
- Default Plume Comparison # of vertical slices
 - Sets the default vertical resolution when generating plume images from the Pareto front, for both spatial and temporal analyses. (note – this value does not apply to the "model visualization settings" used within *ModelBuilder™*, which has a separate input value for # of vertical slices)



Plan Details

Figure 45 and Figure 46 show the plan details comparison for a spatial optimization. Note: rank is a GA term showing which front a solution belongs to. The 0th rank consists of the front of non-dominated solutions.

🗞 Plan	Details									
Save		Plan com	parison view	Plan 20 Pla	n 65 Plan 16	i i				
	Select All	Dian # 1	2 MW 004	MW 007	MW 000	MIN 100	MW 101	MM 102	Pennene (un I) Errer Datie	Sampling Cost Des Event
	Plan comparison view	20	0	0 Of	4 Off	00	04	04	1 17022	sampling cost Per Event
	Plan 20 Plan CC	20	00 0	on Or off Of	1 0ff	00	Off	Off	0.41962	\$2,700
	Plan 65 Plan 16	16	01 0	n Of	f Off	01	Off	Off	0.23619	\$3,000
	Save to .cov									
	Export to GEMS		*							4

Figure 45. Plan Details - Comparison (Spatial)

💰 Plan Details		x
Save	Nan comparison view Plan 16	_
Select All Plan comparison view Plan 16	On Off 1 MW-012 MW-001 3 MW-021 MW-003 4 MW-022 MW-003 5 MW-022 MW-007 6 MW-031 MW-006 7 MW-022 MW-010 8 MW-024 MW-016 10 MW-041 MW-031 11 MW-044 MW-031 12 MW-045 MW-032 13 MW-064 MW-031 14 MW-064 MW-032 15 MW-066 MW-072 17 MW-073 MW-087 19 MW-084 MW-082 19 MW-084 MW-082 19 MW-084 MW-082 19 MW-084 MW-082 19 MW-085 MW-082 19 MW-084 MW-082 10 MW-084 MW-082 10 MW-084 MW-082 10 MW-084 MW-087	
Save to .csv		

Figure 46. Plan Details - Individual (Spatial)



Figure 47 and Figure 48 show the plan details comparison for a temporal optimization.

\delta Plai	n Details							
Save		Plan comparison	view Plan 16 P	lan 28 Plan 27 I	Plan 43 Plan 42	Plan 5 Plan 13		
	Select All Plan comparison view	Plan # V-072	MW-078	MW-086	MW-091	MW-100	Benzene (µg/l) Error Ratio	Sampling Cost Per Year
	Plan 16	16 Annually	Semi-Annually	Semi-Annually	Semi-Annually	Annually	0.80344	\$2,550
	Plan 28	28 Annually	Semi-Annually	Semi-Annually	Semi-Annually	Every 5 years	0.86886	\$2,440
-	Plan 27	27 Annually	Every 4 years	Semi-Annually	Semi-Annually	Every 5 years	1.04297	\$2,345
	Plan 43	43 Annually	Semi-Annually	Semi-Annually	Semi-Annually	Every 5 years	0.85382	\$2,520
[FT]	Plan 42	42 Annually	Semi-Annually	Semi-Annually	Semi-Annually	Semi-Annually	0.51341	\$2,750
[[1]	Plan 5	5 Annually	Semi-Annually	Semi-Annually	Semi-Annually	Semi-Annually	0.64989	\$2,700
	Plan 13	13 Annually	Every 4 years	Semi-Annually	Semi-Annually	Every 5 years	1.05745	\$2,265
[Save to .csv							

Figure 47. Plan Details - Comparison (Temporal)

Save	Plan comparison vie	w Plan 28 Plan	5					
Select All	Quarterly	Semi-Annually	Annually	Every 2 years	Every 3 years	Every 4 years	Every 5 years	Off
Plan 29	1	MW-010	MW-007				MW-004	
Plan 5	2	MW-013	MW-072				MW-100	
	3	MW-024						
	4	MW-027						
	5	MW-033						
	6	MW-040						
	7	MW-049						
	8	MW-060						
	9	MW-078						
	10	MW-086						
	11	MW-091						

Figure 48. Plan Details - Individual (Temporal)

To export the plan details to .csv format data files, select the details you would like to save in the **Save** portion of the window and click **Save to .csv**.



