7 CONCLUDING REMARKS

7.1 Thesis Contributions

This thesis contributes a highly adaptable groundwater monitoring design methodology that can reduce the fiscal burden of long term stewardship for contaminated sites while allowing stakeholders to select, understand, and balance their design objectives. The methodology has immediate potential for application at real world sites where it can be used to save millions of dollars that are currently spent on sampling spatially redundant data. The underlying structure of this thesis is based on decomposing the LTM problem into three steps: (1) select and balance design objectives, (2) choose a plume interpolation method, and (3) quantify design tradeoffs using multiobjective genetic algorithms (GAs). This thesis focused on the specific challenges practitioners will face in implementing the methodology proposed in this research.

The first step of the methodology involves specification, understanding, and negotiated balance of LTM design objectives to assess the "quality" of designs. Properly assessing the quality of designs is the most pivotal and challenging issue facing designers from any discipline because they must capture all of their preferences for a system. This is particularly challenging for LTM design and, in general, environmental design because quality is defined by the conflicting preferences of contaminated sites' owners, multidisciplinary teams of experts, regulatory agents, and the public. This thesis suggests that high order Pareto optimization (i.e., optimizing a system for more than 2 objectives) combined with visualization can be used for selecting and balancing design objectives. Combining higher order Pareto optimization with visualization allows stakeholders to assess the mathematical models used to represent their objectives, discover how their objectives are affecting designs, and negotiate a final design that balances their conflicting design preferences.

In this thesis, plume interpolation is assumed to be required for both evaluating and visualizing design objectives. The second step of the LTM design methodology focuses on identifying which plume interpolation scheme should be used for these purposes. The basic goal of plume interpolation is to attain unbiased maps of contaminant concentrations and local uncertainties. Preferential sampling and the highly skewed nature of groundwater contamination can combine to severely bias interpolation estimates and any subsequent decisions that must be made using these estimates. This thesis directly illustrates that these factors bias the standard practice of using cross-validation to rank the relative performances of interpolation methods. These biased rankings can mislead the practitioner into using a less effective interpolation method and could adversely affect subsequent LTM design decisions. Quantile kriging was selected to evaluate the LTM designs in this research because it was the most robust of the interpolation methods, showing the least bias from both preferential sampling and the variability of contaminant data. These results warrant further studies into the applicability of quantile kriging to data sets from other fields ranging from mining to life sciences, where nonstationary interpolation also plays a vital role.

The final step of the LTM design methodology seeks to quantify high order Pareto surfaces using effectively designed multiobjective GAs. A multiobjective algorithm that is capable of high order Pareto optimization can be used as a tool of discovery for designers, serving as an interface between the human decision process and the design of a physical system. This capability was demonstrated in this thesis when the Nondominated Sorted Genetic Algorithm-II (NSGA-II) identified important system properties that had been previously unknown and would have remained unknown without the algorithm.

Real world applications of high order multiobjective optimization prior to this thesis have been constrained by computational limitations. This thesis overcomes these limitations by developing a design methodology for the NSGA-II that fully exploits the algorithm's efficiency to enable the automatic solution of this new class of problems. The NSGA-II design methodology was used in Chapter 6 to solve the first application of evolutionary multiobjective algorithms to a real-world problem with 4 objectives. The design methodology treats the NSGA-II as a system with discernable properties and utilizes valuable theoretical work from the field of genetic and evolutionary computation to discern these properties. Theoretical relationships allow practitioners to gain direct insights into the algorithm's performance as a function of its input parameters while also reducing the number of parameter combinations that must be considered relative to the trial-and-error methods that have traditionally been used in real world applications. Trial-and-error methods inherently treat evolution-based solution methods as randomized black boxes and obfuscate important trends in performance. In this thesis, recognizing the influence of population sizing and selection pressure on performance facilitated the selection of the NSGA-II and its efficient use as a tool for design and discovery. The NSGA-II design methodology developed in this thesis can be used by designers in any discipline to efficiently discover, understand, and balance tradeoffs among their design objectives.

