Walking the Tightrope: Long-Term Monitoring Design for Multiple Objectives

Patrick M. Reed[1] and Barbara S. Minsker[1]

Abstract

Tools such as multiobjective genetic algorithms 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 physical system being designed and the human decision process. This paper demonstrates the use of high order Pareto optimization for long-term monitoring (LTM) design, combining quantile kriging and the Nondominated Sorted Genetic Algorithm-II (NSGA-II) to successfully balance four objectives. Optimizing the LTM application with respect to these objectives reduced the decision space of the problem from a total of 500 million designs to the set of 1156 designs identified on the 4 dimensional Pareto surface. Although the 4-dimensional Pareto surface cannot be visualized, this study demonstrates how the set of 1156 designs can inform decision making. First, we analyzed pairs of the objectives that were known to conflict. Visualizing the 7 designs from these tradeoffs provided a better understanding of how these objectives affect sampling designs and aided in discovering additional objective conflicts. Once these conflicts were discovered, they were then used to identify acceptable objective bounds and negotiate a single compromise design.

Introduction

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. This paper demonstrates 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. The paper demonstrates that high order Pareto optimization can provide a means of selecting objectives, discovering objective conflicts, and helping stakeholders in the negotiation process.

The optimization methodology is demonstrated 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

[1] Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, 3230d NCEL, MC. 250, 205 N. Mathews Ave., Urbana, Illinois, USA, 61801, phone (217) 333-9017, email: preed@uiuc.edu, minsker@uiuc.edu

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.

Test Case Data

The test case developed for this study uses data drawn from a 50 million-node flow-and-transport simulation performed by Maxwell et al. (2000). These data were also used by Reed et al. (2001) and Reed et al. (2002a); the concentration data set used in this study corresponds to the medium test case from Reed et al. (2002a).The simulation provided realistic 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 (1) and (2). If the i^{th} monitoring well was selected for sampling then PCE is sampled at all the possible sampling locations along its vertical axis.

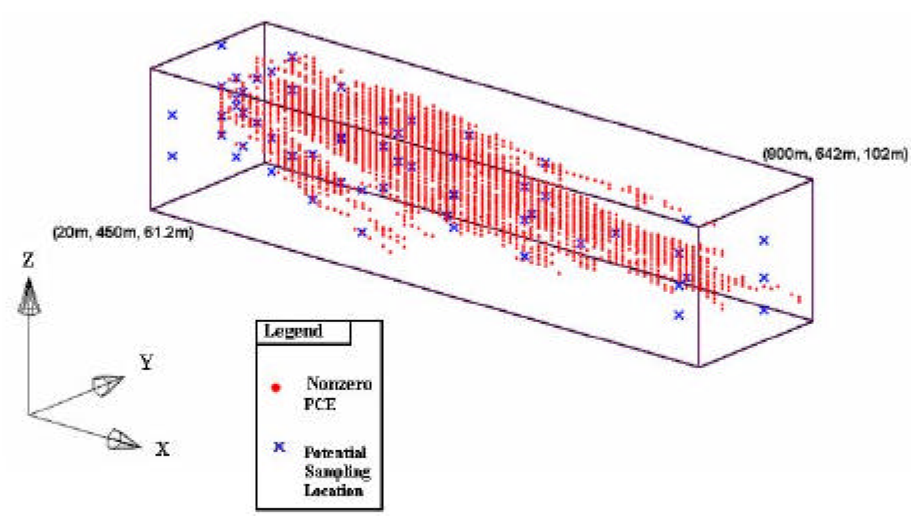

Figure 1. Monitoring network sampling a total of 58 locations.

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. This paper addresses only spatial redundancy analysis, which seeks to identify and remove sampling locations that contribute minimally to understanding the plume's extent in space, time, or both. This study assumes that the spatial sampling plans will be re-evaluated periodically as site conditions change. 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).

