Simulation and Optimization of Large Scale Subsurface Environmental Impacts; Investigations, Remedial Design and Long Term Monitoring - 8330

L. M. Deschaine, PE
Science Applications International Cooperation
35 Varden Street, Aiken, SC
USA

and

Chalmers University of Technology
Department of Physical Resources
Complex Systems Group
Goteborg, Sweden

Larry.M.Deschaine@alum.mit.edu

ABSTRACT

The global impact to human health and the environment from large scale chemical / radionuclide releases is well documented. Examples are the widespread release of radionuclides from the Chernobyl nuclear reactors, the mobilization of arsenic in Bangladesh, the formation of Environmental Protection Agencies in the United States, Canada and Europe, and the like. The fiscal costs of addressing and remediating these issues on a global scale are astronomical, but then so are the fiscal and human health costs of ignoring them.

An integrated methodology for optimizing the response(s) to these issues is needed. This work addresses development of optimal policy design for large scale, complex, environmental issues. It discusses the development, capabilities, and application of a hybrid system of algorithms that optimizes the environmental response. It is important to note that “optimization” does not singularly refer to cost minimization, but to the effective and efficient balance of cost, performance, risk, management, and societal priorities along with uncertainty analysis. This tool integrates all of these elements into a single decision framework. It provides a consistent approach to designing optimal solutions that are tractable, traceable, and defensible.

The system is modular and scalable. It can be applied either as individual components or in total. By developing the approach in a complex systems framework, a solution methodology represents a significant improvement over the non-optimal “trial and error” approach to environmental response(s).

Subsurface environmental processes are represented by linear and non-linear, elliptic and parabolic equations. The state equations solved using numerical methods include multi-phase flow (water, soil gas, NAPL), and multicomponent transport (radionuclides, heavy metals, volatile organics, explosives, etc.). Genetic programming is used to generate the simulators either when simulation models do not exist, or to extend the accuracy of them. The uncertainty and sparse nature of information in earth science simulations necessitate stochastic representations. For discussion purposes, the solution to these site-wide challenges is divided into three sub-components; plume finding, long term monitoring, and site-wide remediation.

Plume finding is the optimal estimation of the plume fringe(s) at a specified time. It is optimized by fusing geo-stochastic flow and transport simulations with the information content of data using a Kalman filter. The result is an optimal monitoring sensor network; the decision variable is location(s) of sensor in three dimensions. Long term monitoring extends this approach concept, and integrates the spatial-time correlations to optimize the decision variables of where to sample and when to sample over the project.
life cycle. Optimization of location and timing of samples to meet the desired accuracy of temporal plume movement is accomplished using enumeration or genetic algorithms.

The remediation optimization solves the multi-component, multiphase system of equations and incorporates constraints on life-cycle costs, maximum annual costs, maximum allowable annual discharge (for assessing the monitored natural attenuation solution) and constraints on where remedial system component(s) can be located, including management overrides to force certain solutions to be chosen are incorporated for solution design. It uses a suite of optimization techniques, including the outer approximation method, Lipchitz global optimization, genetic algorithms, and the like. The automated optimal remedial design algorithm requires a stable simulator be available for the simulated process. This is commonly the case for all above specifications sans true three-dimensional multiphase flow. Much work is currently being conducted in the industry to develop stable 3D, three-phase simulators. If needed, an interim heuristic algorithm is available to get close to optimal for these conditions.

This system process provides the full capability to optimize multi-source, multiphase, and multicomponent sites. The results of applying just components of these algorithms have produced predicted savings of as much as $90,000,000(US), when compared to alternative solutions. Investment in a pilot program to test the model saved 100% of the $20,000,000 predicted for the smaller test implementation. This was done without loss of effectiveness, and received an award from the Vice President – and now Nobel peace prize winner - Al Gore of the United States.

INTRODUCTION

The cost to clean up environmentally impacted sites is high, both for businesses and for taxpayers. For many locations, environmental cleanup is not the core function of the facility. While a very important activity, the cleanup efforts remove dollars from operations budgets. This expense restricts productive capability in the short term and reduces investment core mission improvement in the long term. The methods of remediating aquifers consist of primarily in-situ treatment in which the contaminants are treated in place (via natural attenuation and/or engineered solutions), ex-situ treatment in which the contaminants are removed from the aquifer, and institutional controls through which the sources of contaminants are managed rather than treated. These costs are significant. The U.S. Department of Energy estimated its environmental remediation costs at more than $150 billion [DOE 1997]. About 15 years ago, the General Accounting Office identified this issue by estimating the cost of remediating just the sites on the Superfund list at $26 billion [GAO 1993]. These are only two examples of remediation cost estimates, and do not include all of the sites, such as gasoline stations, dry cleaners, manufacturing facilities, landfills, and the like, requiring remediation. In response, research into the optimization of site investigation programs and remediation designs has been intense. Optimization of solutions for these sites is both mathematically challenging and complex. Aquifers can vary from simple, single-layer geologic units to multiunit perched systems with variable saturated flow. Multiphase impacts may result, including dense, nonaqueous-phase liquids (DNAPLs) such as chlorinated solvents; light, nonaqueous-phase liquids (LNAPLs) such as petroleum products; dissolved-phase contamination in which residual material migration is driven by groundwater flow; vapor-phase contamination; and residual soil contamination. Plume delineation is expensive. Solution options for contaminant migration management are many, ranging from hydraulic pump-and-treat or air sparging systems to optimization of in situ biogeochemical oxidation-reduction potential (Redox) zone design of monitored natural attenuation or engineered systems.

