6 BALANCING PERFORMANCE CRITERIA

6.1 Introduction

This chapter builds upon the previous chapters by combining an effectively designed multiobjective GA [see Chapter 4] with quantile kriging [see Chapter 5] into a highly adaptable LTM design methodology that has significant potential for aiding stakeholders in selecting, understanding, and balancing monitoring objectives for contaminated sites.

6.1.1 Motivation & Scope

Buras (2001) contends that one of the unresolved issues in water resources is the inclusion of multiobjective formulations in the design of engineered systems. Multiobjective problem formulation implicitly requires decision makers to select, understand, and balance performance objectives for the physical systems being designed. The goal of this chapter is to demonstrate that tools such as multiobjective genetic algorithms (GAs) that are capable of high order Pareto optimization (i.e., optimizing a system for more than 2 objectives) can serve as an interface between the design of the physical system and the human decision process. *Buras* (2001) recommends that multi-attribute utility analysis (MAUA) be used in this role. MAUA implicitly requires a preferential weighting of design objectives as well as a conversion of the objective values into the single metric of "utility". The major difficulties encountered when applying MAUA are (1) ensuring that only incommensurate, conflicting objectives are used and (2) assigning a unique utility function for group decision making (i.e., the Arrow paradox) [see *de Neufville* 1990 for more details]. Because of these difficulties, *de Neufville* (1990) states that multiobjective optimization is the preferred solution technique in cases where there are vast numbers of design alternatives. However, multiobjective optimization and MAUA can be combined sequentially; *Horn* (1997) recommends that Pareto optimization be used to quantify a

small set of optimal solutions from which a final solution can be selected using MAUA. This chapter will demonstrate that high order Pareto optimization can provide a means of selecting objectives, discovering objective conflicts, and helping stakeholders in the negotiation process, whether or not combined with MAUA. The highly adaptable optimization methodology proposed in this chapter has significant potential for aiding the design of water resources systems by enabling decision makers to discover, understand, and balance tradeoffs among their design objectives.

The optimization methodology is demonstrated in this chapter using a long-term monitoring (LTM) application. The application addresses the two most important problems LTM practitioners face in the design process: (1) selecting monitoring objectives and (2) balancing these objectives. Both the *ASCE Task Committee on Geostatistical Techniques* (1990b) and *Loaiciga et al*. (1992) concur that the selection of performance criteria is the most important component of any monitoring design methodology. The problem of selecting monitoring performance criteria requires stakeholders to abstract their design preferences into mathematical functions and understand how these functions affect sampling strategies. *Loaiciga et al*. (1992) state that "[o]ne of the key difficulties in the design of ground water monitoring networks via mathematical models is to choose objective functions that faithfully represent a [stakeholder's] objective". Moreover, stakeholders must be able to assess how these mathematical models interact and how these interactions affect the final design of a monitoring system.

For example, there is an obvious conflict between cost and uncertainty. As the number of sample locations used decreases, sampling costs also decrease but uncertainty increases. Now consider uncertainty and contaminant mass estimation error: both quantities increase as the number of sample locations decrease. Does a conflict exist between these objectives? Do both

objectives have a significant effect on the final design of a monitoring network? High order Pareto optimization can serve to answer these questions by enabling stakeholders and regulators to isolate and visualize a small number of sampling strategies that are optimal with respect to multiple objectives. Through visualization of these sampling strategies, stakeholders can discover how their objectives are affecting designs and select only those objectives that best fit their design preferences.

6.1.2 Background on LTM Design

Monitoring network design has been studied extensively in the past, but previous studies have primarily focused on two problems: (1) the use of geostatistics to augment or design monitoring networks for site characterization (for a review, see *ASCE Task Committee on Geostatistical Techniques* 1990a,b) and (2) the use of optimization and numerical simulation to site new monitoring points for contaminant plume detection at landfills and hazardous waste sites (for a review, see *Loaiciga et al*. 1992). Recently, a third problem has emerged that seeks to reduce spatial and temporal redundancies in pre-existing well networks for sites undergoing long term monitoring (LTM). The LTM design methodology proposed in this chapter combines concepts from the geostatistical design and the redundancy analysis studies described below.

The geostatistical approach to monitoring network design utilizes geostatistical estimation procedures (kriging) to evaluate alternative sampling schemes. Geostatistics provides minimum error estimates of contaminant concentrations at unsampled locations using linear combinations of sample values (for an introduction see *Kitanidis* 1997, *Goovaerts* 1997, *Chilès & Delfiner* 1999). In addition to providing the expected values of contaminant concentrations, the estimation variance is computed that represents the uncertainty of estimates at unsampled locations. Estimation variances are independent of the sampled data and vary only as a function

of the monitoring well network's geometry, enabling monitoring network designs to be evaluated for their ability to characterize a site before any sampled data are available.

Hughes & Lettenmaier (1981) recognized the potential of the estimation variance attained from the geostatistical estimation for guiding sampling design. Early studies focused on methodologies that used estimation variance to either augment or design monitoring networks for site characterization (for reviews see *ASCE Task Committee on Geostatistical Techniques* 1990b, *Loaiciga et al*. 1992). Additional sampling points are added based on an analysis of which locations will maximally decrease the estimation variance attained in geostatistical interpolation (*Rouhani* 1985, *Rouhani and Hall* 1988).

Spatial redundancy analysis has garnered more attention as the number of contaminated sites requiring LTM has increased over the previous decade. Sampling-and-analysis costs have been recognized as significant contributors to the overall cost of remediating sites with groundwater contamination. LTM at many sites can require decades of expensive sampling at tens or even hundreds of existing monitoring wells, resulting in hundreds of thousands or millions of dollars for sampling and data management per year. The tremendous costs associated with the collection and management of LTM data has motivated the development of design methodologies that seek to identify and remove sampling locations that contribute minimally to understanding the plume's extent in space, time, or both. Evidence of the emerging importance of redundancy analysis can be seen in the Federal Remediation Technologies Roundtable's focus on tools that directly account for redundancies in monitoring data to reduce sampling costs (see http://www.frtr.gov/optimization/monitoring/).

Previous spatial redundancy approaches can be classified as being either global or local in nature. *Reed et al.* (2000a) present a global redundancy analysis that combined a simple GA,

fate-and-transport simulation, and plume interpolation. The study presents a single objective methodology for monitoring design, in which minimum cost sampling plans are identified that quantify the total (global) contaminant mass as accurately as when all available monitoring locations are utilized. *Reed et al.* (2001) presents a local redundancy analysis that combined deterministic plume interpolation and a multiobjective genetic algorithm (GA) to seek the optimal tradeoff between sampling costs and the relative accuracy of local concentration estimates. Other local approaches have combined trial-and-error analysis with plume interpolation to eliminate spatial redundancies (see *Johnson et al*. 1996, *Cameron & Hunter* 2000, *Aziz et al.* 2000). None of these studies directly quantify how eliminating monitoring points increases uncertainty.

The LTM design methodology proposed in this chapter combines both the spatial redundancy and geostatistical approaches to monitoring design. Quantile kriging and the Nondominated Sorted Genetic Algorithm-II (NSGA-II) are combined to quantify the tradeoffs among the following four performance criteria: (1) cost, (2) the relative accuracy of local concentration estimates, (3) the relative accuracy of global mass estimates, and (4) local uncertainty as measured by kriging estimation variances. These four criteria not only combine spatial redundancy analysis with geostatistical approaches, but they also combine both the local and global spatial redundancy measures used separately in *Reed et al.* (2001) and in *Reed et al.* (2000a), respectively. The test case described in Section 6.2 corresponds to the medium test case described in Chapter 5. Readers familiar with the previous chapter can go directly to Section 6.3.

6.2 Test Case Data

The test case developed for this chapter uses data drawn from a 50 million-node flow-and-transport simulation performed by *Maxwell et al.* (2000). The simulation provided realistic