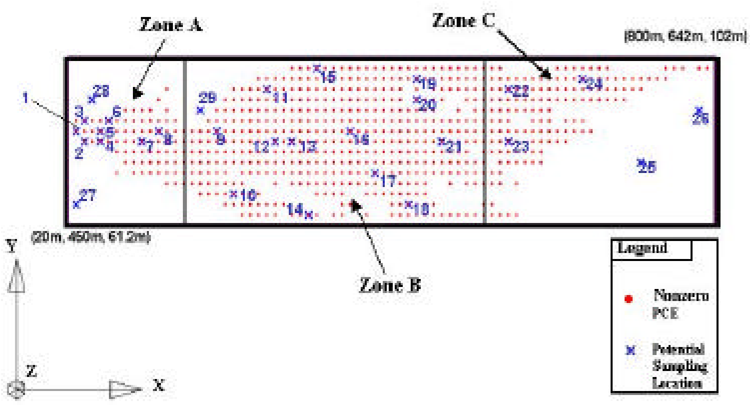

Figure 2. Cross-sectional view of contaminated area showing identification numbers for the 29 monitoring wells.

Methodology

The LTM design methodology proposed in this paper 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 (objectives): (1) cost, (2) squared relative estimation error (SREE), (3) the relative global mass error, and (4) local uncertainty as measured by kriging estimation variances. Cost is a linear function of the number of PCE samples that are used in a given monitoring design. SREE measures how the interpolated picture of the plume using data only from wells included in the k^{th} sampling plan compares to the result attained using data from all available sampling locations. Likewise, the global mass objective error in the total mass of PCE in the subsurface. Lastly, local uncertainty is estimated using the sum of the estimation standard deviations (i.e., the square root of estimation variances) from kriging (for more details see Reed et al. 2002a and Reed et al. 2002b).

Plume Interpolation using Quantile Kriging. Quantile kriging was selected for plume interpolation in this study based on the findings of Reed et al. (2002a), who present 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. For more details on quantile kriging see Journel & Deutsch (1997), Juang et al. (2001), and Reed et al. (2002a).

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.

Reed (2002) 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 Reed (2002) for more details.

Results

In the LTM application a single constituent is being monitored 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 the 4 objectives 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.

First, we analyzed pairs of the objectives that were 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. A total of 7 sampling designs were then visualized from across the following tradeoffs shown in Figure (3): (1) Cost—SREE, (2) Cost—Mass Error, and (3) Cost—Uncertainty. Visualizing the 7 designs from these tradeoffs provided a better understanding of how these objectives affect sampling designs and aided in discovering and understanding conflicts among the following additional pairs of objectives: (1) SREE—Mass Error, (2) SREE—Uncertainty, and (3) Uncertainty—Mass Error.

(a)

(b)

(c)