7.2 Future Research

Although this thesis successfully addresses spatial redundancy analysis using a highly adaptive multiobjective methodology, several issues deserve further consideration:

- How can the methodology be extended to account for contaminant reaction and migration in both space and time?

- How can “soft” data (e.g., numerical modeling data) be combined with “hard” data (e.g., groundwater monitoring samples) to improve LTM design?
- How can multiple contaminants and multiple sampling technologies be integrated into LTM optimization?
- Can both LTM and remediation be optimized simultaneously at contaminated groundwater sites?
- Can the frequency of sampling for costly contaminant constituents be reduced using predictive relationships derived from lower cost measures of groundwater quality?
- When and how should negotiation be integrated into multiobjective design frameworks?

Spatial redundancy analysis assumes periodic re-assessment of the effectiveness of the monitoring network to determine whether locations previously classified as redundant still minimally affect LTM design objectives. Re-assessment reflects the temporal component of the LTM design problem, which results because the evolution of groundwater contaminant plumes occurs in both time and space. Existing redundancy studies have most often considered only time or only space. The few studies that have considered both space and time have generally used only single objective problem formulations. A potential extension of this thesis and previous studies lies in investigating a multiobjective framework that evaluates designs in both time and space. This framework would require the investigation of novel methods for creating spatiotemporal plume maps for single or multiple contaminant species that could range from statistical data-driven approaches to physics-based flow-and-transport simulation.

Flow-and-transport simulation is a tremendous, but costly source of spatiotemporal "soft" data that has significant potential for helping to address several of the outstanding LTM issues listed above. The first question LTM designers face is "Does the size and complexity of my site justify or preclude the costs of physics-based simulation?". The answer to this question for large sites with tens or hundreds of monitoring wells is more straightforward because long-term stewardship will likely cost several million dollars. The maintenance of a simulation model would require a relatively small percentage of the overall cost and has the benefit of allowing stakeholders to jointly optimize remediation and monitoring.

For smaller sites, the cost of maintaining a simulation model may be too large to justify, although the persistence and number of contaminants at a site should be weighed as carefully as a site's physical size and number of monitoring points. It is common for small sites in the United States to monitor several contaminants of concern. The cost of analyzing groundwater for these chemicals can range from tens to thousands of dollars per sample (*Rast* 1997). If the contaminants are persistent, LTM may be required for 30 years or more at a cost of several million dollars. The Department of Energy (DOE) has estimated that it will spend more than $5 billion on long-term site management within the six year period between 2000 and 2006 (*DOE* 2001). Moreover, the DOE estimates that long-term site management will cost $100 million per year for the next 70 years. The argument for using spatiotemporal LTM optimization techniques that integrate flow-and-transport simulation gains credence even for small sites when multiple, persistent contaminants are present.

After determining that fate-and-transport simulation is justified for a site, LTM designers then face the task of deciding how to combine model predictions with groundwater samples (i.e., hard data). Bayesian approaches hold significant potential for addressing this issue by providing a "living model" approach where model predictions are continually conditioned to contaminant observations over the course of a site's lifetime. The primary difficulty of using these methods in combination with multiobjective optimization is their computational complexity. The complexity of these methods grows markedly with the number of contaminants and data sources considered. In these cases, data mining tools such neural networks or decision trees maybe useful for reducing the computational complexity of predictions (see *Michael et al.*, 2002). Moreover, these tools could be used to develop approaches for limiting expensive contaminant observations by replacing them with lower cost groundwater quality measures.

The "living model" approach could greatly improve LTM design, especially if it used in the context of a spatiotemporal multiobjective optimization framework. All multiobjective frameworks for LTM design will result in tradeoff surfaces. Real-world implementations of these frameworks will require the negotiated selection of compromise solutions. Therefore a relevant question is "How should negotiation be integrated into environmental decision making?". Visualization tools that provide a better understanding of how designs evolve in space and time as a consequence of stakeholders' objectives could aid the negotiation process. Research is needed into the relative merits of normative approaches (e.g., multi-attribute utility analysis) versus descriptive approaches (e.g., fuzzy analysis) for aiding negotiation and environmental decision making. Improving our understanding of the intricacies of environmental negotiation could potentially revolutionize current industry policies and practices.