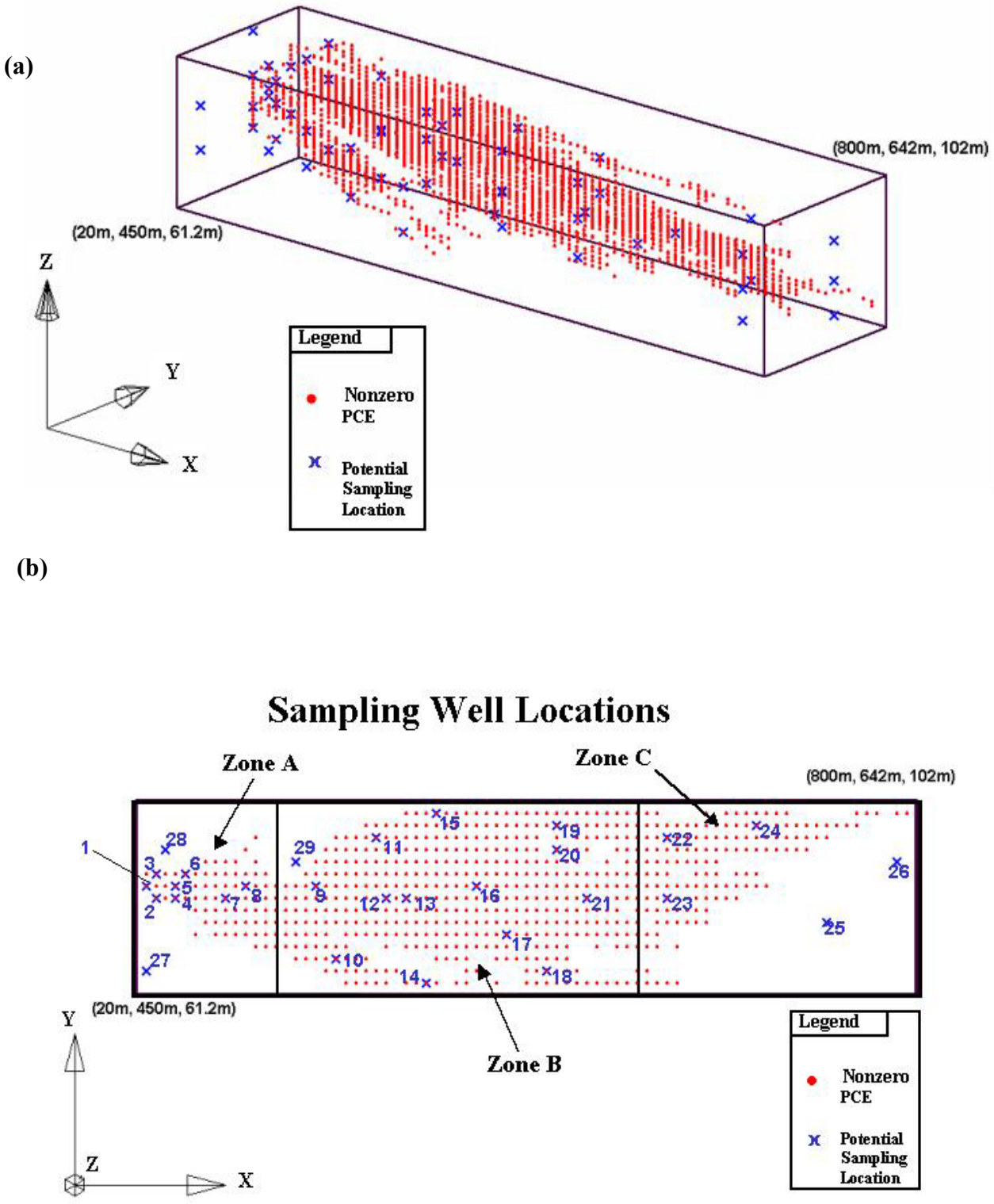


Figure 6.1 (a) 3-dimensional view of the monitoring test case (b) XY-plane view of monitoring wells

historical data for the migration of a hypothetical perchloroethylene (PCE) plume in a highly heterogeneous alluvial aquifer. The hydrogeology of the test case is based on an actual site located at the Lawrence Livermore National Laboratory (LLNL) in Livermore, California. Data were provided for a total of 58 hypothetical sampling locations within a 29-well multi-level monitoring network shown in Figures (6.1a) and (6.1b). Note that Figure (6.1b) and all subsequent XY-views of the plume are divided into zones A, B, and C to show how the sampling designs change in the source area, the midsection, and the leading edge of the plume, respectively. The data represent a snapshot in time, 8 years after an underground storage tank has continuously released contamination into the aquifer system. The monitoring wells can sample from 1 to 3 locations along their vertical axis and have a minimum spacing of 10 m between wells in the horizontal plane.

The site is assumed to be undergoing long-term monitoring, in which groundwater samples are used to assess the effectiveness of current remediation strategies. Quarterly sampling of the entire network has a potential cost of over $85,000 annually for PCE testing alone, which could translate into millions of dollars if the site had a typical life span of 20 to 30 years (*Rast* 1997). This chapter addresses only spatial redundancy, assuming that the spatial sampling plan will be re-evaluated periodically as site conditions change. As noted previously, this type of approach has been applied in several trial-and-error field applications (*Johnson et al*. 1996, *Cameron & Hunter* 2000, *Aziz et al.* 2000).

6.3 Methodology

The methodology has 3 primary components: (1) a high order, constrained multiobjective problem formulation, (2) plume interpolation using quantile kriging, and (3) search and

optimization using the NSGA-II. Each of these components is described in more detail in the following sections of this chapter.

6.3.1 Generalized Multiobjective Formulation

Equation (6.1) presents the generalized multiobjective problem formulation used in this chapter to search for optimal sampling strategies.

$$Minimize\, F(\bar{x}_\kappa) = \left(f_1(\bar{x}_\kappa),\, f_2(\bar{x}_\kappa),\ldots,f_\beta(\bar{x}_\kappa)\right), \quad \forall \kappa \in \Omega$$

Subject to

$$U(\bar{x}_\kappa) = 0 \tag{6.1}$$

In the above equation, $F(\bar{x}_\kappa)$ is a vector-valued performance function in which the β components, $f_1(\bar{x}_\kappa),\, f_2(\bar{x}_\kappa),\ldots,f_\beta(\bar{x}_\kappa)$ are all minimized. Each of the component objectives are functions of the vector $\bar{x}_\kappa$ which represents the κ^{th} sampling plan within the overall decision space Ω. Equation (6.2) defines the i^{th} component of the decision variable vector $\bar{x}_\kappa$ representing the κ^{th} monitoring scheme.

$$x_{\kappa,i}^{samp} = \begin{cases} 1, \textit{ if the } i^{th} \textit{ well is sampled} \\ 0, \quad \textit{otherwise} \end{cases}, \quad \forall \kappa, i \tag{6.2}$$

The vector-valued performance function $F(\bar{x}_\kappa)$ is minimized subject to the constraint shown in equation (6.1), which stipulates that the total number of unestimated points in the interpolation domain $U(\bar{x}_\kappa)$, described in Table 1 for this case chapter, is equal to zero. Kriging contaminant data requires the specification of maximum radii of search for known concentrations in the vicinity of an unestimated point. If no known concentrations are found within these radii of search, then the current grid point remains unestimated. The presence of unestimated points within the interpolation domain prevents reliable calculations of the relative accuracy of local concentration estimates, the relative accuracy of global mass estimates, and uncertainty.

Minimum X-Coordinate	40 meters
Minimum Y-Coordinate	474 meters
Minimum Z-Coordinate	70.8 meters
Size of Grid Blocks in X-Direction	20 meters
Size of Grid Blocks in Y-Direction	24 meters
Size of Grid Blocks in Z-Direction	2.4 meters
Number of Grid Blocks in X-Direction	34
Number of Grid Blocks in Y-Direction	7
Number of Grid Blocks in Z-Direction	7

Table 6.1 Interpolation grid definition

6.3.2 Design Objectives Suite

In this chapter, $F(\bar{x}_\kappa)$ in equation (6.1) contains four component objectives: (1) cost, (2) relative error of local concentration estimates, (3) relative error of global mass estimates, and (4) local uncertainty. Details on each of these four objectives are given below.

6.3.2.1 Cost

The cost of each sampling design is evaluated using equation (6.3) below, which shows that the cost is a discrete function of the number of wells sampled.

$$f_{COST}(\bar{x}_\kappa) = \sum_{i=1}^{nwell} C_S(i) x_{\kappa,i}^{samp} \tag{6.3}$$

Dollar-valued monitoring costs are computed by multiplying a constant coefficient defined in terms of dollars per sample by the total number of samples. In this chapter, costs have been normalized (i.e., the constant cost coefficient has been eliminated) and are presented in terms of the total number of samples required in a given sampling plan. In equation (6.3), if the i^{th} well is sampled then all available locations along the vertical axis of that well will be sampled. The normalized cost, $C_S(i)$ of sampling the i^{th} well ranges between 1 and 3 depending on the number sampling locations along its vertical axis. The total cost of the κ^{th} sampling plan is determined

by summing the product of $x_{\kappa,i}^{samp}$ from equation (6.2) and $C_S(i)$ over the total of *nwell* monitoring wells. In this chapter, there are a total of 29 monitoring wells that sample 58 locations in space.

6.3.2.2 Relative Error of Local Concentration Estimates

The accuracy of local concentration estimates is quantified in terms of squared relative estimation error (SREE) using equation (6.4). SREE is the local redundancy measure that was used in Chapters 3 and 4.

$$f_{SREE}(\bar{x}_\kappa) = \sum_{j=1}^{n_{est}} \left(c_{all}^{*}(\bar{u}_j) - c_{est}^{\kappa}(\bar{u}_j) \right)^2 \tag{4}$$

SREE provides a measure of how the interpolated picture of the plume using data only from wells included in the κ^{th} sampling plan compares to the result attained using data from all available sampling locations. The measure is computed by summing the squared deviations between the local concentration estimates attained using data from all available sampling locations, $c_{all}^{*}(\bar{u}_j)$, and the estimates based on the κ^{th} sampling plan $c_{est}^{\kappa}(\bar{u}_j)$ at each location $\bar{u}_j$ in the interpolation domain. Each $\bar{u}_j$ specifies coordinates for the j^{th} grid point in the interpolation domain.