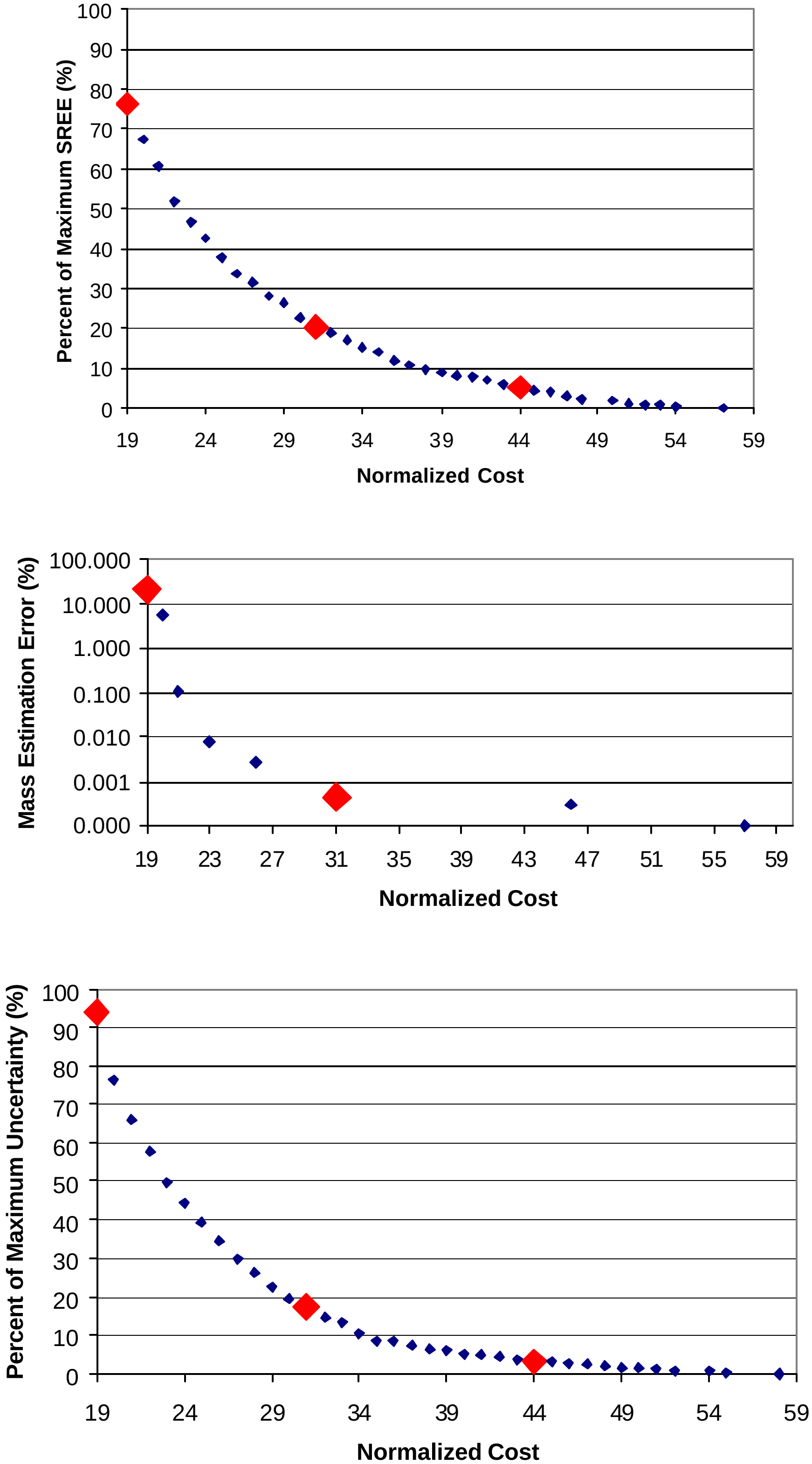

Figure 3. (a) Cost—SREE tradeoff (b) Cost—Mass tradeoff (c) Cost—Uncertainty tradeoff**.**

For example, the Cost—Mass Error tradeoff shown in Figure (3b) shows that only 21 sampling locations were required to attain an accurate mass estimate. This result could motivate the decision maker to eliminate the mass objective from consideration, which would be a mistake because Figure (3b) only considers Cost—Mass Error interactions. The two solutions shown in Figures (4a) and (4b) are from the Cost—Uncertainty tradeoff presented in Figure (3c). The 44 sample solution shown in Figure (4b) uses 13 more sampling locations to decrease uncertainty relative to the solution shown in Figure (4a), but note the surprising increase in mass estimation error relative to the 31 sample solution in Figure (4a).

A very interesting multi-well interaction between well numbers 1, 27, and 28 [see Figure (2)] 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 sampling wells 27 and 28 in the absence of well 1 cause the mass in Zone A to be severely underestimated. This result identifies an unexpected Uncertainty—Mass Error conflict and demonstrates how visualizing the designs highlighted in Figure (3) helped to inform the decision process.

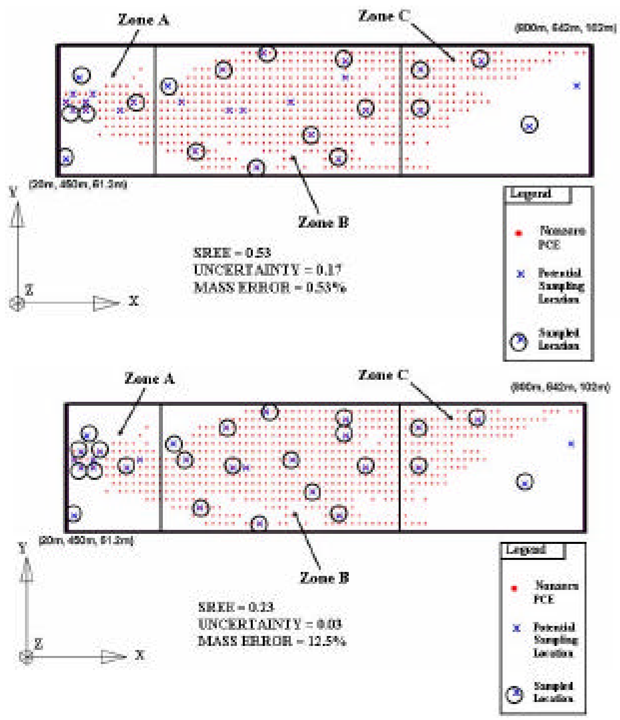

Figure 4. Sampling locations corresponding to solutions from the Cost—Uncertainty tradeoff (a) 31 sample solution (c) 44 sample solution.