So Choose direct	ory to save ind	ividual(s) in	x
Save in:	🕕 Tutorial	▼	¢ 🕫 🗉 📰
Recent Items			
Desktop			
Documents			
Maria Terrapin			
Network	File <u>n</u> ame: Files of <u>t</u> ype:	Envirosolutions\SampleOptimizer & SampleTracker\Tutorial All Files	Save to directory

Figure 49. Save Plan Details

Here is what the saved files look like when opened in *Microsoft Excel*:

	Microsoft Excel									
	Home	Insert	Page Layo	ut Form	nulas Di	ata Revi	ew Viev	v Add-I	ns	0
Pas	te 🖋	alibri B I U ~ E ~ 🖓 ~ Font		■ = = E = = 律 律 ≫ Alignmen	Ge ⊡	eneral • • % • 0 \$00 Jumber •	Styles	r Insert ▼ Delete ▼ Format ▼ Cells	Σ ÷ ŽΥ	•
	A1	-		<i>f</i> ∞ On						×
200	9-01-01-Pla	n_compariso	n_view.csv							
	Α	В	С	D	E	F	G	Н	1	
1	Plan #	MW-001	MW-002	MW-003	MW-004	MW-007	MW-008	MW-010	MW-011	M١
2	64	Off	Off	Off	Off	Off	Off	Off	Off	Or
3	27	Off	Off	Off	Off	Off	Off	Off	Off	Or
4	14	Off	Off	Off	Off	Off	Off	Off	Off	Or
5	16	Off	Off	Off	Off	Off	Off	Off	Off	Or
I	< → → 20	09-01-01-	Plan_comp	arison_vie	w 🦄					
	2000 01 01 1	Diam 64 any								
	2009-01-01-1	Plan 04.CSV	C	D	E	E	G	ш		
1	On	Off	C	U	L		U			
2	MW-012	MW-001								
3	MW-013	MW-002								
4	MW-021	MW-003								
5	MW-024	MW-004								
6	MW-027	MW-007								-
H.	()) 20	09-01-01-	Plan 64 🦯] 4			•	.::
Read	ly						100%	9	Ū (÷:

Figure 50. Saved Plan Viewing



In addition to saving a plan to a .csv, you can also export it into a Summit *GEMS* database location group. Once in *GEMS*, the location group can be exported to a shapefile for viewing in ESRI *ArcView*.

So Choose GEMS database to create location group(s) in								
Save in:	🕕 Tutorial			- 🥬	P 💷 📰			
Recent Items	르) GEMS DB.	mdb						
Desktop								
Documents								
i 🌉 Terrapin								
2	File <u>n</u> ame:	GEMS DB.mdb			Save to database			
Network	Files of type:	MS Access Database (.mc	b)	•	Cancel			

Figure 51. GEMS Export



Plume Maps

Clicking on the Plumes button will instruct the software to generate plume maps for viewing. They will automatically pop up when they are ready. Please note that while you are viewing plumes, you cannot interact with the **SampleOptimizerTM Dashboard**.

Here is what plume comparison looks like with a spatial optimization, showing a 50% sampling cost reduction (note that a "+" symbol in the figure indicates a location that was proposed for removal in the sampling plan):



Figure 52. Plume Map Comparison (Spatial)



Here is what plume comparison looks like with a temporal optimization (the actual plume map illustrated is for one sampling event, but all sampling events can be accessed and illustrated):

80 Plume Visualizations			
Color	Benzene (µg/l)		
Auto 1 range/site Reset	5 -	•	
Barran (up ()	4.5 -		
Maximum 5.0		0	
Minimum 0.0	ling 4 -	•	
Resolution	Ще 3.5 — С		
Use Defaults Recalculate	Жад	- • • •	
# of vertical slices 150	2.5 -	· · · · · · · · · · · · · · · · · · ·	
# of border slices 10	Loca	0	
	E 2		
Q4 2005	କୁ 1.5 – ଓ		
	व 1 –		
Q2 2005 Q4 2005 Q2 2006 Q4 2006	0.5		
	0_		
	5 -	•	
	4.5 -		
	1	•	
Legends, Symbols, & Labels			
🔽 Draw Legends	3.5 -		
Draw Sampling Location Symbols	3 -	- + • •	
Add Unsampled Locations	8 c 2.5 -	- ⁰ 0 o	
Draw Sampling Location Labels	۲ <u>۲</u>	0	
Location IDs O Values O Interpolations			
Font Size 🗸 Use Default 🛛 8 📩	1.5 -	- · · · · · · · · · · · · · · · · · · ·	
Details	1 -		
Plan comparison view Plan 28	0.5 —		
Plan # MW-004 MW-007 MW-010	0 -		
28 Every 5 years Annually Semi-Annually			
(
Save			
Select all Save All Sampling Locations Max Sa Save			