6.3.2.3 Relative Error of Global Mass Estimates

Reed et al. (2000a) analyzed spatial redundancies using global mass interpolation within a single objective LTM monitoring methodology. The study sought minimal cost sampling plans that would quantify the mass of dissolved groundwater contaminant as accurately as if all available sampling locations had been used. Quantifying the total mass of contaminant within the interpolated domain is equivalent to computing the zero$^{\text{th}}$ moment of the contaminant plume

at time t_s when samples are taken. Equation (6.5) gives relative mass estimation error measure ($f_{MASS}(\bar{x}_\kappa)$) used both in this chapter and by *Reed et al.* (2000a).

$$f_{MASS}(\bar{x}_\kappa) = \left| \frac{Mass_{All} - Mass_\kappa}{Mass_{All}} \right| * 100 \tag{6.5}$$

$Mass_{All}$ is the best mass estimate based on sampling all available monitoring points in the model domain and has a value of 37 Kg for this application. Equation (6.5) is the absolute relative difference between $Mass_{All}$ and the mass estimate, $Mass_\kappa$, computed using the data attained from the κ^{th} monitoring plan $\bar{x}_\kappa$.

6.3.2.4 Local Uncertainty

Local uncertainty is quantified using estimation standard deviations (i.e., the square root of estimation variances) from kriging using equation (6.6).

$$f_{UNCERT}(\bar{x}_\kappa) = \sum_{j=1}^{nest} A\sigma(\bar{u}_j) \tag{6.6}$$

The local uncertainty measure $f_{UNCERT}(\bar{x}_\kappa)$ is a weighted sum of the *nest* estimation standard error $\sigma(\bar{u}_j)$ for location $\bar{u}_j$ in the interpolation domain. For kriging-based estimation, *Journel & Rossi* (1989) showed that the estimation variance is a variogram model dependent ranking of sampling configurations. Chapter 5 shows that kriging estimation variances can successfully predict which areas of the plume have the highest estimation errors. The underlying goal of equation (6.6) is to identify solutions that sample subsets of the 58 total sampling locations while minimally increasing local uncertainty. Readers familiar with both quantile kriging and the NSGA-II should proceed directly to Section 6.3.4.3, which discusses how constraint violations were penalized.

6.3.3 Plume Interpolation using Quantile Kriging

Quantile kriging was selected for plume interpolation in this chapter based on the findings of Chapter 5, which presents a comprehensive performance analysis of 6 interpolation methods for scatter-point concentration data, ranging in complexity from intrinsic kriging based on intrinsic random function theory to a traditional implementation of inverse-distance weighting. Quantile kriging was shown to be the most robust and least biased of the interpolation methods they studied. Additionally, the method's non-parametric uncertainty estimates successfully predicted zones of high estimation error for each test case.

Quantile kriging is a transformation-based variant of the kriging estimator in which the i^{th} concentration sample is assigned the probability that its value is less than or equal to all of the remaining concentration samples. These probabilities create an empirical cumulative distribution function for concentration for a given site and are known to be uniformly distributed (*Hogg and Tanis* 1997). For this reason, the constant A shown in equation (6.6) was assigned a value equal to $2\sqrt{3}$ for computing local uncertainty estimates for a uniformly distributed quantity (for details see *Juang et al.* 2001). The term $2\sqrt{3}\,\sigma(\overline{u}_j)$ in equation (6.6) then computes the distance between the upper and lower bound estimates from the 95th confidence interval for concentration at location $\overline{u}_j$ within the interpolation domain. For more details on quantile kriging see *Journel & Deutsch* (1997), *Juang et al.* (2001), and Chapter 5. For more details on the spatial modeling used in the test case presented in this chapter, see Chapter 5 and Appendix B.

6.3.4 Multiobjective Search & Optimization

NSGA-II is used to identify high order Pareto surfaces in the LTM methodology. NSGA-II is a second generation evolutionary multiobjective GA developed by *Deb et al.* (2000). It significantly improves upon the original NSGA by (1) invoking a more efficient

nondomination sorting algorithm, (2) eliminating the sharing parameter, and (3) adding an implicitly elitist selection method that greatly aids in capturing high order Pareto surfaces. *Zitzler et al.* (2001) and *Deb et al.* (2001) show that the NSGA-II performs as well or better than the other second generation evolutionary multiobjective algorithms on difficult, high order problems. The following section details basic principals and operators of the NSGA-II.

6.3.4.1 NSGA(II) Basics

Genetic algorithms search a decision space using a process that is analogous to Darwin's "natural selection". The decision variables are first encoded as 0-1 binary strings, or chromosomes. The fitness of each member of a randomly generated initial population of these strings is determined by how well the design satisfies specified objectives and constraints. After each individual is assigned a fitness value, GAs find optimal solutions using three basic operators: (1) selection, (2) crossover (mating), and (3) mutation.

The NSGA-II uses a two-step selection process, which combines both binary tournament selection and $(\mu + \lambda)$ selection. First, tournament selection allows only the fittest individual from a group of strings randomly drawn from the current population to be placed into the mating population. Next, the crossover operator couples members of the mating population to mate with a specified crossover probability (P_c). Mating consists of randomly selecting one or more crossover points at which the strings exchange bit values with each other. Tournament selection and crossover are repeated until a population of *N* children has been produced. Mutation then randomly flips binary bits from 1 to 0 or vice versa within the new child population with a given probability of mutation (P_m). Lastly, the NSGA-II uses $(\mu + \lambda)$ selection to choose which of the parent and child designs will survive. In this selection scheme, the populations of *N* parent designs and *N* child solutions in the current generation *t* are combined to yield a selection pool of

2N individuals, from which the *N* best individuals are allowed to pass to generation *t+1*. This selection method aids the algorithm in efficiently identifying high order Pareto surfaces because it is implicitly elitist (i.e., the best designs are guaranteed to survive into the next generation).

These three operators act to create a new population (or generation) of individual sampling plans with improved average fitness. The Schema Theorem is the general theory describing how these three operators combine to evolve high quality near-optimal solutions [see *Goldberg*, 1989 or *Holland*, 1975 for more information]. It states that highly fit strings are composed of small chunks of information (or building blocks) that are relevant to the solution of the problem. The GA exerts a selection pressure where only highly fit members are allowed to pass their traits or building blocks to the next generation. Highly fit parent strings are allowed to mate, yielding offspring that inherit building blocks from both parents. It is in this manner that the GA assembles optimal or near-optimal solutions to a problem.

The primary difference between the NSGA-II and single objective GAs is in how fitness is assigned. The NSGA-II evaluates sampling designs in terms of a vector of objectives. A sampling design cannot be assessed in terms of its performance in any single objective because it may perform poorly with respect to the remaining objectives. Instead, the concepts of *Pareto dominance* and *crowding* are used to assign fitness values to sampling designs in the two steps described below (*Deb et al.* 2000).

The first step in fitness assignment employs the concept of *Pareto dominance* in which a design $\bar{x}$ dominates another design $\bar{x}'$ if and only if it performs as well as $\bar{x}'$ in all n_{obj} objectives and better in at least one. In the NSGA-II's improved nondomination sorting, the algorithm first ranks each design by the number of designs that dominate it. The second step in fitness assignments utilizes the concept of *crowding* (for more details see *Deb et al.* 2001) in

which the average distance between an individual design and those designs within the population that have been assigned the same rank is computed. The crowding operator eliminates the sharing parameter calculations (for more details see Chapter 3) that were required for the original form of the algorithm.

After these two steps, the fitness value of an individual design can be assigned either its rank or its crowding distance as follows. In cases were two designs have different ranks, the individual with the lower rank is preferred (i.e., individuals that are dominated by fewer solutions). Alternatively, if both solutions possess the same rank then the individual with larger crowding distance is preferred (i.e., the individual that adds the most diversity to the population). A diverse population ensures that the NSGA-II will find solutions along the full extent of the Pareto surface.

6.3.4.2 Design & Parameterization

Chapter 4 introduces a multi-population approach for automating parameter specification for the NSGA-II. The methodology combines concepts from previous GA design methodologies (*Reed et al.* 2000b, *Reed et al.* 2001) and the "parameter-less GA" methodology presented by *Lobo* (2000). The methodology utilizes GA design theory to automatically set the probabilities of crossover and mutation as well as the maximum number of generations. The probabilities of crossover and mutation are set equal to 50 percent and $1/N$, respectively, where N is the population size. The maximum number of generations was set equal to 60. Four runs with increasing population sizes from 500 to 4000 members were completed to identify the nondominated set. The runs were halted automatically when further increases in population size resulted in less than a 10 percent increase in the number of nondominated solutions identified. See Chapter 4 for more details.

6.3.4.3 Penalizing Constraint Violations

The constrained problem formulation given in equation (6.1) requires that the number of unestimated points $U(\bar{x}_\kappa)$ be equal to zero. For any design with unestimated points, the penalty function shown in equation (6.7) was used in place of the fitness function in equation (6.1) to ensure that the solution would be dominated by feasible solutions (i.e., any solution with no unestimated points).

$$F_{PENALTY}(\bar{x}_\kappa) = \begin{pmatrix} f_{COST}^{PENALTY} = f_{COST}^{MAX} + f_{COST} \\ f_{SREE}^{PENALTY} = f_{SREE}^{MAX} + f_{SREE} \\ f_{MASS}^{PENALTY} = f_{MASS}^{MAX} + f_{MASS} \\ f_{UNCERT}^{PENALTY} = f_{UNCERT}^{MAX} + f_{UNCERT} \end{pmatrix}^T, \; given\; U(\bar{x}_\kappa) > 0 \qquad (6.7)$$

Equation (6.7) shows that the 4 component functions of $F_{PENALTY}$ are the sum of the maximum possible values for the objectives and their actual values from equations (6.3) – (6.6), using only those points where estimates exist. The maximum values of every objective are used because every feasible design is guaranteed to have lower objective values and dominate these solutions. The actual objective values are important to allow the infeasible solutions to compete with one another, even though the estimates lack validity, because feasible solutions can be very near infeasible solutions in decision space and could be lost if the infeasible solutions are immediately removed from consideration.

