

Assessing the Value of Environmental Observations in a Changing World: Nonstationarity, Complexity, and Hierarchical Dependencies

Patrick M. Reed^a, Ruchit Shah^a, and Joshua B. Kollat^a
The Pennsylvania State University (preed@enr.psu.edu)

Abstract: This paper explores three related propositions for designing environmental observation systems: (1) Nonstationarity in environmental data series has the consequent impact of making observation network design itself a nonstationary, stochastic planning problem where the value of alternative observation strategies should be evaluated based on planners' evolving conception of Pareto efficiency given new knowledge, technologies, and policies over long time-scales. (2) Real-world budgetary constraints within observation network design problems yield resource allocation conflicts across space, time, and competing foci that are equivalent in form to the multiobjective d-dimensional knapsack problem (MO-dKP). Consequently, the Pareto efficiency of observation networks can only be determined approximately for non-trivial problem instances. (3) Multiobjective hierarchical Bayesian optimization provides a very promising tool for identifying observation alternatives that are approximately Pareto efficient while simultaneously providing insights into the emergent dependencies of our decisions (both science and management oriented) on critical observations.

Keywords: observation networks, many-objective analysis, Bayesian networks

1. INTRODUCTION

Our detection, prediction and management of critical environmental gradients is fundamentally dependent on our ability to design and manage observation networks. As noted by Reed et al. [2006], environmental change necessitates a shift from myopic, non-adaptive long-term observation strategies towards adaptive design frameworks that link our observations and predictions of evolving human—natural systems. Key to this challenge is properly posing and analyzing the question: *what environmental observations are necessary to detect, predict, and manage the risks posed by environmental change?* Although a more holistic assessment is justified, at present our national, regional, and local observation strategies are largely ad-hoc, non-adaptive, and generally disconnected from evolving water resources policy and management needs, a condition that has long been recognized [Davis et al., 1979; Langbein, 1979; Moss, 1979; United States Geological Survey, 1999].

Thirty years ago Marshall Moss [1979] eloquently acknowledged these challenges and framed the need for future observation network design strategies to use a "...more integrated measure of information...[that] results from a complex interaction of both the hydrologic knowledge and the procedures that are used to incorporate the knowledge into...decisions" (p. 1673). Moss's recommendation represents a major departure from the more commonly employed statistical information measures [e.g., Shannon, 1948; Kiefer, 1959] by seeking to understand the value of observables for advancing knowledge while simultaneously characterizing their value to the procedures used to make decisions. This discussion paper draws on our recent research results to highlight three related challenges that must be considered when judging the value of observation systems as well as their gaps through their space, time, and management dimensions:

1. Nonstationarity in hydrologic systems has the consequent impact of making observation network design itself a nonstationary, stochastic planning problem where the value of alternative observation strategies should be evaluated based on planners' evolving conception of Pareto efficiency given new knowledge, technologies, and policies over long time-scales.
2. Real-world budgetary constraints within observation network design problems yield resource allocation conflicts across space, time, and competing foci that are at least as challenging as the multiobjective d-dimensional knapsack problem (MO-dKP). Consequently, determining the Pareto efficiency of observation networks has a NP-Complete complexity (Nondeterministic Polynomial time-complete), which means that globally optimal tradeoffs cannot be attained with modern computers for non-trivial problem instances.
3. Multiobjective hierarchical Bayesian optimization represents a very promising tool for identifying observation alternatives that are approximately Pareto efficient while simultaneously providing insights into the emergent dependencies of our decisions (both science and management oriented) on critical observations.

The remainder of this paper explores each of these challenges in more detail.

2. TRADEOFFS & NONSTATIONARITY IN AN INTEGRATED MEASURE

It is very difficult to estimate the value of observations in terms of their impacts for improving our understanding of the evolving risks and environmental services of water resources systems. We propose that Pareto efficiency provides a mechanism for discovering and understanding information worth tradeoffs across the broad range of incommensurate objectives that could be of interest for observation network design problems (minimize costs, minimize risk, maximize coverage, minimize uncertainty, etc). The performance of any potential monitoring alternative \mathbf{X}_k must be evaluated in a manner that considers performance across the total component objectives that relate to investments, prediction goals, and management needs. Feasible solutions to the problem are evaluated in terms of their nondomination. Assuming minimization of all objectives: an observation alternative \mathbf{X}_k^1 dominates \mathbf{X}_k^2 , $\mathbf{X}_k^1 \succ \mathbf{X}_k^2$, if its objectives' values are less than or equal to those of \mathbf{X}_k^2 and there exists at least one objective where \mathbf{X}_k^1 attains a lower objective value. This mathematical partitioning rule then serves to identify the feasible space of sampling alternatives that do not have their performance exceeded in all objectives [Pareto, 1896]. Figure 1 provides a two-objective illustration of the concepts of nondomination and Pareto efficient fronts. In the figure, assuming minimization of both cost and error, the goal is to attain the minimum level error for each level of cost. The shaded boxes designate the objective space dominated by solutions 1, 2, and 3 in the figure. The full set of nondominated solutions as plotted represent the Pareto efficient frontier (i.e., the optimal tradeoff between cost and error).