Figure 53. Plume Map Comparison (Temporal, Ordinary Kriging)



To export the plan details to .csv format data files, select the details you would like to save in the **Save** portion of the window and click **Save to .csv**.

So Choose direct	ory to save file:	; in				×
Save in:	🕕 Tutorial			- 🥬	Þ 🗉 📰	
Recent Items						
Desktop						
Documents						
Terrapin						
Network	File <u>n</u> ame: Files of <u>typ</u> e:	Envirosolutions\Sample(Optimizer & SampleTrack	er\Tutorial	Save to direct	ory

Figure 54. Plume Map Saving

Here is what the saved files look like when opened in *Microsoft Windows Vista's Windows Photo Gallery*:



Figure 55. Saved Plume Map Viewing



SampleTracker™

Getting Started

Before you begin using *SampleTracker*TM (abbreviated as ST), you will need both a historical dataset and current dataset to work with. If you have generated output files from *ModelBuilder*TM for Mass Metric and/or Mass Flux, you need to combine them into your historical and current datasets if you wish to track those metrics. In this case, please pay attention to the report dates used in *ModelBuilder*TM to insure they agree with the other data you input to *SampleTracker*TM.

As with *SampleOptimizer*[™], the currently supported data standard used by *SampleTracker*[™] for data input & output is .csv.

Tips and Tricks

If you see an option in the software and are unsure about it, allow your mouse cursor to hover over it for a few seconds and a "Tool tip" will pop up and give you more information about that option.

Overview

SampleTrackerTM reviews current monitoring data against historical data and identifies cases where current data deviates from expectations that are based on the historical dataset. These expectations are formulated in the form of **bounds**, which represent a statistical estimate of the range of values that might be expected in a specific sampling event, based on the historical data at that location. Any current data value which is outside the bounds range for that location is reported to the user as being out-of-bounds.

ST includes two types of bounds calculators: static and time-dependent. The software automatically makes a recommendation for the bounds type to be used based on the historical data. Advanced users can change the bounds type from the default based on the expectations for a given sampling location.



Electronic Data Deliverable (EDD)

The EDD for *SampleTracker*TM is similar to that used by *SampleOptimizer*TM; the main difference is that in SampleTrackerTM, historical data and current data must be separated into two files: one for the historical dataset, and another for the new (current) dataset to be evaluated.

Mass Metric and Mass Flux Tracking

SampleTracker[™] can track Mass Metric and Mass Flux data via the *ModelBuilder*[™] component of *SampleOptimizer*[™]. Please see Mass Metric and Mass Flux for more information on how to use these features.

Non-Detects in SampleTracker™

- For best performance of the bounds calculators, zero values should not be used. Instead, non-detect data with typical RLs should be replaced by a numerical substitute value. ¹/₂ RL is commonly used, but keep in mind that there is no optimal substitute value.
 - In particular, zero values cannot be used with time-dependent decreasing bounds, since those use log-transformed data, and very small artificial or substitute values can distort any actual trends that may be present.
- If the typical RLs vary from event to event, using a common value such as ¹/₂ the median of typical RLs across events may be useful in order to avoid introducing trends where there aren't any.

EDD Notes

- ND values with "high" reporting limits should be screened by the user because they can potentially cause the upper bounds calculated by the software to be higher than may be desired. When 4 or more historical detects exist for a COC at a location, any current or historical ND for that COC at that location with reporting limit (RL) above the 25th percentile of the detected historical values for that COC at that location should be eliminated from the analysis.
- Follow the duplicate sample guidelines in the <u>SampleOptimizerTM EDD</u> for any duplicate samples you find in your historical dataset. If your current dataset has duplicates, you can screen them separately by adding them into the current dataset one at a time.