The maximum value of the normalized cost in this chapter is equal to 58, which is the total number of sampling locations that are available. The use of quantile kriging made computing the maximum value of SREE, f_{SREE}^{MAX}, very easy because concentrations are transformed to have values in the interval [0, 1], so the maximum squared deviation at each grid point is 1. This results in a maximum SREE value equal to the number of nodes in the interpolation domain (or *nest* = 1666). f_{UNCERT}^{MAX} was attained by assuming all kriging estimation

variances within the interpolation domain were equal to 1, their maximum possible value. This simplifies equation (6.6) to be the product of *nest* = 1666 and the constant $A = 2\sqrt{3}$, resulting in a maximum value of uncertainty, f_{UNCERT}^{MAX}, equal to 5770. The maximum relative mass estimation error, f_{MASS}^{MAX}, was assumed to be 3300. This value was computed using $Mass_{All}$ and the mass estimate that would be attained if every location in the interpolation domain had a concentration value equal to 6500 mg per m^3, which is the maximum concentration in the monitoring dataset used in this chapter.

6.4 Results & Discussion

Recall that if the i^{th} monitoring well is selected for sampling then PCE is sampled at all the possible sampling locations along its vertical axis. This is equivalent to monitoring a single constituent at 29 monitoring wells, which results in a decision space of more than 500 million possible sampling designs (i.e., 2^{29} sampling designs). Using the NSGA-II to identify the subset of sampling designs that are optimal with respect to cost, SREE, mass estimation error, and uncertainty reduces the set of designs that must be considered from 500 million to 1156 designs identified on the Pareto surface. Although the 4-dimensional Pareto surface cannot be visualized, the set of 1156 designs can inform decision making as follows. Initially, interactions and conflicts among the design objectives are identified using a variety of visualizations derived from the full set of nondominated solutions. Once these interactions and conflicts are discovered, they are then used to identify acceptable objective bounds and negotiate a compromise design. This decision process is illustrated for the LTM design application in the following sections.

6.4.1 Understanding the Interactions of Design Objectives

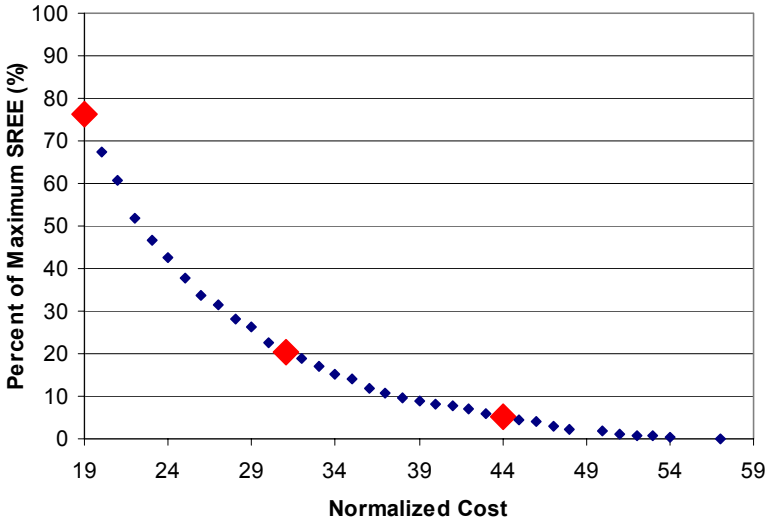
This section of the chapter illustrates the process of learning how cost, SREE, mass estimation error, and uncertainty interact and affect sampling designs. This process begins by

analyzing pairs of the objectives that are known to conflict. These 2-dimensional tradeoffs are subsets of the overall 1156 member nondominated set. These tradeoffs are found by identifying only those solutions that are nondominated in terms of cost and one other objective, independent of the remaining objectives' values. Three designs from each of these tradeoffs have been visualized to demonstrate their effect on sampling schemes as well as to promote the discovery of any additional objective conflicts. The next two sections demonstrate this process of discovery for the LTM application.

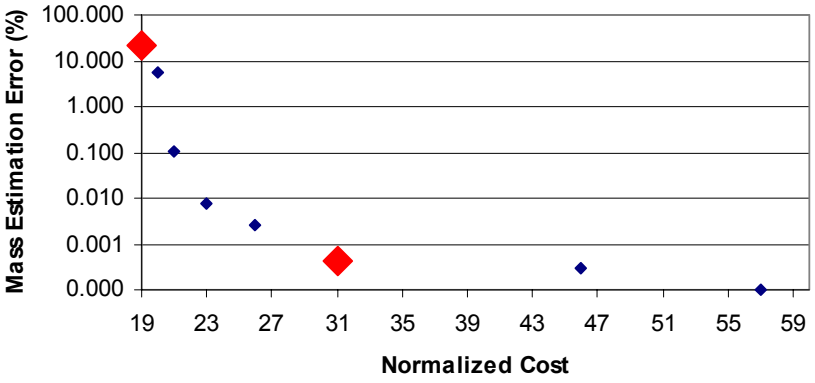
6.4.1.1 Visualizing Known Objective Conflicts

Figures (6.2a) – (6.2c) present tradeoff curves for the following pairs of known conflicting objectives: (1) Cost—SREE, (2) Cost—Mass, and (3) Cost—Uncertainty. The solutions on these figures represent the lowest possible SREE, mass error, and uncertainty values attainable for each level of sampling cost. Both the Cost—SREE and Cost—Uncertainty tradeoffs in Figures (6.2a) and (6.2c) sample between 19 and 58 locations. Figure (6.2b) shows that there are significantly fewer nondominated solutions in the Cost—Mass tradeoff, which is presented using a log scale because mass estimation errors ranged over 5 orders of magnitude. The gaps present in Figure (6.2b) are not due to a shortcoming of the NSGA-II; they result because nondominated sorting considered only cost and mass estimation error. For example, the gap that exists between 31 and 46 sample solutions occurs because the 31 sample solution dominates all of the solutions that sample less than 46 locations (i.e., it is lower in both cost and mass estimation error). Only 21 sampling locations were required to attain a mass estimate that fell within 0.1 percent of $Mass_{All}$, showing that this objective is far less constraining than either SREE or uncertainty.

(a)



(b)



(c)

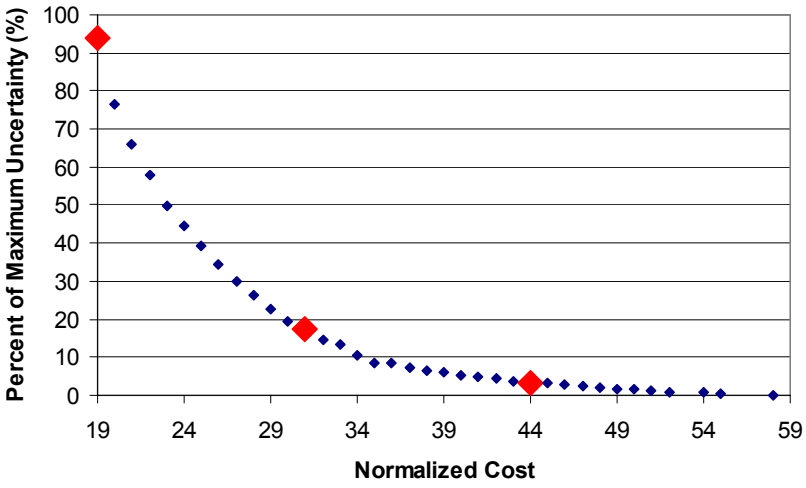


Figure 6.2 (a) Cost—SREE tradeoff (b) Cost—Mass tradeoff (c) Cost—Uncertainty tradeoff