The Cost—Error Pareto front is classified as an a posteriori decision analysis tool, which means that decision makers are provided with an explicit representation (see Figure 1) of their design tradeoffs when seeking compromises between conflicting objectives. Although the water resources literature has largely focused on two-objective formulations, there is a growing body of literature that is advancing a more generalized “many-objective” version of planning and design using problem formulations with 3 or more objectives [Balling, 1999; Reed and Minsker, 2004; Bekele and Nicklow, 2005; Fleming et al., 2005; di Pierro, 2006; Kollat and Reed, 2007; Kollat et al., 2008; Zhang et al., 2008; Kasprzyk et al., In-Press; Hadka and Reed, In-Review]. Kollat et al. [2007] use an observation network example to demonstrate how many-objective search and interactive visualization provide a new a posteriori decision aid for discovering tradeoffs, decision interactions, and design consequences across a range of objectives. These issues link strongly to Moss's [1979] proposal for an integrated information measure that links advances in our scientific knowledge to “...the procedures that are used to incorporate the knowledge...into

decision”. Assessing Pareto efficiency serves as a unifying decision analysis where decision makers can discover and visually explore tradeoffs in how investments in observations impact our knowledge and management objectives.

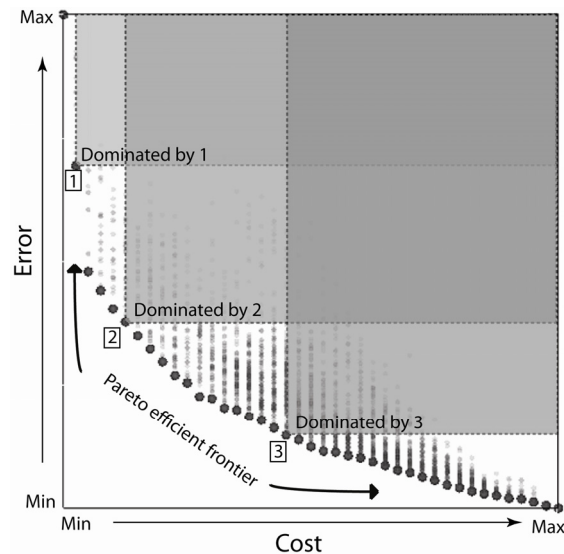


Figure 1. Illustration of the concepts of domination and Pareto efficiency assuming minimization of Cost and Error. Solutions 1, 2, and 3 dominate all of the solutions within their respective grey shaded regions in all objectives. The Pareto efficient frontier is the non-dominated solutions that attain the minimum level of Error at each Cost.

Using the concept of Pareto efficiency as an integrated information measure will require innovations in how we define many-objective network design problem formulations, advancements in promising new solution algorithms [e.g., see [Coello Coello et al., 2007]], and new decision analysis technologies incorporating interactive visual analytics [Russell et al., 1993; Keim et al., 2006; Kollat and Reed, 2007; Sanfilippo et al., 2009; Kasprzyk et al., In-Press]. Objectives and decisions are contextually driven by the needs within each management period and would be expected to evolve with new needs and problem insights. Broadly, this issue highlights that the mathematical spaces (or topologies) that define observation systems’ tradeoffs are in fact nonstationary and often highly uncertain. They are nonstationary in the sense that as designers make new discoveries, these discoveries feedback to new hypotheses which then motivate human-guided structural changes in mathematical formulations [e.g., see the de novo planning concepts of Zeleny, 2005]. Given the increasingly severe uncertainties, dependencies and decision tradeoffs for complex environmental systems, the many-objective Pareto efficiency information measure explicitly elucidates the consequences, compromises, and hypotheses that emerge with new information and knowledge.

3. COMPLEXITY OF MANY-OBJECTIVE INFORMATION MEASURE

Beyond the number of objective conflicts, sampling decisions also strongly influence the computational complexity of the environmental observation network design. The n -dimensional binary decisions $\mathbf{X}_k \in \{0, 1\}^n$ represents a lower bound in the complexity of the problem where yes/no decisions are made across space, time, and different environmental states (flow, water quality species, etc.). A linear increase in sampling decisions yields a 2^n exponential growth rate of the number of design alternatives. Other defensible formulations that include real-valued and/or integer decisions could have far more severe growth rates (e.g., factorial or potentially infinite). Our goal in analyzing the lower bound complexity of the problem is to demonstrate the strong computational challenge that environmental observation network design poses. When solving many-

objective formulations of network design problems, their overall problem difficulty will be governed by maximally difficult sub-problem(s).