Date	SiteID	EastCoordinate	NorthCoordinate	Benzene	Chlorobenzene
4/9/1995	BL003	110	502	4.1	240.0
4/9/1995	OS004	233	438	0.9	50.0
4/17/1995	OS003	220	224	100.0	1800.0
4/16/1995	OS005	157	525	0.5	
3/11/1996	MWSL001	197	386	2.7	8.2

Figure 56. Example SampleTrackerTM Historical Data File

Date	SiteID	EastCoordinate	NorthCoordinate	Benzene	Chlorobenzene
4/12/1997	BL003	110	502	1.0	180.0
4/13/1997	OS004	233	438		50.5
4/11/1997	OS005	220	224	0.5	0.5
4/14/1997	MWSL001	157	525	1.5	0.5

Figure 57. Example SampleTracker[™] Current Data File



Managing Historical Data

A minimum number of historical samples are required for a sampling location / COC pair to be analyzed; the user may specify what the minimum number is, but it must be at least four samples (see Figure 56). In particular, for formal RCRA detection monitoring programs a minimum of 8 historical samples is recommended $\frac{1}{2}$.

In general, we recommend that the user add new samples to the historical set for each sampling location / COC pair until it contains at least 8 historical samples present per sampling location / COC pair, keeping in mind that the background data should be representative of values that can be expected in the future.

However, in some monitoring situations the interested parties agree that data from a certain period is representative of what should be expected from that location. This comes up particularly in regulatory settings, where the agency reviews data and approves, or when by regulation the first two years of data forms the "official" historical dataset. In these cases you would keep the historical dataset the same until there is good reason to update it, such as if you start getting repeated exceedences and the interested parties agree that the situation is understood and the old historical data are no longer relevant.

Please note that regularly adding new samples to the historical set may cause the software to not catch any mild trend(s) which may be occurring. If one takes the position that such data are essentially consistent with an in-control process, this is acceptable. Otherwise, the historical set should be kept the same until an evaluation is triggered, and then it should be updated (keeping the last 8, say, observations) after the interested parties agree.



Loading Data

Upon running *SampleTracker*TM, you will be presented with the following screen.

St. Untitled	
<u>File Edit H</u> elp	
Sampling Data	Settings A Results
Data Import	
	Load Historical Data

Figure 58. SampleTracker[™] Initial Screen



Application Options

At any time, you can adjust the applications options. There are four settings which may be changed here.

- Program Path
 - This is for technical support use only.
- Zero substitute value
 - Time-dependent bounds use natural log scale, and since the log of zero is undefined, a small value must be substituted for zero values in the dataset. That value is specified with this parameter. Note that we do not recommend using zero values at all; as stated earlier, non detects in SampleTracker[™] should be replaced by a fraction of the RL.
- Slope test upper tail %
 - This is the confidence value for the slope test done when recommending the bounds type for each location / COC pair.
- Number of digits in result tables
 - The number of digits displayed in the bounds result tables can be adjusted here.

Dr. Application Options				-	x
		Rev	ert to Defaults		
Program Path			C:/Program Files/Summit Envirosolutions/Monitoring Tools/		
Zero substitute value (Time-dependent bounds)	0.0	≤	1.0E-10		
Significance level for one-tailed test of slope	0.0	≤	0.05		
Number of digits in result tables	1	≤	8	≤	14

Figure 59. Application Options



Loading Historical Data

Before loading the historical data, you can adjust the historical data settings if desired (Figure 60). This parameter controls the minimum number of historical samples which are required for SampleTrackerTM to analyze a COC at a given location.



Figure 60. Historical Data Settings

To start a new analysis, click **Load Historical Data** and locate the file you wish to load, or just drag the historical data file onto the application window.



Figure 61. Loading Historical Data



Figure 62 shows an example screen after a historical data file has been input into the software.