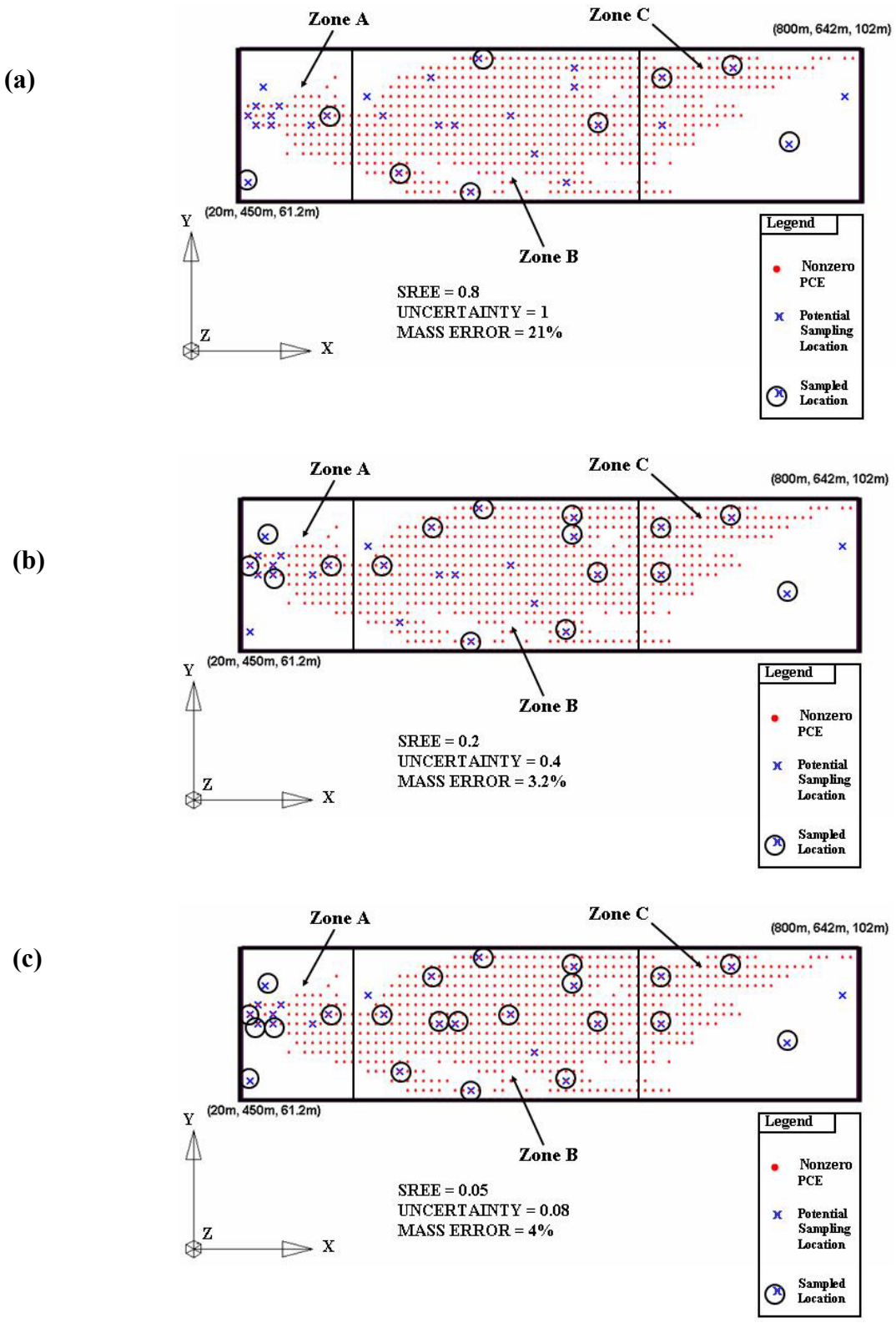


Figure 6.3 Sampling locations corresponding to solutions from the Cost—SREE tradeoff (a) 19 sample solution (b) 31 sample solution (c) 44 sample solution

The solutions highlighted in the tradeoffs of Figure (6.2) sample between 19 and 44 locations and are visualized in Figures (6.3) – (6.5) to investigate how the known objective conflicts are affecting sampling geometries. Comparison of the 19 sample solutions in Figures (6.3a) and (6.5a) show that the least cost solution for the Cost—Uncertainty tradeoff is very similar to the least cost solution for the Cost—SREE curve. The 19 sample solution for the Cost—Mass curve is identical to the 19 sample solution for the Cost—SREE curve and is not shown. The only differences between the solutions in Figures (6.3a) and (6.5a) lie in Zone B, where well 17 is sampled in place of well 15, and in Zone C where well 23 is sampled in place of well 22, resulting in a 6 percent reduction in the uncertainty measure. Chapter 5 shows for the same test case presented in this chapter that the zone of highest uncertainty is located near X = 400m and Y = 550m, very near wells 17 and 23, which explains why these sampling locations helped to reduce uncertainty.

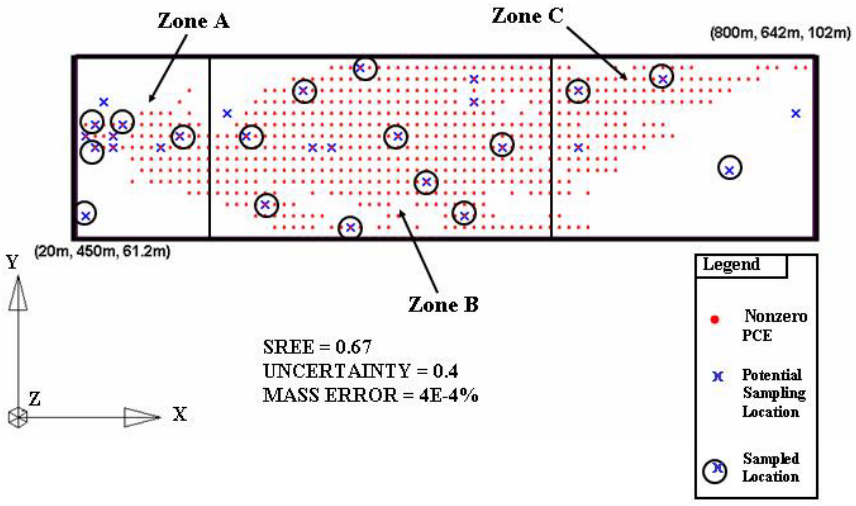


Figure 6.4 Sampling locations corresponding to the 31 sample solution from the Cost—Mass tradeoff

Figures (6.3b), (6.4), and (6.5b) present the 31 sample, nondominated solutions from the Cost—SREE, Cost—Mass, and Cost—Uncertainty tradeoffs, respectively. The 31 sample

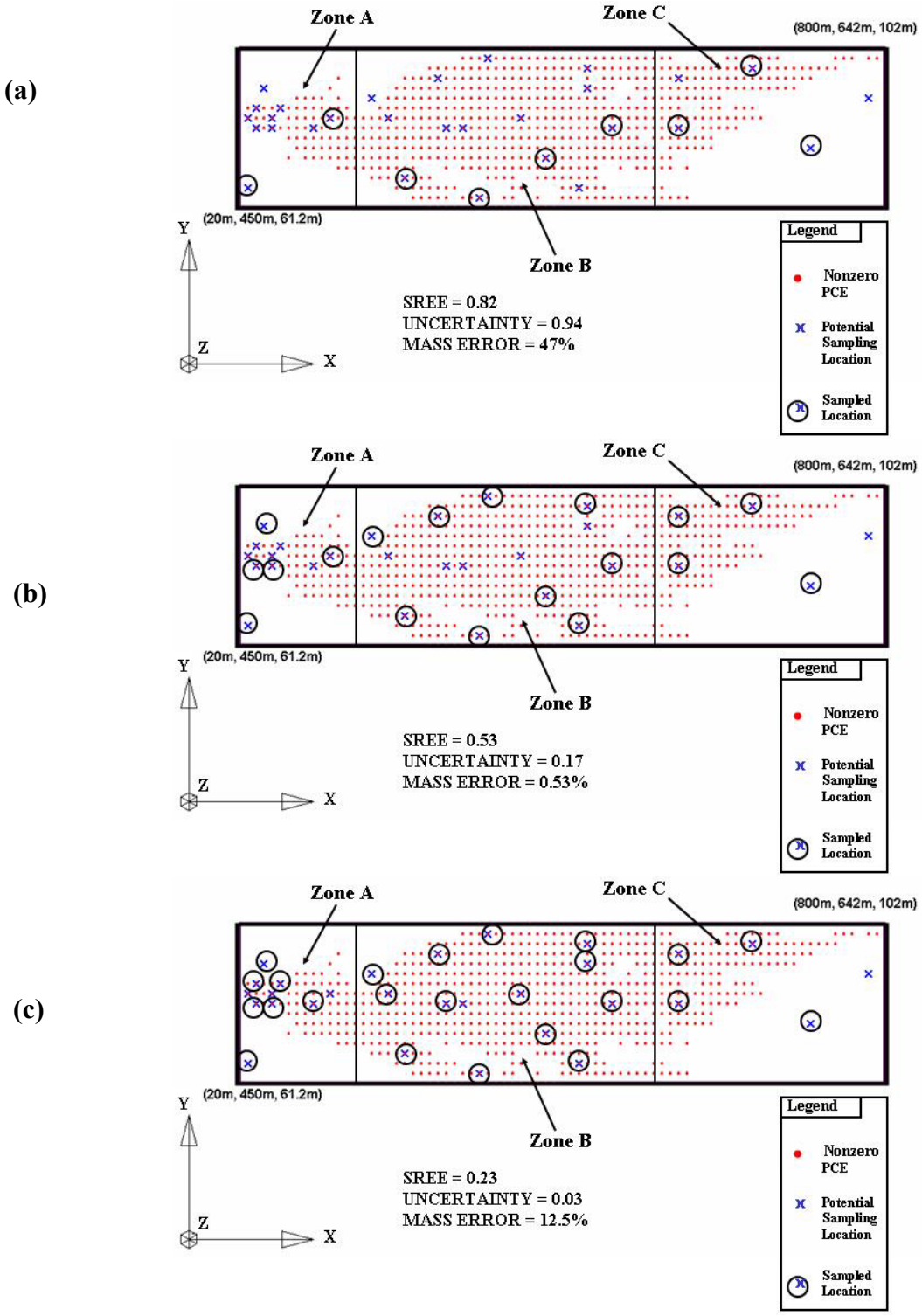


Figure 6.5 Sampling locations corresponding to solutions from the Cost—Uncertainty tradeoff (a) 19 sample solution (b) 31 sample solution (c) 44 sample solution