Consequently, we can build on known results from computational complexity theory to show that the environmental observation network design problem is NP-complete (Nondeterministic Polynomial time-complete). In brief, an NP-complete problem [for a more formal discussion see Cook, 1971; Garey and Johnson, 1979] represents a severely difficult problem that cannot be solved exactly using modern computers (i.e., deterministic Turing machines) for instances that cannot be enumerated. Therefore non-trivial instances can only be solved approximately and global optimality is not attainable for any algorithm. The NP-complete computational complexity of network design can be surmised by considering a simple accounting of cost objectives $\mathbf{j}_{cost}(\mathbf{A}_{k-1}^f, \mathbf{X}_k)$ used for the constrained allocation of sampling investments between D system states. Equation (1) provides a highly simplified D-objective constrained cost formulation which would be a subset problem of a full formulation that could include prediction and/or management objectives. Equation (1) is a simple representation of investment tradeoffs. Moreover, equation (1) is identical to the multiobjective D-dimensional knapsack allocation problem [Martello and Toth, 1990; Shah and Reed, In-Review].

$$\begin{aligned} \text{Min } \mathbf{j}_{cost} &= \left[j_{cost}^1(\mathbf{A}_{k-1}^f, \mathbf{X}_k), j_{cost}^2(\mathbf{A}_{k-1}^f, \mathbf{X}_k), \dots, j_{cost}^D(\mathbf{A}_{k-1}^f, \mathbf{X}_k) \right] \\ \text{where } j_{cost}^i(\mathbf{A}_{k-1}^f, \mathbf{X}_k) &= \sum_{j=1}^n p_{ij} x_j, \quad \forall i \in [1, \dots, D] \end{aligned} \quad (1)$$

Subject to:

$$\begin{aligned} \left[\sum_{j=1}^n p_{ij} x_j \right]_k &\leq c_{ik}, \quad \forall i \in [1, \dots, D] \\ \mathbf{X}_k &\in \{0, 1\}^n \end{aligned}$$

Although more complex cost equations that incorporate the nonlinearities and complexities that could be associated with a more economic-oriented formulation that accounts for the time value of investment would be defensible, equation (1) provides arguably the simplest meaningful accounting for cost as a simple linear summation of discrete costs p_{ij} for the j^{th} sample of the i^{th} state. Taken as a whole, the objectives and constraints of equation (1) represents the subset sum instance of the knapsack where capacity constraints' weights equal items' respective profit coefficients, see [Martello and Toth, 1990; Pisinger, 2005; Shah and Reed, In-Review].

Consequently, given that the knapsack problem is a classic NP-complete problem, equation (1) implies that environmental observation design is an NP-complete problem class. Returning to our proposal of using many-objective Pareto efficiency as an integrated information measure, equation (1) implies that robust computational tools are needed to attain high quality approximations to the optimal tradeoffs. Moreover, alternative instances of the knapsack problem can be vastly more difficult than others when seeking high quality approximate solutions. So the immediate concern for environmental observation network design is answering the question, how hard is our instance of the knapsack?

The historical theoretical work for the knapsack provides insights into the difficulty of the observation networks problem class. Prior studies [Martello and Toth, 1990; Pisinger, 2005] have clearly shown that a high degree of correlation or interdependence between the knapsack problem's binary decisions and/or constraints often dramatically increases the difficulty of finding high quality approximate solutions. These findings represent a severe concern for environmental network design because observation decisions across space-and-time are fundamentally linked and interdependent due to hydrologic systems' socio-physical organization. Mathematically the concept of hierarchy provides a useful means of capturing how a particular decision to observe at the current location and time influences the impacts of observations at other times and locations. In simple terms, hierarchy may be

viewed as a series of probabilistic if-then-else observation rules across space-and-time. It remains an important challenge to discover and exploit these dependencies observation design frameworks.