Since the mid 1980s to the early 1990s, optimization of groundwater remediation design for groundwater extraction/reinjection systems has resulted in significant cost reductions. Formal optimization of the remedial design problem has been examined by many researchers who directly couple saturated groundwater flow and transport simulation models with optimization algorithms [Ahlfeld 1986; Ahlfeld and Heidari 1994; Wagner 1995]. In the early 1990s, work was started on an algorithm to optimize remedial designs by integrating advanced multiphase, multicomponent flow and transport simulators,
formal optimization algorithms, heuristic optimization algorithms, and economic functions into a multisource, multiplume, sitewide risk-based optimization tool [Deschaine 1992; Karatzas and Pinder 1993; Karatzas and Pinder 1996]. Along the way the top three deployments of these tools have saved $100M, $90M, and $72.6M(US).

ALGORITHM DESCRIPTION

This formal site wide environmental impact remedial design optimization system is a system of algorithms that fills the gap in available aquifer optimization tools. It forms a bridge between the “trial and error” and the saturated groundwater pump-and-treat-only approaches to remedial system design by providing a formal, structured approach with which to find a least-cost solution, even when the costs and the flow and transport aquifer properties are uncertain. Figure 1 provides an overview of the tool.

![Fig.1. Overview of the site wide optimization tool](image)

- The Plume Manager module controls the overall solution. It integrates information on cost functions and spend-out constraints with environmental simulations to develop the best solution for the challenge at hand for a site-wide integrated solution.
- The Plume Investigator module integrates the flow and transport simulations with information-content modeling (deterministic or stochastic) to optimally design sample programs that define the extent of the impacts and monitor them long term.
- The Plume Simulation module consists of a large number of subsurface flow and transport simulators used to evaluate baseline conditions as well as evaluate potential remedial responses.
- The Plume Cleaner module integrates optimization algorithms with the Plume Simulation algorithms to design least-cost solutions to various remediation options, including monitored natural attenuation and active remediation.
- Information needed to maximize the effective deployment of this tool include the following:
  - A mathematically correct statement of the flow properties of the aquifer. This model can be deterministic or stochastic.
  - A mathematically correct statement of the transport properties of the aquifer, including the biogeochemical processes affecting them in space and time. This model can be deterministic or stochastic.
  - An annual and project life cycle cost function that represents both the capital and operational costs of the various remedial options under consideration.
The constraints on the solution include the maximum desired annual costs, the limits on contaminant discharge or concentrations at point(s) of compliance, and the constraints as to where the remedial action can be physically located and where it cannot. Management overrides—specifying or prohibiting a remedial option at a specific source—are important and are accommodated by the tool.

The integrated optimization algorithm reads the above information and provides feasible and optimal or near-optimal solutions when all things are considered. The main technologies consist of genetic and other evolutionary computation algorithms, Tabu search algorithms, and Monte Carlo and Latin Hypercube simulation algorithms.

SITE-WIDE RISK-BASED OPTIMAL MANAGEMENT MODULE

The site-wide risk-based optimizer integrates the physics of flow and transport with the economics of project management. Acceptance is gained using the verification, validation, accreditation and credibility guidance of the U. S. Defense Modeling and Simulation Office (DMSO). The uses the subsurface flow and transport models as subroutines, so it is physics model independent. The most suitable model can be chosen for the site / question, expanding the flexibility, adaptability, and solution correctness of the optimizer. This approach leverages advances in subsurface simulation code, which have been substantial over the tool’s 23-year development cycle, beginning with a 114 site state-wide optimization program (Deschaine, et. al 1985).

Fig. 2. Site-wide optimization tool’s spreadsheet GUI. The stochastic constraints (discharge limits for each CoC, annual funding, etc.) are located to the right of the figure and represent literally hundreds of constraints on the optimal solution.
The algorithm links the science and economics—the feasible solutions with the business, regulatory, physical and social constraints. Specifically, these business algorithms provide the functionality to optimize the integrated site-wide remediation decision. Because each algorithm must, by necessity, be customized for a specific site, only a general description of the algorithm is possible. Such a description is provided below.

1. Using a suitable subsurface flow and transport model, develop a mathematically correct statement of the source-specific flow and transport system as is and predict future impacts if the contaminant is allowed to migrate unabated. This output can be the result of deterministic simulations (i.e., simulate one “best” aquifer) or of stochastic simulation of equally likely aquifers (i.e., use GSLIB to simulate the realizations of equally likely aquifers). Store these results.

2. Screen remedial options using standard European or U.S. Environmental Protection Agency’s guidance, for example, and select the most likely options per source area if multiple sources are present. Perform value engineering on new or existing solutions, as appropriate. Assess whether the treatment effectiveness is known (deterministic) or estimated. If estimated, build an option-specific stochastic function. Store these results.

3. Assemble the cost constraints, which can be per source area, per year, or per life cycle. Assess whether the cost constraints are known or estimated. If estimated, build an option-specific stochastic function. Store these results.

4. Assemble the environmental and physical constraints. These constraints can be the point of compliance, cleanup level, discharge mass per year, or treatment system component location and can be per source area, per year, or per life cycle. Assess whether the environmental constraints are known or estimated. If estimated, build an option-specific stochastic function. Store these results.

5. Assemble the management override constraints. These constraints are typically to force a certain remediation at one location or prohibit one at another and often relate to land use. For example, force capping on a land parcel for sale or prohibiting pumping and treating using vertical wells on an airport runway.