தீ⊱ * Untitled *		1.00						_				Į	_	x
<u>File E</u> dit <u>H</u> elp														
Sampling <u>D</u> ata	a 💥 Set	tings 🔼	<u>R</u> esults											
Data Import														
		-												
			Load	Historical Da	ta 🔪	🖁 Historical	Data Setting	s [Load Curr	ent Data				
Data Summary														
Sampling Location	s: 49						COC's:	1						
Historical Data	Current Data													
Benzene														
Denzene														-1
Date	MW-001	MW-002	MW-003	MW-004	MW-007	MW-008	MW-010	MW-011	MW-012	MW-013	MW-016	MW-017	MW-018	
Dec 11, 1990									1.0E-5					
May 18, 1993									1.0E-5					
Jul 7, 1993	1.0E-5									1.0E-5	1.0E-5			
Aug 10, 1994	1.0E-5									1.0E-5	1.0E-5			
May 3, 1995	1.0E-5									1.0E-5	1.0E-5			
Aug 10, 1995	1.0E-5								1.0E-5	1.0E-5	1.0E-5			
Nov 7, 1995	1.0E-5								1.0E-5	1.0E-5	1.0E-5			
Feb 19, 1996	1.0E-5								1.0E-5	1.0E-5	1.0E-5			
Aug 19, 1997	1.0E-5								1.4		1.0E-5			
Dec 9, 1997														
Jul 31, 1998	1.0E-5								1.0E-5		1.0E-5			
Nov 5, 1998														E
Apr 8, 1999	1.0E-5								1.0E-5		1.0E-5			
Sep 8, 1999	1.0E-5								1.0E-5	1.0E-5	1.0E-5			
Dec 20, 1999		1.0E-5	8.2											
Apr 12, 2000	1.0E-5	1.0E-5	5.1				1.0E-5		1.0E-5	1.0E-5	1.0E-5			
Aug 29, 2000							1.0E-5							
Jan 24, 2001	1.0E-5	1.0E-5	2.5				1.0E-5		1.0E-5	1.0E-5	1.0E-5			
Jun 4, 2001	1.0E-5	1.0E-5	1.9				1.0E-5		1.0E-5	1.0E-5	1.0E-5			
Jan 30, 2002	1.0E-5	1.0E-5	1.3				1.0E-5		1.0E-5	1.0E-5	1.0E-5			
Jan 8, 2003	1.0E-5	2.0	2.9	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5		1.0E-5	1.0E-	
May 31, 2005	1.0E-5	1.0E-5	1.0E-5	1.0E-5	2.7	1.0E-5	1202.0	1.0E-5	1.0E-5	1.0E-5		1.0E-5	1.0E-	
Nov 7, 2005	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5		1.0E-5	1.0E-	
May 30, 2006	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5		1.0E-5	1.0E-	
	•													
L														

Figure 62. Historical Data Loaded

Next, click on Load Current Data.

斎 Open						×
Look in:	🕕 Tutorial			•	ø 🕫 🗉	
Recent Items	Current.cs Historical	sv .csv				
Desktop						
Documents						
Terrapin						
Network	File <u>n</u> ame: Files of <u>t</u> ype:	Current.csv Comma-Separa	ited Files (.csv)		•	<u>O</u> pen Cancel

Figure 63. Loading Current Data



Here is an example showing the screen after both a historical data file and a current data file have been loaded into the software.

ő	* Untitled *		100	-			_		_					_ 0	x
[ile <u>E</u> dit <u>H</u> elp														
	Sampling Da	ata 💥 S	ettings	<u>R</u> esults											
Γ	Data Import														
				Loan	d Historical (Data	X Historic	al Data Settir	ngs	Load Cur	rrent Data]			
L	Data Summary														
	Sampling Location	ons: 49						COC	s: 1						
	Historical Data	Current Da	ta												
	Benzene														
	Date	MW-001	MW-002	MW-003	MW-004	MW-007	MW-008	MW-010	MW-011	MW-012	MW-013	MW-016	MW-017	MW-018	MW
	Dec 4, 2006	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5	1.0E-5	
		•	1	11											•
IL.															

Figure 64. Current Data Loaded



Settings

After loading the historical and current data, you can view and/or change the settings for the bounds calculation in the Settings tab (Figure 65). SampleTrackerTM automatically recommends static or time-dependent bounds based on the data history at each location. You can also specify the parameters used in the bounds calculators themselves, by clicking **Bounds Calculation Settings** (Figure 66).