4. DISCOVERING HIEARCHICAL DEPENDENCIES

The discovery of hierarchical if-then-else probabilistic rules for complex, adaptive systems has been a focus of the artificial intelligence field since its inception [Simon, 1968]. For environmental observation network problems, we add to this ambition the challenge posed by the NP-complete knapsack problem structure detailed in equation (1). While the socio-physical organization of environmental systems would be expected to engender interdependencies in observation decisions, computationally this knowledge is not present in traditional multiobjective optimization tools [for a detailed review see Coello Coello et al., 2007]. Recently, Pelikan and Goldberg [2003] introduced a new form of evolutionary optimization tool termed the Hierarchical Bayesian Optimization Algorithm (hBOA) to provide the capability to learn and exploit Bayesian network models of decision interdependencies while solving problems. The hBOA innovations are strongly relevant to the environmental observation network design problem's limiting challenges. Pelikan [2002] demonstrates that the hBOA can attain sub-quadratic computational scaling for severely challenging hierarchically structured problems. In short, this means that the hBOA has reduced computational demands when solving increasingly larger problems with hierarchical dependencies. Moreover, Pelikan and Goldberg [2003] show that more traditional solution tools can have exponentially scaled computational complexities for hierarchically structured problems (i.e., they rapidly fail to attain high quality results unless they resort to enumeration).

At its original inception, the hBOA was a single objective probabilistic model building evolutionary algorithm [Pelikan, 2002; Pelikan et al., 2002]. Expanding the hBOA to address many-objective instances of equation (1), Kollat et al. [2008] introduced the Epsilon Dominance Hierarchical Bayesian Optimization Algorithm (ϵ -hBOA). The ϵ -hBOA represents a new type of multiobjective evolutionary algorithm where during evolution, the ϵ -hBOA selects high performing solutions and builds Bayesian network models of the underlying probabilistic dependencies of their decisions. After learning these dependencies, ϵ -hBOA uses them within its Bayesian network models to generate probabilistic hypotheses on what decision combinations would yield improved candidate solutions.

An important contribution of the ϵ -hBOA is that while solving many-objective problems, the algorithm is explicitly building a joint probabilistic density function (PDF) model of what makes observation decisions likely to be non-dominated with respect to many-objectives, simultaneously. The joint PDF provides the potential for discovering and visualizing the hierarchical dependency structure of environmental observation problems. Kollat et al. [2008] used a many-objective groundwater monitoring application to show that the ϵ -hBOA can dramatically enhance our approximate evaluation of Pareto efficiency for increasingly larger networks. Shah and Reed [In-Review] further showed that the ϵ -hBOA provides a robust approximation technique for many-objective knapsack instances with strong dependency structures.

Figure 2A provides a network visualization of some of the strongest hierarchical rules proposed by the ϵ -hBOA that are expected to be satisfied by greater than 98% of all of the non-dominated solutions for the test case. In reality, the algorithm provides a joint PDF model of rules for a full range of probabilities. In the ϵ -hBOA network illustration in Figure 2A, the numbered circles represent potential decisions for sampling one of the 25 available wells. Green designates wells that are sampled and red represents locations that are not sampled. In Figure 2A, the interior most circles represent the dependent variables in the rules subject to the decisions made in the outer circles. The order of hierarchy for each rule is defined by the number of independent decisions that influence the interior dependent sampling rules. For example, well 22 has a zeroth order hierarchy rule which indicates that it should be sampled by greater than 98-percent of non-dominated solutions independent of all other wells. Alternatively, the ϵ -hBOA proposes rules of up to the 7th order hierarchy

for sampling well 2, which is striking given the simplicity of the test case. Note that well 2 has multiple rule instances which represent alternative independent decisions that would motivate that the location be sampled.