The current implementation of the site-wide optimizer uses a combination of optimization techniques (discussed below) and Monte Carlo and Latin Hypercube simulation to optimize the solution under uncertainty. It designs and answers the following questions:

1. What is the least-cost solution to the site remediation? How does this solution vary if annual funding is constrained or changed?

2. How will the solution change if cleanup levels are relaxed, points of compliance are changed, or time to clean up varies? How will it change if various remedial alternatives are selected or deselected?

3. When multiple projects or opportunities for savings are present, which ones should be implemented first?

4. Given a desire to be 95% confident that a site-wide solution strategy will be protective of the environment and all constraints will be met, which course of action is best? How will this change if desired confidence is 90% or 50%?

A simplified version of this business process optimization algorithm was used to project savings of $90M over a 5-year period [Deschaine 1998b]. This optimization work, conducted at the U.S. Department of Energy’s (DOE’s) Savannah River Site, received the National Performance Review “Hammer Award” in 1997. An extended version of this algorithm was developed for energy systems and technologies research and development program design for DOE’s Vision 21 Energy Development Program [Deschaine 2001b].
PSUEDO-CODE FOR THE SITE-WIDE OPTIMIZER

The generalized pseudocode for the optimization tool that incorporates the above is as follows:

- Minimize Total Economic and Social Cost of Plume Presence
  - Total Cost = f(investigation, remediation, monitoring reporting, administration, increase in health services, loss of recreational and other use, loss of natural resources, etc.).

- Subject to:

  - The risks to human health and the environment are within acceptable levels within the specified time frame
    - Concentrations Ci,t for all spatial locations (i) within exposure time (t).
  - The mass flux into receptors are less than assimilative capacity (natural attenuation) of the receiving system, such as a receiving surface water body or aquifer.
    - Mass fluxes Mi,t for all spatial locations (i) within exposure time (t).
  - The annual costs are below a specific value, with certainty “p”, for all time periods (t) for example 95% certain that costs will be below budget in any given year.
    - Dollar requirements for all annual times (t).
  - Expert consensus expectation of success greater than minimum value, with probability Pse that success not less than Smin. This constraint balances the “risky” but potentially high reward options with the tried and true methods, similar to balancing a portfolio of investments while minimizing the downside risk, for example constrain 25th percentile of selected portfolio of projects to some absolute downside value. It capture the uncertainty of the experts to know priory the success of the various site management policies.
    - (Pse > Psemin1 and 25th percentile Pse > Psemin2)

Modifying a well-known and widely used decision support algorithm, the analytical hierarchy process to accept stochastic inputs, captures the expert’s uncertainty. The AHP algorithm is presented in [Saaty, 1996]. Other methods reviewed, and viable options to this approach to represent human behavior include fuzzy logic, non-monotonic reasoning, expert systems, reinforcement learning and the like. The specific human behavior method is flexible in this tool, so long as the results pass the DSMO requirements for validation of human behavior in simulations.

Life-Cycle Cost Function

The cost of the remediation project can essentially be broken down into two components: the capital cost and the variable or annual cost. In the cost function discussed above, there are costs that the polluter pays, and costs borne by society. The cost function is developed by an assembly of stakeholders who decide or negotiate which cost components are included in the analysis and optimization of the response policy. This may not be a trivial exercise in many cases. This work allows for the flexible and comprehensive complete or partial analysis of cost functions.

The capital portion of a response policy consists of the cost of the infrastructure, such as the treatment building(s), permits, ancillary components, and the like. The variable cost portion (cost of operations and maintenance over the project life cycle) consists of items such as the recovery wells and pump systems; the pipe and trenching; and the cost of the cleanup operation, including system operation, maintenance, and environmental sampling. The total project life-cycle cost, which consists of both components, is the cost function that is optimized within the annual funding and other project constraints discussed above. The “balance” question is what capital improvements to plume management to make when, recognizing...
the action and in action have economic and social costs when the total cost of plume existence is considered. The total cost of ownership is developed in the value engineering phase. In this phase, the stakeholders are assembled and each aspect of the problem and potential solution is quantified to the best of the group’s ability – recognizing and cataloging uncertain aspects. Solving for the optimal solution based on this quantification and acknowledged uncertainties is discussed below.

Value Engineering

Value engineering is used to specifically define the cost function(s) of the various site management options, including scope of investigation, remediation alternatives and long term monitoring. This consists of essential meetings with the various stakeholders to develop the goals of the management policy. The stakeholder group(s) heuristically optimize, to the extent practical, the decision components. This step is the precursor to the formal optimization, and supplies the accreditation and facilitates acceptance of any optimal solutions developed. The investigation and remedial options are evaluated, and using a combination of costs in the project databases or actual operation costs, a very detailed cost matrix is developed. The cost function is developed such that changes in one aspect of a solution policy under consideration propagate throughout the integrated site-wide optimization system.

PLUME SIMULATION MODULE

Many groundwater flow and transport codes currently exist that employ either the finite difference or finite element solution techniques. They can solve either single- or multiphase flow and single- or multicomponent transport. The choice to model a site using finite differences or finite elements is site specific—it depends greatly on the ability of the technique to represent the subsurface flow system with a mathematically correct statement. Guidance is available on model selection criteria and solutions for many of the state-of-the-art models available at this time [BWXT and SAIC 2002]. Examples of the models used, along with a brief summary description of their capabilities, are presented below.