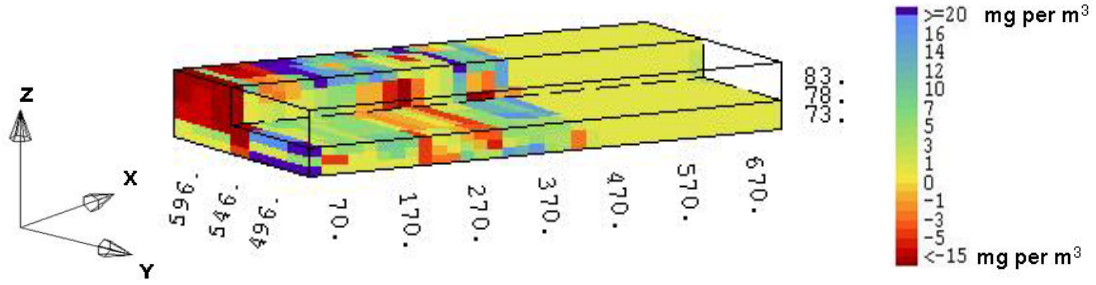
solutions all affirm that sampling wells along the plume's outer boundaries and near its leading edge are important for reducing SREE, mass error, and uncertainty. The 44 sample solutions presented in Figures (6.3c) and (6.5c) from the Cost—SREE and Cost—Uncertainty tradeoffs, respectively, show that both SREE and uncertainty are further reduced by sampling areas of high concentrations within Zone A and wells within the interior of Zone B.

As expected, the 44 sample solution shown in Figure (6.5c) uses more locations to decrease uncertainty, but note the surprising increase in mass estimation error relative to the 31 sample solution in Figure (6.5b). A very interesting multi-well interaction between well numbers 1, 27, and 28 [see Figure (6.1b)] causes this unexpected increase in mass estimation error. These wells sample the minimum and maximum concentrations within Zone A. Specifically, well 1 provides concentration values that exceed 4500 mg per m^3 while wells 27 and 28 sample locations where there is no PCE. All of the nondominated designs that sample 44 locations and include wells 27 and 28 but not well 1 resulted in mass estimation errors that exceed 9 percent. The increased mass estimation error results because wells 27 and 28 in the absence of well 1 cause the mass in Zone A to be severely underestimated. This result identifies an Uncertainty—Mass Error conflict. The next section demonstrates how visualization can be used to identify additional objective interactions.

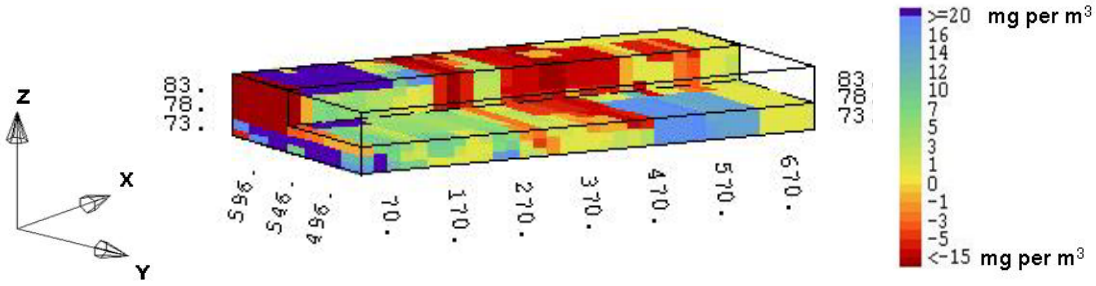
6.4.1.2 Discovering and Understanding Additional Objective Interactions

For the sake of illustration, the 31 sample cost level is used in this chapter to identify potential conflicts among SREE, mass error, and uncertainty. Figures (6.3b), (6.4), and (6.5b) presented the 31 sample, nondominated solutions from the Cost—SREE, Cost—Mass, and Cost—Uncertainty tradeoffs, respectively. To improve the clarity of the subsequent comparative discussion of these designs, they are referred to as SREE(31), Mass(31), and Uncertainty(31).

(a) SREE = 20 percent of maximum value



(b) SREE = 67 percent of maximum value



(c) SREE = 53 percent of maximum value

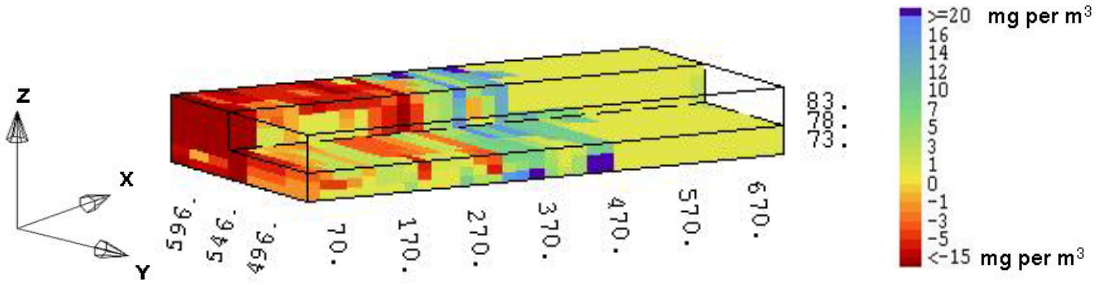


Figure 6.6 Three dimensional cross sectional views of how SREE changes for (a) the SREE(31) solution (b) the Mass(31) solution (c) the Uncertainty(31) solution

SREE increases from its minimum value of 0.2 in the SREE(31) solution to 0.67 and 0.53 for the Mass(31) and Uncertainty(31) solutions, respectively. Although the SREE performance measure provides a ranking of the sampling designs according to the relative accuracy of their local estimates, it does not provide a clear understanding of what these relative rankings

represent. Figures (6.6a) – (6.6c) provide a direct assessment of how the sampling geometries presented in Figures (6.3b), (6.4), and (6.5b) are affecting SREE. Figure (6.6) visualizes how the relative local errors in SREE [i.e., $-\left(c^{*}_{all}(\overline{u}_j) - c^{\kappa}_{est}(\overline{u}_j)\right)$ in equation (6.4)] vary throughout the interpolation domain in terms of milligrams per m^3. Positive values represent the case where the PCE concentration estimate $c^{\kappa}_{est}(\overline{u}_j)$ is overestimated relative to the estimate attained using all 58 available sampling locations. Conversely, negative values represent underestimates. For SREE, more than 50 percent of the $c^{*}_{all}(\overline{u})$ values were less than 12 mg per m^3 and 70 percent where less than 20 mg per m^3, hence any local errors exceeding ± 5 mg per m^3 represent significant increases. The 31 sample solutions visualized in Figure (6.6) clearly illustrate significant local errors, which motivated the selection of higher cost 44 sample designs in the negotiation stage of this methodology (described below) to reduce local errors within acceptable limits.

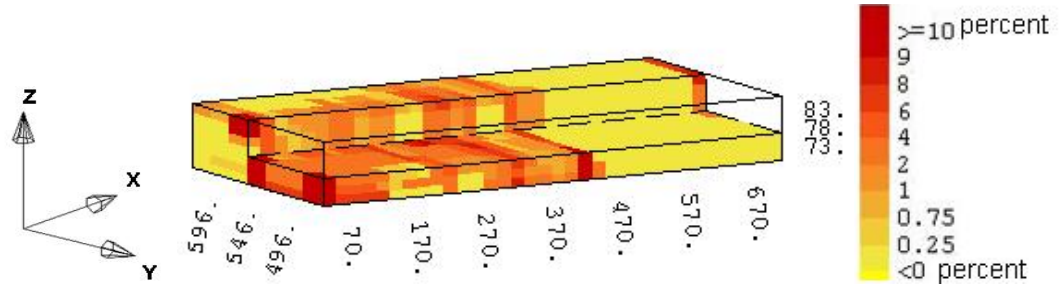
Figure (6.6a) shows that the majority of local errors for the SREE(31) solution are found in Zones A and B of the interpolation domain, while the local errors are nearly zero in the leading edge of the plume. The SREE(31)'s sampling plan in Figure (6.3b) shows 4 of the 5 the wells nearest to the leading edge of the plume are sampled while several wells within Zones A and B are unsampled, resulting in increased local estimation errors in these areas of the domain. Although this design has the lowest SREE value of the three 31 sample solutions being considered, it has the highest mass estimation error at 3.2 percent, showing a potential SREE—Mass tradeoff. This result is surprising because minimizing local estimation error throughout the interpolation domain would seem to implicitly result in accurate mass estimates. This contradiction can be understood by looking at the Mass(31) solution, which has negligible mass estimation error while having the highest SREE value of the three 31 sample solutions. Figure

(6.6b) visualizes the Mass(31) solution's local estimation errors and provides insight into how the SREE and mass error objectives conflict. The figure shows that local concentrations are significantly overestimated in the source area of the plume (Zone A) and severely underestimated in large portions of Zones B and C. In effect, the sampling plan shown in Figure (6.4) balances both overestimation and underestimation throughout the plume to yield a mass estimate with a coincidentally high relative accuracy while severely sacrificing the quality of the interpolated map of PCE (i.e., SREE). Severe overestimation and underestimation of contaminant concentrations can have very negative consequences in correctly assessing risk associated with a site. These results show that global measures such as mass should be counter-balanced with localized measures such as SREE and uncertainty.