ற்ஃ * Untitled *		
<u>F</u> ile <u>E</u> dit <u>H</u> elp		
Sampling <u>D</u> ata	Settings	
Bounds Advanced Setti	ings	
		Rounds Calculation Settions
Bounds Type		
Sampling location ID	Benzene	
MW-001	Steady-State 👻	A
MW-002	Steady-State 👻	
MW-003	Time-Dependent 👻	
MW-004	Steady-State 👻	
MW-007	Steady-State 👻	_
MW-008	Steady-State 👻	-
MW-010	Steady-State 👻	
MW-011	Steady-State 👻	
MW-012	Steady-State 👻	
MW-013	Steady-State 👻	
MW-016	Steady-State 👻	
MW-017	Steady-State 👻	
MW-018	Steady-State 👻	
MW-020	Steady-State 👻	
MW-021	Time-Dependent 🗸	
MW-024	Time-Dependent 👻	
MW-027	Steady-State 👻	
MW-028	Steady-State 👻	
MW-031	Steady-State 👻	
MW-032	Steady-State 👻	
MW-033	Steady-State 👻	-

Figure 65. Bounds Type




Figure 66. Bounds Calculation Settings



Results

After you have input the data and adjusted any necessary settings, you are ready to have the software compute the results.

🖧 Untitled					
<u>Eile Edit Help</u>					
Sampling Data 💥 Settings 🗛 Results					
Calculate					
Calculate Results					
-Summary-					
Export					
Export In-Bounds Table Export Out-Of-Bounds Table					
Details					
View Graph					

Figure 67. SampleTrackerTM (Ready to Compute Results)



After you click on **Calculate Results**, you will see a screen similar to this one:

A * Untitled *			_				
Eile Edit Help							
Sampling Data	🖇 Settings	A Results					
Calculate							
🛞 Calculate Results							
Summary							
Export							
		r					
			Export In-Bo	ounds Table	Export Out-Of-Bounds Table		
		l	-				
Details							
In Bounds Out Of Bou	unds						
Sampling Location ID	COC	Lower Bound	Concentration	Upper Bound			
MW-001	Benzene	0.00001	0.00001	0.00001	A		
MW-002	Benzene	0	0.00001	1.842/185			
MW-003	Benzene	0.0000001	0.00001	0.00060845			
MW-004	Benzene	0.00001	0.00001	0.00001 E 47940159			
MW-007	Renzene	0 00001	0.00001	0.00001	Ξ		
MW-010	Benzene	0.00001	0.00001	1 107 47334621			
MW-010	Benzene	0.00001	0.00001	0.00001			
MW-012	Benzene	0	0.00001	0.82303916			
MW-013	Benzene	0.00001	0.00001	0.00001			
MW-016	Benzene	0.00001	0.00001	0.00001			
MW-017	Benzene	0.00001	0.00001	0.00001			
MW-018	Benzene	0.00001	0.00001	0.00001			
MW-020	Benzene	0.00001	0.00001	0.00001			
MW-021	Benzene	3.64497689	12.5	34.99620206			
MW-027	Benzene	0	2.2	307.59724436			
MW-028	Benzene	0	0.00001	3.50116444			
MW-031	Benzene	531.99247876	751	1,204.00752124			
MW-032	Benzene	0.00001	0.00001	0.00001	▼		
View Graph							

Figure 68. In Bounds Results



If any COC's were out-of-bounds at any sampling locations, they will be visible in the **Out-of-Bounds** tab.

த் * Untitled *					-	-	1.00	-	
<u>File E</u> dit <u>H</u> elp									
Sampling Data	😵 Settings	A Results							
Calculate									
	🛞 Calculate Results								
Summary									
Export									
	Export In-Bounds Table								
Details									
In Bounds Out Of Bo	unds								
Sampling Location ID	COC	Lower Bound	Concentration	Upper Bound					
MW-024 MW-073	Benzene Benzene	364.2509086	989 210	966.84086394 141.50769488					
		1			1				
I View Graph									

Figure 69. Out-Of-Bounds Results

To save the result tables, hit either the **Export In-Bounds Table** button, or the **Export Out-Of-Bounds Table** buttons.





Figure 70. Exporting In-Bounds Table



Figure 71. Exporting Out-Of-Bounds Table



To view a graph, simply select a row in the table and hit View Graph.



Figure 72. Viewing Graph

To export a graph to a .png graphics file, click on **Export graph to .png** from the **Export** menu.