Finite Element Method

- Princeton Transport Code with Plume-Finding Technology: a saturated/unsaturated flow and single-component transport model, very fast and robust, excels when many calls are needed to evaluate options.
- BioF&T3D & SA_MAPS: a variably saturated, multiphase, multicomponent flow and transport model that includes biochemistry and dual-porosity formulations such as air sparging and groundwater/surface water interactions.
- FEHM & PORFLOW: Variably saturated flow and multicomponent flow and transport models that include varying degrees of biogeochemistry. FEHM also allows for multiphase and dual-porosity/dual-permeability formulations.
Finite Difference Method

- MODFLOW: a saturated flow model from which several transport and optimization codes are linked.
- MT3DMS & MOC3D: single-component transport codes linked with MODFLOW.
- BioRedox, SEAM3D, TRACR3D, and RT3D: biologic transport simulators linked with MODFLOW (except TRACR3D, which has a stand-alone flow code). OS3D links with MODFLOW and simulates metals transport, as does the CRUNCH geochemical code.
- UTCHEM and Tough2v2: multiphase simulators with various transport capabilities, with UTCHEM being the more comprehensive.

It is important to note that for various remediation technologies, various simulation codes may be used. No one code can always meet the needs of the analysis, whether they be complexity of flow or transport or solution time. For example, a fully three-dimensional, multiphase, multicomponent simulator would not usually be chosen to solve a saturated regional flow system with a nonreactive contaminant. Not only would use of such a simulator be overkill, but it would also be inefficient and require much longer run times. Conversely, bioairsparging of an LNAPL or surfactant flushing of a DNAPL would not be solved with MODFLOW. It would be an inappropriate use of the simulator.

PLUME INVESTIGATION MODULE

Determining the nature and extent of a plume requires finding the fringe in three dimensions. The plume investigation module of this site-wide optimizer uses a combination of tools to accomplish this effectively and efficiently [McGrath and Pinder 1996; McGrath, et al. 1997; McGrath 1997]. The plume-finding module is a program for guiding the location of sensors (i.e., soil borings, cone-penetrometer probes, or monitoring wells) in groundwater investigations of contaminated aquifers. The objective of the investigation is to provide the groundwater professional and stakeholder assistance in determining the number of samples and locations needed to delineate the boundary of a contaminant plume in a three-dimensional aquifer. The “boundary” of the plume is defined by the aqueous-phase contaminant concentration standard, such as a maximum contaminant level or other risk-based standard.

The goal is to quantify the information that an existing monitoring well network provides in establishing knowledge of the location of the groundwater plume and to identify (prior to installing a sensor) the next sampling location(s) in three-dimensional space that, when sampled, minimizes the uncertainty of the plume boundary location. This will provide a prioritization of the sensor installation activity with each new proposed sampling location at the location where it will provide the maximum amount of information for solving the plume location challenge. It will also quantify the reduction in incremental knowledge gained by each new sensor; thereby providing a stopping point when the sensor network is adequate. Adequacy of plume delineation is defined by the stakeholder specified data quality objectives. The following describes the major elements in the plume finding process.

1. Determine the plume fringe by assessing where the fringe of the contaminant of concern’s isocontour(s) lines are. Generating multiple realizations of the aquifer and plotting the expected concentrations and distributions at each node complete this plume line. The fringe is added by identifying which nodes contain the iso-contour with 95% certainty, for example. This is straightforward statistical analysis.
2. Assess the reduction in the uncertainty of the knowledge of the plume location given an existing monitoring well network. This is accomplished by combining the physical subsurface flow and transport model and the Kalman filtering approach (Note: The Kalman filter is discussed below in the long term monitoring section. The major difference is that here we specify the time and solve for the best location to have monitoring information. The long term monitoring solves for the best space and time for the information being collected.)

3. Assess the best location(s) to add new sensors (monitoring wells, soil gas data collection, etc.) to reduce the uncertainty in plume location definition by the maximal amount.

The procedure below is used to implement the plume-finding algorithm.
Construct a preliminary subsurface flow and transport model as discussed above. This initial model can be quite simple and does not require detailed site knowledge to be effective.

1. Apply the plume-finding technology to generate a rank-ordered list of sample locations of monitoring points/wells.
2. Collect the data and add this information to the observation database.
3. Update the plume finder statistical information.
4. Repeat until the confidence in the knowledge of the plume location meets project requirements.

By maximizing the benefit of each sampling point, optimal sampling programs (maximum information for the least cost) can be designed to best meet project goals, and the design can be defended. An optimal network can be designed by selecting which location(s) for new monitoring well(s) reduce the volume under the uncertainty surface by the greatest amount.

The plume finding example, shown in Figure 3, is produced using the Princeton Transport Code as the subsurface simulator, the sgsim aquifer conductivity package from GSLIB, and the Kalman filter as implemented by [McGrath and Pinder 2003]. Only the results that were shown publicly are reproduced here, as we are in process of preparing and obtaining clearance for detailed publication. A general description of the problem follows.

The goal of the study was to assess the placement of between six and eight new monitoring wells to assess the effectiveness of the existing monitoring well network on delineating the plume fringe. About a dozen monitoring wells exist at the site. The aquifer of concern is 500 ft below the ground surface. The cost of each well is $150,000 to install, with additional life cycle costs resulting from the periodic sampling and reporting of results. An existing site model was available, as well as a regional variogram. The first step was to assess the maximum uncertainty of the plume without any monitoring wells installed. In figure 3, the direction of flow is along the longer x-axis (0 to 24,000 ft). The y-axis extends from 0 to 12,000 ft. The worth of sample data values is plotted on the z-axis. The highest value represents the location where the most value is obtained by collecting information at the peak of that surface. With no monitoring well network, the maximum uncertainty was determined to be 40 “worth of sample data” units. This condition is shown on Figure 3a.