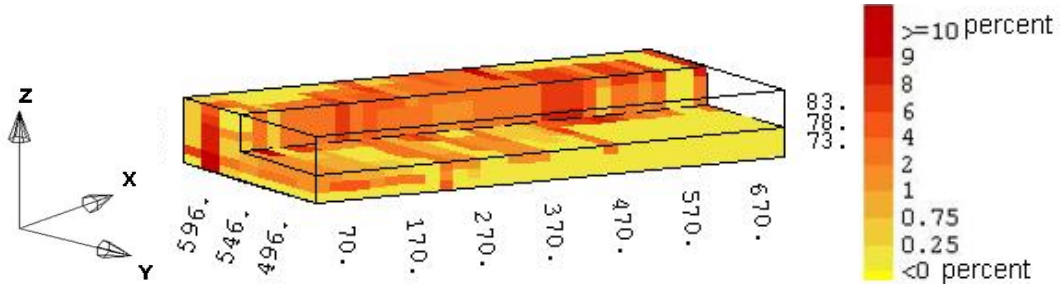
Figures (6.7a) – (6.7c) show how the local uncertainty estimates change in each of the three 31 sample solutions. These figures show the absolute percentage change of $\sigma(\overline{u}_j)$ for the j^{th} location within the interpolation domain [see equation (6.6)]. It should be noted that a small percentage change in the kriging standard error [i.e., $\sigma(\overline{u}_j)$] for the j^{th} location yields potentially huge changes in uncertainty calculations. For example, consider a location $\overline{u}_j$ with a mean concentration estimate equal to 7 mg per m^3 and kriging standard error $\sigma(\overline{u}_j)$ = 0.2 (typical for this case study). A 4 percent increase in $\sigma(\overline{u}_j)$ causes the 95th percentile estimate for location $\overline{u}_j$ to increase from 182 mg per m^3 to 482 mg per m^3, representing a 260 percent increase.

Although the Uncertainty(31) solution has the minimum uncertainty value of the three 31 sample solutions, it has increased SREE and mass estimation errors relative to the SREE(31) and Mass(31) solutions. The SREE—Uncertainty and Mass—Uncertainty conflicts occur because both SREE and mass error are heavily affected by extreme concentrations, especially

(a) Uncertainty = 40 percent of maximum value



(b) Uncertainty = 45 percent of maximum value



(c) Uncertainty = 17 percent of maximum value

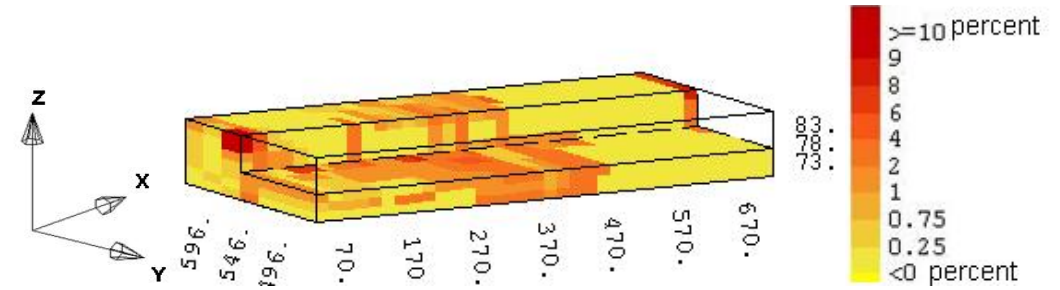


Figure 6.7 Three-dimensional cross sectional views of how Uncertainty changes for (a) the SREE(31) solution (b) the Mass(31) solution (c) the Uncertainty(31) solution

within the plume's source area, while uncertainty is independent of the concentration values and is solely a function of how well the sampling locations are distributed throughout the interpolation domain. Figure (6.7b) shows that the Mass(31) solution greatly increased uncertainty throughout the interpolation domain relative to the Uncertainty(31) solution in Figure (6.7c). Figures (6.6c) and (6.7a) confirm the SREE—Uncertainty conflict. The Uncertainty(31)

solution in Figure (6.6c) severely underestimates PCE concentrations within the source area of the plume, resulting in its increased SREE value of 0.53, while Figure (6.7a) shows that the SREE(31) solution significantly increased uncertainty within Zones A and B of the plume. The increased local uncertainties for the 31 sample solutions from the Cost—SREE and Cost—Mass tradeoffs resulted because they lack samples along the plume's outer boundary or near the zone of highest uncertainty centered at X = 400m and Y = 550m [see Figures (6.3)-(6.5)].

6.4.2 Negotiating a Final Design

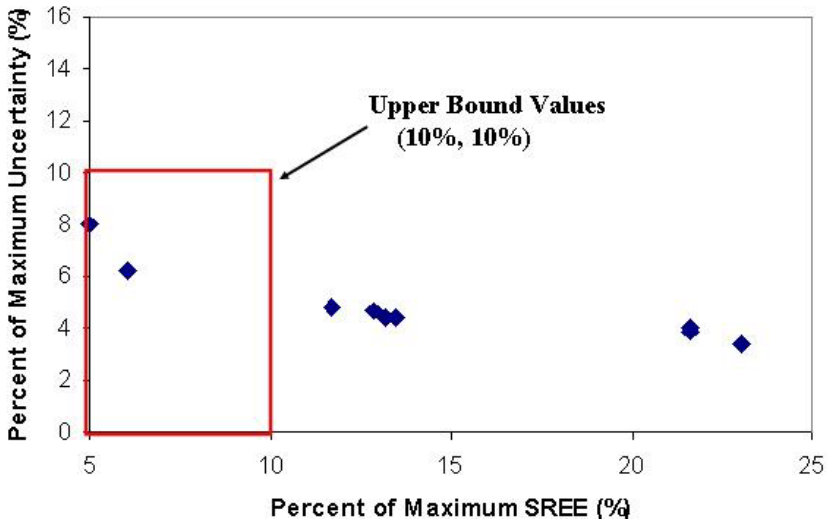
Recall that the LTM application presented in this chapter has more than 500 million possible sampling designs. High order Pareto optimization using the NSGA-II reduced the set of designs to be considered to the 1,156-member nondominated set that are optimal in terms of cost, SREE, mass error, and uncertainty. The previous sections visualized a total of 7 sampling schemes to provide stakeholders and regulators with an improved understanding of their design objectives. The next two sections provide an illustrative example of how stakeholders and regulators can then negotiate a final compromise sampling scheme.

6.4.2.1 Bounding Stakeholder Expectations

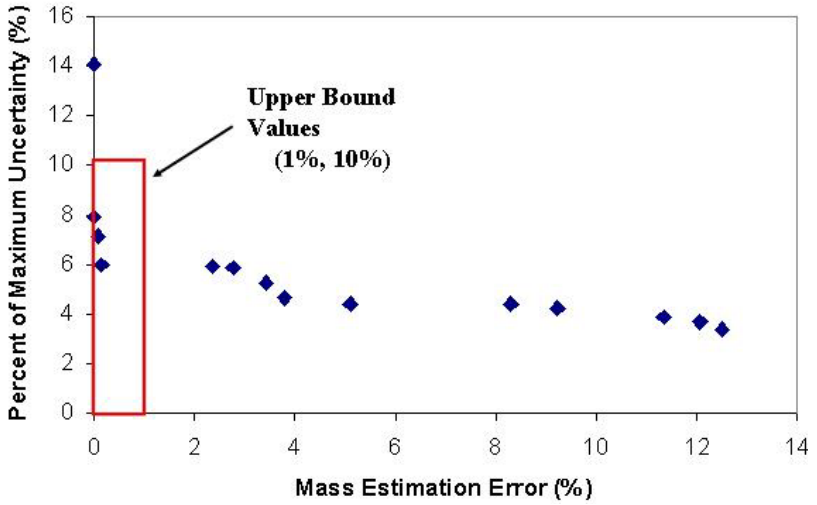
The first decision that faces stakeholders is selecting acceptable bounds on each objective. For this example, a conservative cost level of 44 sampling locations was selected to reduce costs by nearly 25 percent while minimally increasing the remaining objectives. Note that this example merely illustrates the negotiation process and that higher or lower cost levels could be selected. By considering only those designs at the 44 sample cost level, the set of potential designs further reduces from 1156 to 46 potential monitoring designs.