Figure 73. Exporting Graph (1 of 2)





Figure 74. Exporting Graph (2 of 2)



Index

bounds	
bounds calculator	
GA	
Inverse Distance Weighting	
Kriging	
license file	
logarithmic transformation	
non-detect	
out-of-bounds	
Pareto front	
quantile transformation	
rank	
spatial optimization	
spatio-temporal	
spatio-temporal optimization	
temporal	
time-dependent bounds	
tradeoff curve	



Appendices

Appendix A – Installation prerequisites

Microsoft .net Framework Runtime 2.0:

http://www.microsoft.com/downloads/details.aspx?familyid=0856EACB-4362-4B0D-8EDD-AAB15C5E04F5&displaylang=en

Sun Microsystems Java JRE 6.0:

http://java.sun.com/javase/downloads/index.jsp



Appendix B – Additional References

Websites

- 1. LTMO
 - a. Federal Remediation Technologies Roundtable <u>http://www.frtr.gov/optimization/monitoring/ltm.htm</u>
- 2. Conceptual Site Model Development
 - a. EPA Triad http://www.triadcentral.org/mgmt/splan/sitemodel/index.cfm
 - b. Lawrence Livermore ERD: <u>http://www-erd.llnl.gov/rescue/Topics/SCModel/outline.html</u>
- 3. CSV
 - a. Wikipedia http://en.wikipedia.org/wiki/Comma-separated_values

Books

- EWRI Task Committee on the State of the Art in Long-Term Groundwater Monitoring Design, *Long-Term Groundwater Monitoring: The State of the Art*, ed. by B.S. Minsker, American Society of Civil Engineers, Reston, VA, 2003.
- 2. Isaaks, E.H. and Srivastava, R.M. *An Introduction to Applied Geostatistics*. Oxford University Press, New York, 1990.
- 3. Goldberg, D.E., *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, 1989.

Journal Papers

 Davis, C. B. and McNichols, R. J., 1994. Ground Water Monitoring Statistics Update: Part I: Progress Since 1988. Ground Water Monitoring and Remediation, Vol. 14, No. 4, pp.148-158; this is cited also in the MAROS manual in the references at the end of Appendix VII.

Documents

1. Roadmap to Long-Term Monitoring Optimization, USEPA, May 2005. <u>http://www.cluin.org/download/char/542-r-05-003.pdf</u>



Appendix C – Troubleshooting

- If you get the error dialog shown in Figure 75 ("Could not find the main class."), there may be a problem with the installation of Sun's <u>Java Runtime Environment</u> (JRE) on your computer.
 - Please uninstall the JRE and reinstall it with the latest version available from (<u>http://java.sun.com/</u>).
 - If you have trouble uninstalling the JRE, try using Microsoft's <u>Windows Installer</u> <u>CleanUp</u> (available from <u>http://support.microsoft.com/kb/290301</u>) to uninstall it.



Figure 75. Java Error Dialog – Could not find the main class.

- If you have difficulty finding solutions with sufficiently high accuracy (aka low error), try increasing the population size in the GA settings.
- If automated model building is taking too long, try decreasing the resolution.



Appendix D – Technical Support, Training, Sales, and Consulting Information

User Forum

A free web forum is available for users to communicate with each other about $SampleOptimizer^{TM}$ related issues, including long-term monitoring optimization in general. To sign up, all you need is an e-mail account; go to <u>http://www.sampleoptimizer.com/forum</u>.

Individualized Support

- If you have a retail license, it may include phone and/or e-mail technical support. For more details, please see the sales receipt that came with your license purchase. If it is not available, please contact the license support & billing department for assistance.
- If you have a non-commercial license (such as an educational, governmental, or demo license), or your retail license support contract is expired, support is available on a paid basis. For more information, please contact Matt Zavislak (see below).

Group Training

Summit plans to offer training seminars available to the general public. Additionally, Summit can come to your company or organization to offer a customized group training event.

Consulting

Summit Envirosolutions, Inc. is a full-service environmental engineering and cultural resources consulting firm which has been helping governments, businesses, and individuals find cost-effective and safe solutions to their environmental concerns for over 15 years. Please contact John Dustman to find out how Summit can assist your organization.

Contact Information

Sales & Consulti	ng	
John Dustman	jdustman@summite.com	651-842-4203
Technical Suppo	rt	
Matt Zavislak	mzavislak@summite.com	651-842-4248
License Support	& Billing	
Stacey Scarcella	sscarcella@summite.com	651-842-4201