The next part of the analysis was to assess the robustness of the existing monitoring well network on meeting the project objectives of plume delineation. About 12 wells are present. The value these well provide on reducing the uncertainty is substantial. The maximum uncertainty measure has been reduced to less then three, and occurs primarily downgradient of the existing monitoring well network. If further reduction in uncertainty is desired, then the optimal location to install one additional well is at the top of the surface shown in Figure 3a. The addition of a well at that location will reduce the uncertainty at that point to almost zero (the sampling error). The added information will also propagate in space around that location. The resulting new surface is shown on Figure 3b.
Note how the surface in Figure 3b was “pulled down” resulting in Figure 3c when the well was hypothetically added to the existing monitoring well network at the peak in Figure 3b. This shows how the information obtained from a monitoring well not only provides information at the sampling point, but aerially as well. This quantification of spatial information from a point estimate is a direct result of the linkage of flow and transport simulator, multiple aquifer realizations, and the Kalman filter. If another far field well is desired, the optimal location is the location of the peak in Figure 3c. Adding information at this location results in the uncertainty surface depicted in Figure 3d. The percentages shown in Figure 3 relate to the maximum residual uncertainty after a monitoring well is installed.

![Figure 3. Summary of PlumeFinder Analysis Value of Additional Wells, Scale Exaggerated To See Results [Results of 4000 flow and transport simulations]](image)

At this point, the uncertainty in plume location is more focused near the source area. This indicates that the efforts to characterize the source will have more value that adding monitoring wells. Should constraints on well locations be present, the technique allows the placement of a well anywhere and will quantify the reduced benefit of not being able to locate a well in the precise optimal location.

The key finding is that knowing the concentration value in a far field well reduces all the uncertainty between the well and the source area. It also propagates this information in the perpendicular direction to that line. Hence, the benefit of adding distal wells, and the information they provide is tremendous. This is in contrast to approaches where the space-time correlation is not fully exploited, and a well only reduces uncertainty near the sampling point.

The plume finding technique is very valuable when allocating resources among the various total costs of an environmental contamination challenge. Specifically, this enables balancing investigation with remediation and long term monitoring. It also quantifies the level of certainty between investigations or various source areas on a site. This balances the site-wide investigations so that one area is not over assessed at that cost of large uncertainty in another part of the site.
Long Term Plume Monitoring Module

Once the solution to the plume fringe finding has been designed and accepted, the long-term operations and monitoring often represent a very expensive and significant portion of the investigation portion of the life-cycle cost. Solutions to this challenge are found in Herrera, et al. 1998; Herrera and Pinder 1998; Herrera 1998; Zhang and Pinder 2000; and Zhang and Pinder 2002; Minsker, 2003. The long-term monitoring program is optimized through an integration of plume simulation, statistical regression, Kalman filter, and a genetic algorithm. The objective of this analysis is to determine the location and timing of groundwater quality sampling events (in existing or proposed sensor locations) so as to achieve specified target goals of statistical accuracy over the specified period of analysis at the minimum cost. In this case the target goal is the coefficient of variation (CV) divided by the mean concentration at the required time, and the location should be less than a certain number (e.g., 20%).

Essentially, this technique can be thought of as an extension of the plume finding algorithm. Whereas the plume finding algorithm uses an estimate of site conditions from a physical simulator and the field data to estimate plume location and maximum uncertainty in that knowledge, the long term monitoring algorithm incorporates the time dimension as well. In this case, the information content of a sample decays as time passes, so that you can visualize a concentration graph as below, with a predicted smooth concentration transitions and sample point values proximate to it. This results in errors bars on the time-concentration curve, so that when a sample is taken the error is at a minimum. As time passes, the error bars will increase until it reaches a preset limit that triggers the next sampling event, reducing the error bars to a minimum value again. By optimizing the maximum error as a system, the least cost accuracy constrained long term monitoring policy is developed.

The procedure below is used to implement the long-term monitoring optimization algorithm.

Objective Function:

$$Min \sum_{j} \left( \sum_{i} (c \cdot \omega_{i,j} + c_0) \cdot x_j \right)$$

Subject to:

$$\omega_{i,j} \leq x_j \quad (\forall i \in I, j \in J)$$

$$CV_{i,j} < CV_0 \quad (\forall i \in I, k \in K)$$

$$x_j = \begin{cases} 1 & \text{if a well is installed at location } j \quad \forall j \in J \\ 0 & \text{if no well is installed at location } j \quad \forall j \in J \end{cases}$$

$$\omega_{i,j} = \begin{cases} 1 & \text{if a sample is taken at time } i, \text{ location } j \quad \forall i, j \in J \\ 0 & \text{if no sample is taken at time } i, \text{ at location } j \quad \forall i, j \in J \end{cases}$$

Where *J*: sets of potential well locations and *i*: sets of sampling periods.

The results are presented in Figure 4.
Figure 4. Long-term Monitoring Optimization Algorithm Concept. When the uncertainty becomes excessive, a signal to collect is sample is triggered. Note: In this formulation, the uncertainty is formally computed as a correlated function of both space and time.

The Kalman filter is an optimal estimator that combines the measurements and system information to achieve a minimum error estimate. It is used to update the statistics of the concentration field after a sample or samples are taken. By combining the information from the physical models with the field measured data sets, the site knowledge and analysis techniques are used to a maximal extent.