Next, the insights gained from examining tradeoffs can be used to negotiate a final compromise sampling scheme by selecting acceptable bounds on each objective. For example, a conservative cost level of 44 sampling locations could be selected, which would reduce costs by nearly 25 percent while minimally increasing the remaining objectives. 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 for the remaining objectives. Upper bound values for SREE, mass error, and uncertainty were set equal to 10%, 1%, and 10% of the maximum values in the nondominated set, respectively. The objective bounds were set to exploit potential decreases in each objective's value that would minimally increase the 3 remaining objectives' values. For example, the Mass—Uncertainty tradeoff presented in Figure (5) shows that uncertainty can be significantly reduced with less than a 1 percent increase in mass estimation error.

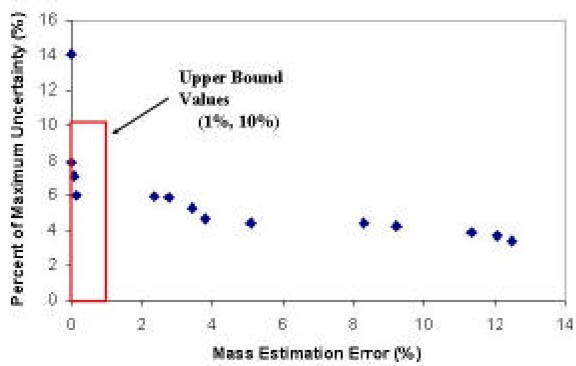

Figure 5. The Uncertainty—Mass Error tradeoff for 44 sample solutions.

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). 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) 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.

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) 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.

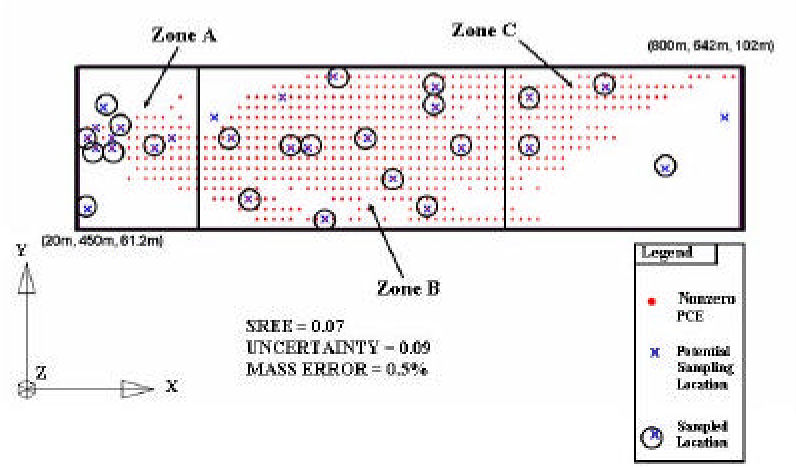

Figure 5. The 44 sample compromise solution.

Conclusions

The optimization methodology presented in this paper 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. Demonstration of the methodology on an LTM design problem shows how multiobjective optimization combined with visualization can aid practitioners in selecting, understanding, and balancing performance objectives when seeking a single compromise solution. The monitoring application results 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 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 study 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.

Acknowledgments

This work was supported by the United States Environmental Protection Agency's Science to Achieve Results (STAR) graduate fellowship program. The authors are grateful to Dr. Albert J. Valocchi for aiding in the development of the monitoring methodology. Dr. David E. Goldberg is gratefully acknowledged for his guidance in the selection and implementation of the NSGA-II. Dr. Tim Ellsworth is acknowledged for his guidance in the implementation of quantile kriging. Additionally, the authors are grateful to Dr. Reed Maxwell for providing the simulation data used in the monitoring test case.

References

ASCE Task Committee on Geostatistical Techniques (1990). "Review of Geostatistics in Geohydrology II: Applications." *Journal of Hydraulic Engineering*, 116(5), 633-658.

Aziz, J. J., Newell, C. J., Rifai, H. S., Ling, M., & Gonzales, J. R. (2000) *Monitoring and Remediation Optimization System (MAROS)*: *Software User's Guide*, Version 1, United States Air Force Center for Environmental Excellence, Brooks AFB: San Antonio, TX.