The next step in the negotiation process is to use the objective conflicts that occur in the 44 sample designs to bound stakeholder expectations and set "acceptable" upper bound values

(a)



(b)



(c)

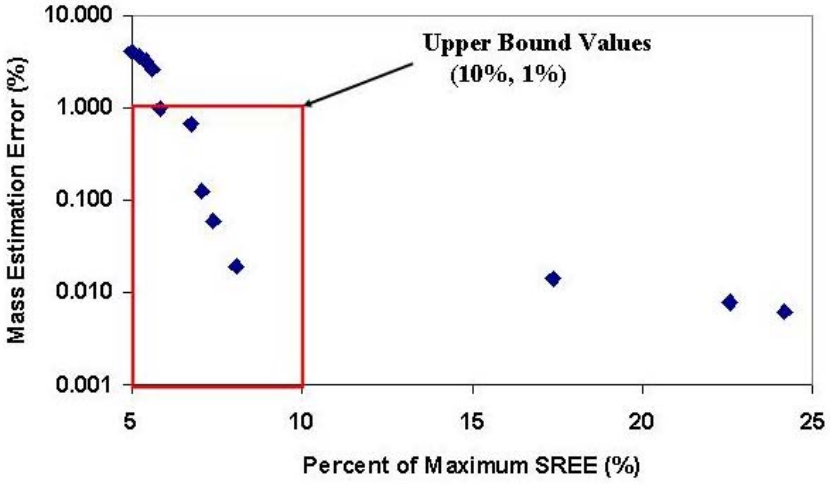


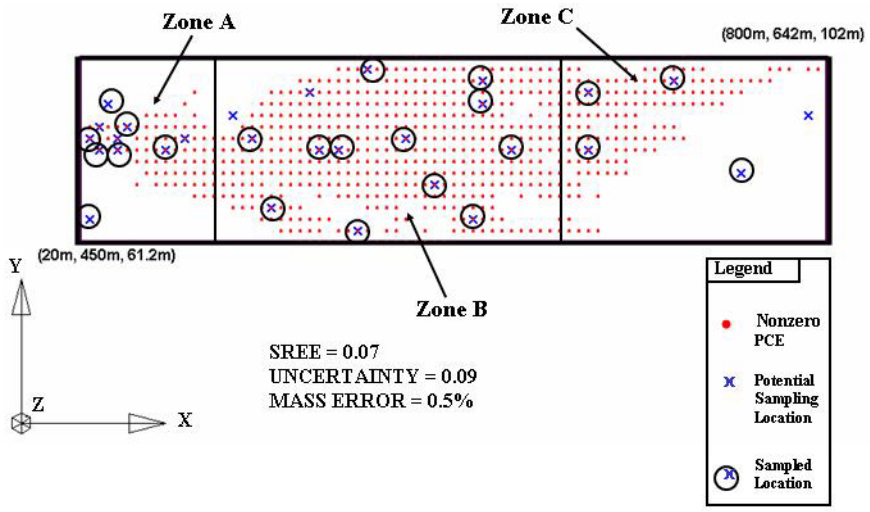
Figure 6.8 Objective conflicts between (a) SREE and Uncertainty (b) Mass Error and Uncertainty (c) SREE and Mass Error for the 44 sample solutions

for the remaining objectives. The SREE—Uncertainty tradeoff shown in Figure (6.8a) shows that increasing SREE beyond 10 percent of its maximum value in the nondominated set has a negligible effect on reducing uncertainty. The Mass—Uncertainty tradeoff presented in Figure (6.8b) shows that uncertainty can be significantly reduced with less than a 1 percent increase in mass estimation error. Lastly, the SREE—Mass tradeoff in Figure (6.8c) shows that a very significant decrease in mass estimation error can be attained for a relatively small increase in SREE. Given these findings, upper bound values for SREE, mass error, and uncertainty were set equal to 10%, 1%, and 10%, respectively, as shown in Figure (6.8).

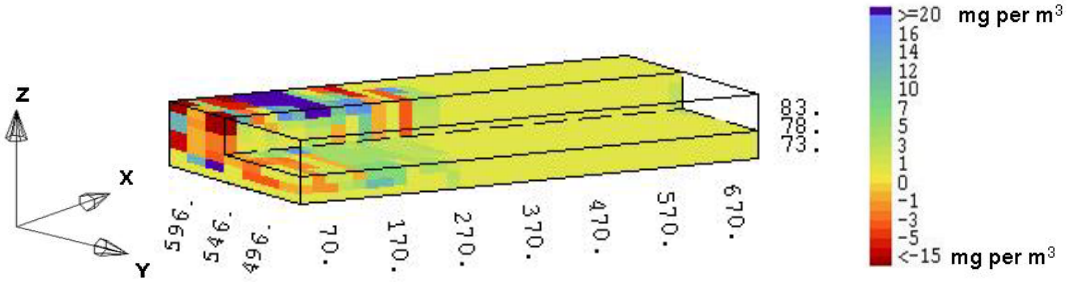
6.4.2.2 Striking the Balance

The final step in the negotiation process is to search the nondominated set for designs that satisfy the objective bounds [i.e., (Cost = 44, SREE < 10%, Mass Error < 1%, Uncertainty < 10%)]. Setting these objective bounds reduces the number of possible sampling designs that must be considered from 46 to the single compromise solution illustrated in Figure (6.9). Although setting objective bounds will not always yield a single solution, it will vastly limit the number of designs that must be considered. The solution shown in Figure (6.9a) reduces sampling costs by nearly 25 percent in any given monitoring period while minimally increasing uncertainty, maintaining a high quality map of the plume, and accurately quantifying the mass of PCE within the subsurface. Figure (6.9b) shows that the majority of the local PCE estimation errors throughout the plume were less than 5 mg per m^3. Figure (6.9c) shows that local uncertainty estimates remained unchanged for a large portion of the plume. Although these observations are important, the most important contribution of Figure (6.9) is how the figure provides regulators and stakeholders with a direct understanding of the effects of their negotiated objective bounds.

(a)



(b) SREE = 7 percent of maximum value



(c) Uncertainty = 9 percent of maximum value

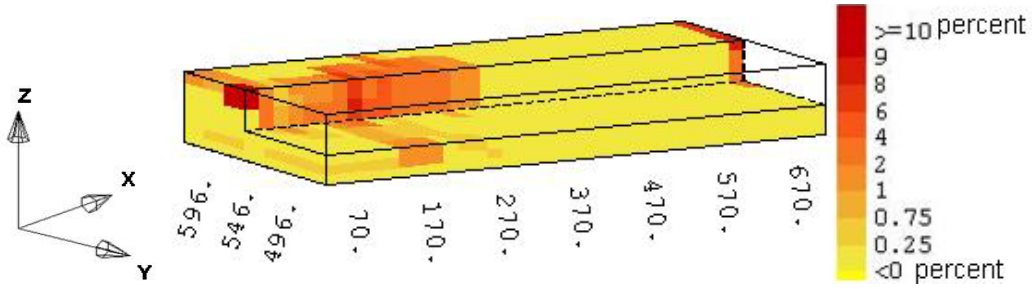


Figure 6.9 Compromise solution's (a) XY-plane view of sampling locations (b) three dimensional cross section of how SREE changes (c) three dimensional cross section of how Uncertainty changes

Current standard practice in redundancy analysis uses trial-and-error analysis to eliminate sampling locations (see *Johnson et al*. 1996, *Cameron & Hunter* 2000, *Aziz et al.* 2000). In these

methodologies, locations thought to be redundant are eliminated and visualization is used to determine the effect of these locations on the quality of the interpolated plume map. The process is repeated for tens if not hundreds of designs until the practitioner is satisfied. This time consuming process does not comprehensively search the decision space or account for multiple objectives. The single compromise solution shown in Figure (6.9) explicitly balances the stakeholders' objectives, required less than a day of computing time, and visualization of only 8 designs. Moreover, the optimization methodology used to attain the compromise solution provides practitioners with a better understanding of how their design preferences interact with the physical monitoring system.

6.5 Conclusions

The optimization methodology presented in this chapter demonstrates that algorithms such as the NSGA-II that are capable of high order Pareto optimization can serve as interfaces between the human decision process and engineered water resources systems. The LTM application used to demonstrate the methodology shows how multiobjective optimization combined with visualization can aid practitioners in selecting, understanding, and balancing these performance objectives when seeking a single compromise solution. The monitoring application successfully balances the following four objectives: (1) minimizing sampling costs, (2) maximizing the quality of interpolated plume maps (i.e., SREE), (3) maximizing the relative accuracy of contaminant mass estimates, and (4) minimizing estimation uncertainty.

These objectives result in a 4-dimensional Pareto surface that was explored using 2-dimensional tradeoffs between selected pairs of objectives. First, objective pairs that are known to conflict were explored through visualizing 7 sampling designs taken from the Cost—SREE, Cost—Mass, and Cost—Uncertainty tradeoffs. Visualization of these designs enabled

stakeholders to discover additional objective conflicts and their effects on the physical monitoring system. The final step in the methodology builds upon the improved stakeholder understanding of design objective interactions to negotiate acceptable bounds for all of the performance criteria used in a monitoring application. These bounds were then used to search the high dimensional Pareto set of optimal sampling strategies for a final compromise solution. The processes of discovery and negotiation demonstrated in this chapter through the use of high order Pareto optimization hold significant potential as tools that can be used in the balanced design of water resources systems.