In our case, we have a “measurement model” as follows:

\[ Z = Hx + \nu \]

where,

- \( Z \) is a vector of m-noise-corrupted measurements;
- \( H \) is the measurement matrix, the dimension is \( m \times n \); and in this case, \( H \) is the sampling matrix of the form \( \{0,0,1,1,0,...,0,0,1,0,1\} \), where the values of \([0,1]\) indication to take a sample at position \((x,y,z)\) at time \((t)\).
- \( x \) is a vector of dimension \( n \) which is the computed value of the desired quantity;
- \( \nu \) is a vector of normal random noise, \( \nu \sim N(0, R) \).

The Kalman filter is incorporated as follows:

State estimate update:

\[ \hat{x}(+) = \hat{x}(-) + K[Z - H\hat{x}(-)] \]
\( P(+) = [I - HK]P(-) \)

where, (+) and (-) signs are estimates immediately after and before a measurement is take, respectively;
- P is the error covariance matrix, \( P = E[(\hat{x} - x)(\hat{x} - x)^T] \);
- K is the Kalman gain matrix, obtained by minimizing a weighted scalar sum of the diagonal elements of \( P(+) \); \( K = P(-)H^T[HP(-)H^T + R]^{-1} \)

To implement this approach to long term monitoring design:
- Generate realizations using a Latin Hypercube sampling technique assuming:
  - hydraulic conductivity is a random field, and
  - contaminant source is a stochastic process.
- Simulate groundwater flow and transport to generate contaminant-concentration realizations.
- Compute a space-time correlation matrix of concentration.
- Use a genetic algorithm to select target locations and times to reduce most of the CV at the monitoring locations. A Kalman filter is used in this step to update the space-time correlation matrix after a sample is taken.
- Continue the selection process until the project goal is achieved.

Long term monitoring is the process where over time information on the location and characteristics of a subsurface plume in monitored. For further details, see [Zhang 2002] as extended by [Zhang and Deschaine 2004]. Specifically, the confidence of the plume knowledge is specified within a preset level, and the least cost, multi-period, long-term solution that satisfies the accuracy constraints is developed. To develop the long term monitoring program that best suites that needs of a project, several preset levels can be analyzed, and a cost versus accuracy curve generated. The stakeholders can then decide what level of monitoring will be necessary given the quantitative information on how various costs and accuracy are interrelated. The resulting optimization will allow for the development of a monitoring program using optimization algorithms to determine which monitoring wells to sample, the frequency to sample them, and which analytes to sample. The solution will identify which wells are key for the plume monitoring and which have minimal value. Depending on the results, minimal or no value wells can be removed from consideration by the algorithm and the optimization rerun to see if any degradation in accuracy occurs. If not, the stakeholders could consider these wells for abandonment. If so, the monitoring wellhead protection plans can be reassessed to ensure the monitoring sensor assets are adequately protected.

**EXTENSIONS TO DNAPL SOURCE AND UXO FINDING ALGORITHMS**

A DNAPL source algorithm is under development using the above plume and long term monitoring techniques, and is significantly extended to include new formulations regarding geologic uncertainty, incorporation of expert knowledge and the like. Information will be available from the SERDP.ORG web page when available (Pinder, Dakou and Deschaine, 2008). In a similar manner, a UXO finding algorithm is developed (Deschaine, et. al, 2002c), and deployed (Francone, Deschaine and Warren, 2007) on the field scale, and now funded for additional testing by DoD as ESTCP project number MM1-017 (Deschaine, Francone and Keiswetter, 2008). To wit, these techniques – and the ability to add to and extend them in other disciplines - have proven quite beneficial.

**ACTIVE PLUME REMEDIATION MODULE**

The constrained engineering design problem is quite complex and well studied. The site-wide optimizer needs to have the option of using multiple optimization algorithms, depending on the site-specific flow and transport system and the source-specific remedial technology. Aquifer remediation methods consist of
primarily in situ treatment in which the contaminants are treated in place (via natural attenuation and/or engineered solutions), ex situ treatment in which the contaminants are removed from the aquifer, and institutional controls through which the sources of contaminants are managed rather than treated. The objective function is concave, with the minimum occurring on the lower rim of the state-space. Constraints on allowable residual contamination levels are present, as are physical constraints of the aquifer ability to be remediated as well as locations available for remediation infrastructure.

The general formulation for this challenge is:

Minimize \( f(x) = \sum_{i=1}^{N} (\delta_i, \alpha_i, q_i) \)

Subject to:

\[
\sum_{i=1}^{N} q_i \geq Q^* \rightarrow i \in I
\]

\( q_i \leq q_i^* \rightarrow \forall i \in I \)

\( c_j \leq c_j^* \rightarrow \forall j \in J \)

\( h_k \geq h_k^* \rightarrow \forall \in K \)

\( t \leq T_{\text{max}} \)

In the above formulation, the minimization function \( f(x) \) is a simplified version of life cycle cost. This example cost formulation is approximated by the flow from a recovery well, \( q_i \) times a unit cost to treat the water, \( \alpha_i \) times a \((0,1)\) \( \delta_i \) multiplier if a well installation is needed or is preexisting. This simplified version of the cost function is for illustrative purposes. Some solutions to this challenge include quite extensive implementation of a cost function, including using industrial remediation cost estimating software. Formulations exist for other remediation technologies (see Deschaine, 2007 for example), including bioremediation but are beyond the scope of this paper to formally explain. For a simple pump and treat applications, the \( q > Q \) constraint requires that the solution pump at least some minimum amount of water. The \( q < q^* \) constraint sets upper limits on a flow at a well. The \( c < c^* \) constraints ensure the aquifer is cleaned up to a certain acceptable residual level. The \( h > h^* \) constraints set maximum allowable drawdown in the aquifer. These can be expanded to include constraints for differential settling from dewatering by adding constraints on the maximum slope of the water table. The \( t < T \) constraint limits the allowable time for the activity to achieve the specified goals. Different remedial technologies have various and different physics models. Also, the general constraints - like the cost function - can be modified and adapted as specific project needs dictate.