Buras, Nathan (2001). "Water Resources—Unresolved Issues." *Journal of Water Resources Planning and Management*, 127(6), 353.

Cameron, K. & Hunter, P. (2000). "Optimization of LTM Networks: Statistical Approaches to Spatial and Temporal Redundancy." Proceedings from the American Institute of Chemical Engineers, 2000 Spring National Meeting, Remedial Process Optimization Topical Conference, Atlanta, GA.

Deb, K., S. Agrawal, A. Pratap, and T. Meyarivan (2000). "A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-objective Optimization: NSGA-II." Kanpur Genetic Algorithm Laboratory (KanGAL) Report 200001, Indian Institute of Technology, Kanpur, India.

Deb, K., L. Thiele, M. Laumanns, and E. Zitzler (2001). "Scalable Test Problems for Evolutionary Multi-Objective Optimization." Computer Engineering and Networks Laboratory Report (TIK-112), Department of Electrical Engineering, Swiss Federal Institute of Technology, Zurich, Switzerland.

Journel, A. G. and C. V. Deutsch (1997). "Rank Order Geostatistics: A proposal for a unique coding and common processing of diverse data." In E.Y. Baafi and N. A. Schofield (eds.) Geostatistics Wollongong '96, Volume 1, Proceedings of the 5th International Geostatistics Congress, Wollongong, Australia, Kluwer Academy Publ., Dordrecht, The Netherlands.

Johnson, V. M., R. C. Tuckfield, M. N. Ridley, and R. A. Anderson (1996). "Reducing the sampling frequency of groundwater monitoring wells." *Environmental Science &* Technology, 30(1), 355-358.

Juang, K. W., Dar-Yuan Lee, and Timothy R. Ellsworth, (2001). "Delineation of Heavy Metal Contaminated Soils with Highly Skewed Data using Quantile Kriging." *Journal of Environmental Quality* (in press).

Loaiciga, Hugo, A., R. J. Charbeneau, L. G. Everett, G. E. Fogg, B. F. Hobbs, S. Rouhani (1992) "Review of Ground-Water Quality Monitoring Network Design." *Journal of Hydraulic Engineering*, 118(1), 11-37.

Lobo, F., (2000). *The Parameter-less Genetic Algorithm: Rational and Automated Parameter Selection for Simplified Genetic Algorithm Operation*, Doctoral Dissertation, Universidad Nova de Lisboa, Faculdade de Ciências e Tecnologia.

Maxwell, Reed M., F. S. Carle, and F. B. Tompson (2000). *Contamination, Risk, and Heterogeneity: On the Effectiveness of Aquifer Remediation*, Lawrence Livermore National Laboratory Report, UCRL-JC-139664, Livermore, CA.

Reed, P., B. Minsker, and A. J. Valocchi. (2000a). "Cost effective long-term groundwater monitoring design using a genetic algorithm and global mass interpolation." *Water Resources Research*, 36(12): 3731-3741.

Reed, P., B. Minsker, and D. E. Goldberg, (2000b). "Designing a competent simple genetic algorithm for search and optimization." *Water Resources Research*, 36(12): 3757-3761.

Reed, P., B. Minsker, and D. E. Goldberg, (2001). "A multiobjective approach to cost effective long-term groundwater monitoring using an Elitist Nondominated Sorted Genetic Algorithm with historical data." *Journal of Hydroinformatics*, 3(2), 71-90.

Reed, P., T. R. Ellsworth, and B. Minsker (2002a). “Spatial Interpolation Methods for Nonstationary Plume Data.” *Ground Water*, In Review.

Reed, P. (2002). *Striking the Balance: Long-term Groundwater Monitoring Design for Multiple Conflicting Objectives*, Doctoral Dissertation, University of Illinois, Urbana, IL.

Rouhani, S. (1985). “Variance reduction analysis.” *Water Resources Research*, 21(6), 837-846.

Rouhani, S., and T. J. Hall (1988). “Geostatistical schemes for groundwater sampling.” *Journal of Hydrology*, 103, 85-102.

Zitzler, E., M. Laumanns, and L. Thiele (2001). “SPEA2: Improving the Strength Pareto Evolutionary Algorithm.” Computer Engineering and Networks Laboratory Report (TIK-103), Department of Electrical Engineering, Swiss Federal Institute of Technology, Zurich, Switzerland.