The goal of the constrained global optimization algorithm for this challenge is to find at least one point that \( x^* \) within the feasible region that satisfies \( f(x^*) \leq f(x) \) for all \( x \), or show that such a point does not exist. Showing that a point does not exist is instrumental to this challenge, as cases have occurred when the number and type of constraints added precluded a feasible region to the problem (Pinder 2003 – personal communication). This requirement screens out the many of the blind search methods, such as genetic algorithms, as these methods would iterate ad infinitum without providing either a solution or proof that one did not exist.

From a purely mathematical point of view, concave minimization belongs to the hardest class of mathematical programming problems (Pinter 2001). Outer approximation and Lipchitz Global Optimization are proven techniques for this challenge. On the other hand, it is this characteristic that makes many of the interior point methods, such as evolutionary computation methods, not well suited to solve this type of problem from a theoretical standpoint. Other researchers have tried simulated annealing, genetic algorithms, and the like, without much success.
**Outer Approximation Method**

This optimization algorithm, the Outer Approximation Method (a.k.a., the Cutting Plane Technique) is described by Karatzas and Pinder 1993; Karatzas and Pinder 1996; Spiliotopoulos, et al. 2000; and Deschaine, et al. 2001 a. The method is a global minimization technique that uses a cutting plane approach to determine the optimal solution. The algorithm starts by determining a polytope defined by a set of vertices that encloses the feasible region. The feasible region is determined as the space in which all of the constraints are satisfied. It should be noted that a robust simulator of the subsurface physics is necessary for automatic optimization techniques to be successful. Some of the problems, such as multiphase flow of water, NAPL and gas where all three phases are active are inherently difficult to get to converge, are expected to create a challenge during automated optimization.

The objective function, the function to be minimized, is formulated as a concave function. The objective function is also not differentiable at the intersection of the capital and operational costs. This algorithm takes advantage of the fact that a basic property of a concave function (f) is that the minimum of the function over a compact set of constraints is always obtained in at least one extreme point of the set. It does not rely on derivatives and, hence, the discontinuity of the objective function is not a performance issue. It also handles integer programming, as either a well is selected for use or not, but a half well has no physical meaning.

The outer approximation algorithm (Karatzas 1992) is implemented as follows:

1. Define the enclosing polytope.
2. Define the vertex that will minimize the objective function.
3. Calculate the constraints and determine the most violated constraint.
4. Generate a cutting hyperplane at the vertex of the most violated constraint and determine the new set of vertices without cutting through the feasible region.
5. Iterate until optimal solution found or infeasibility (zero vertices) declared.

The outer approximation algorithm determines the vertex of the enclosing polytope that minimizes the objective function. Next, it examines if the selected vertex is feasible. If all constraints are satisfied, it declares this vertex as the optimal solution. Otherwise, a cutting plane is introduced that eliminates this vertex and its surroundings, creating a new enclosing polytope that is a better approximation of the feasible region, and the process is repeated. The goal of this process is not to determine the best approximation of the feasible region but rather to determine the most extreme point of the feasible region without eliminating any part of it.

If the cutting plane technique cuts the solution space until there is nothing left, then an “infeasible solution” will be declared. Infeasibility in the operations research world translates into technical impracticability in the remediation world. That a solution is impracticable with have mathematical basis, as opposed to having an analyst search for a solution and when one is not found, having the decision makers ask them to “try harder.” Once infeasibility is declared, it will not matter how long an analyst attempts to find a solution, because none will exist. To find a feasible solution in this case, the constraint relaxation is necessary. That is, either the point of compliance and/or the concentration at the point of compliance needs to be made less strict. The relaxation of the constraints can be found by the algorithm. See [Deschaine, et. al. 2001] for details of the algorithm and an applied solution technique.

In one trial this method solved a problem about 15 times faster than the MINOS algorithm. While it is already a very efficient, effective approach, solution speed can be increased even more if the simulation model can be accelerated. Between 20 and 100 calls to the simulator are common using this method. In one case it was predicted that application of this technique would result in a savings of $72.6 million through the optimization of extraction and injection rates of a currently installed pump-and-treat system.
when compared with the current operational configuration [Karatzas 2001]. Lirschitz Global optimization is described in (Deschaine and Pinter, 2006).

This technology also allows direct integration of the aboveground treatment system cost function with the below ground remediation solution. For example, it can automatically balance whether it is better to extract a lot of water at a low concentration or to extract less water at a higher concentration. This optimization includes the variable treatment cost for various flow rates and concentrations and is contaminant specific.

HORDA

The Heuristic Optimal Remedial Design Algorithm (HORDA) is an algorithm that provides a consistent, traceable, and logical method for determining a least-cost solution to a remediation challenge. The original algorithm was first developed for saturated flow systems [Deschaine 1992] in the early 1990s and since then has been extended (and is continually being enhanced) for multiphase/multicomponent, variably saturated flow and biogeochemically reactive systems. In its simple form, it consists of a set of criteria that systematically directs the answer to the least-cost solution. The basic criteria are listed below:

1. Minimize the travel time of the chemical from its initial location to the in situ or ex situ treatment location. The in situ treatment location can move in space and time, such as in Redox zone optimization. The ex-situ treatment point is often an extraction well.

2. Transport the subsurface chemical material from locations of lower contamination through areas of higher contamination in route to the treatment location. Natural gradients, concentration field manipulation, and/or extraction/injection wells or galleries can accomplish this.

3. Achieve transport migration control to prevent further spread of the area of impact. This control can be accomplished by air sparging curtains, natural or engineered in situ Redox zone optimization, or extraction and injection wells, for example.

4. Minimize the amount of clean water that is treated.

5. Achieve remediation of each location in the area of impact at the same time.

Each of the above criteria will have site-specific importance (i.e., weights) as it relates to the site remedial technology, source, and remediation-cost function. One application of the HORDA algorithm to a multiphase, single-component, chlorinated organic remediation project resulted in a remedial solution that took less than half the time (8 years versus 29 years) and saved more than 40% of the cost ($730,000 versus $420,000) compared to solutions proposed by others [Deschaine et al. 1998a, 2007]. This is a simple methods that remedial design practitioners without extensive optimization training can use to quickly find good – though not provably globally optimal – solutions.

MODOFC

MODFLOW Optimal Flow Control (MODOFC) couples the U.S. Geological Survey MODFLOW simulation program with an optimization algorithm [Ahlfeld and Mulligan 2000]. MODOFC accommodates the linear pumping costs of wells (dollars per amount pumped or injected); installation of each well; and bounds on hydraulic head and head difference, individual and net pumping rates, and total number of wells. MODOFC converts the groundwater flow control problem into an optimization problem and requires MODFLOW to evaluate the fitness of the proposed solution. It is excellent at solving
complex plume stabilization challenges in saturated aquifers. It has saved $100M at a Superfund site when compared to alternative solutions [Ahlfeld 1998].

SIMULATOR REPLACEMENT – FIDELITY OPTIMIZATION

Integrated simulation and optimization typically requires a sequence of “expensive” function calls. While extremely valuable in concept, when the computation cost of simulations functions is high (hours/days) and or the optimization paradigm is inefficient (thousands of function calls), real-time or timely optimal solutions are elusive. The use of machine learning has developed high fidelity model of a process simulator that executes quickly (milliseconds), as opposed to hours. This function is then optimized using the LGO solver, thus enabling optimization in real-time (Deschaine and Pinter, 2003). Developing general functions that approximate simulators using artificial intelligence/machine learning to accelerate optimization is described using neural networks [Peralta 2000] and linear genetic programming [Deschaine and Francone 2002a; Deschaine and Francone 2002b]. Also, this technique has proven successful for high-accuracy unexploded ordnance discrimination, as discussed in (Deschaine, 2006).

INFORMATION INTEGRATION AND SCALABILITY

Scalability is an important issue if we are ever to develop a system then extends beyond “site-wide” to state, country, or global optimal solution. Yet, precisely this capability is needed to respond to the global sustainability challenge. Philosophically, I define global sustainability generally as making decisions today that ensure future generations a comparable Earth.

A start of the process of integrating information globally is discussed in Deschaine, 2000. This document describes an agent-based system that was used to link environmental databases in the United States and Europe to behave coherently as one. By linking these project files together electronically, updates that are made to the individual projects automatically update the program file. This speeds information transfer, provides a strong configuration management system, and reduces error. Linkage can be accomplished using innovative and ever changing Web-based database connectivity tools.

But actually, am not convinced that this is the correct approach to use in all situations. I see this working on static systems at scales where I can go out, query many distributed databases, retrieve the information, assemble it into a matrix, and use it to make decisions. This is fine for slow-moving subsurface systems; however, some of the environmental analysis and protection applications are better represented as dynamic distributed systems in the scale of a country or worldwide. I am currently investigating the Wave-WP paradigm. It may be the system of choice for optimized reaction on non-local disasters, where there are plenty of highly dynamic data at many locations and time. The Wave-WP will allow effective real-time, non-local solutions over a fully distributed data and or knowledge. These challenges include watersheds for public drinking water supplies. The watershed often has many point and non-point sources of contamination, and drinking water can consist of a mixture of surface and groundwater. A watershed is subject to rainfall events, which result in highly variable flow rates and water qualities. The watershed may be subject to instantaneous and intentional contamination. This problem specification requires spatial solutions over dynamic data, wherever the data resides. While the agent system discussed above will work fine for slow-moving groundwater systems, not handle this dynamic condition where optimal solutions are needed in real time. The solution methodology needs to keep the highest integrity as a system, react to rapid changes in the distributed environment, but without the need for central control. I conjecture that the WAVE-WP paradigm is the best approach for this large-scale solutions where dynamics are important. The author of WAVE-WP concurs that this is a valid conclusion and excellent route to pursue (Sapaty, 1999, Sapaty, 2003).

CONCLUSIONS
A formal site-wide environmental impact remedial design optimization system has been developed. It is important to note that “optimization” does not singularly refer to cost minimization, but to the effective and efficient balance of cost, performance, risk, management, and societal priorities and uncertainty. This tool integrates all of these elements into a single decision framework. It provides a consistent approach to designing optimal remediation systems that are tractable, traceable, and defensible. The system is modular and scalable. It can be applied either as individual components or in total. Component deployment has already produced savings between $100M and $72.6M and timesaving of more than half of original expectations, while not sacrificing effective responses to safety of human health or the environment. These results do not indicate that the initial attempts at developing optimal site-wide response policies by others was necessarily bad, but that optimization of complex environmental challenges is very hard to accomplish without the proper integrated value engineering, simulation, and optimization tools. Extensions of the tool to power plant design and UXO discrimination is demonstrated.

REFERENCES