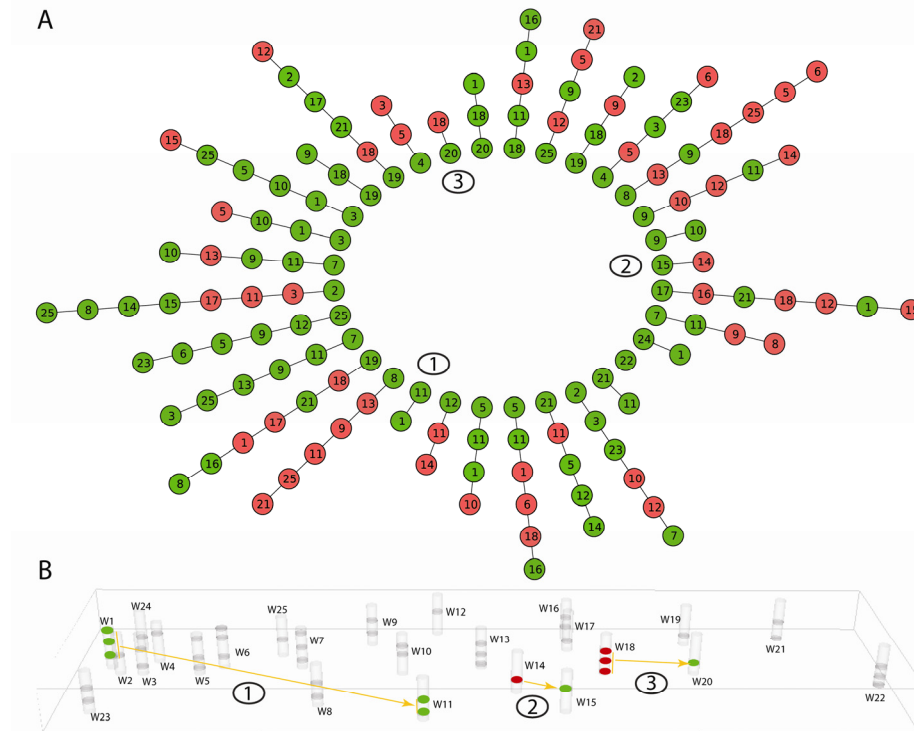


Figure 2. (A) The network graphic depicts examples of the hierarchical Bayesian rules proposed by the ϵ -hBOA when searching for Pareto efficient sampling strategies. The numbering and coloring corresponds to well identifications and sampling decisions (green—sampled, red—not sampled). The interior most green circles are the dependent sampling decisions. Excluding the interior dependent wells, the number of exterior wells in each rule defines its hierarchical order. All rules shown are proposed by the ϵ -hBOA to occur in greater than 98-percent of the solutions that compose the test case’s Pareto efficient front. (B) A spatial representation of the multi-point sampling wells in the test case’s domain with illustrations of three example rules.

Figure 2B provides a spatial visualization of three 1st order hierarchy rules highlighted in Figure 2A. The ϵ -hBOA’s proposed rules are sources of potential hypotheses on system behavior when you place them within the natural space-time contexts of environmental systems. Figure 2B provides contextual meaning for some of the rules discussed in Figure 2A. For example, wells 1 and 11 jointly sample the defining boundaries of the plume. Similarly, if well 18 is not sampled along the longitudinal axis of the plume, then the ϵ -hBOA’s rule proposes that well 20 should be, which makes intuitive since it is the next closest well near the mid-line of the PCE plume. Likewise, well 15 is a suggested substitute for well 14. Broadly, the network graphics in Figure 2 provide a classification of how sensitive non-dominated sampling strategies are to a range of interdependent sampling decisions. The rules provided by the ϵ -hBOA provide decision makers with a means of discerning which sampling decisions have broad impacts over the full plume (e.g., well 1 in the PCE source area) versus those that have more localized effects (well 22 on the edge of the domain). There are two important issues to note in Figure 2. First, although some of the simpler rules may seem intellectually trivial, they represent problem knowledge that traditional solution tools are incapable of capturing, thus enabling the ϵ -hBOA to deal with severely interdependent problems (a form of severe nonlinearity termed epistasis). Secondly, the more complex high order hierarchical rules in Figure 2A pose interesting hypotheses on the relationships and controls impacting how alternative sampling strategies

can attain near optimal tradeoffs between the test case's four objectives. Figures 2 supports our contentions that the general environmental observation network design problem class represents difficult knapsack problems with hierarchical dependency structures.

5. CONCLUSION

Environmental change motivates the consideration of the data intensive metrics for the risks, resilience, and adaptability of water resources systems. It is important to recognize that ad hoc and myopic observation management policies for environmental observation systems may pose substantive risks to our ability to manage environmental change over long-time scales. They lack an integrated view of long-term economic costs, impacts on scientific predictive skill, and long-term risk management consequences. Understanding of the value of information within the evolving network-of-networks that characterize national monitoring efforts requires a more direct assessment of their evolving tradeoffs and dependencies, scalable improvements in observation network design methods, and holistic assessments of water resources risks.

ACKNOWLEDGMENTS

The authors of this paper were partially supported by the United States National Science Foundation (NSF) Career grant CBET-0640443. Any opinions, findings, and conclusions or recommendations in this paper are those of the authors and do not reflect the views of the NSF.

REFERENCES

- Balling, R. (1999), Design by Shopping: A New Paradigm, paper presented at Proceedings of the Third World Congress of Structural and Multidisciplinary Optimization, Buffalo, NY.
- Bekele, E. G., and J. W. Nicklow (2005), Multiobjective management of ecosystem services by integrative watershed modeling and evolutionary algorithms, *Water Resources Research*, 41(W10406), doi:10.1029/2005WR004090.
- Coello Coello, C., G. B. Lamont, and D. A. Van Veldhuizen (2007), *Evolutionary Algorithms for Solving Multi-Objective Problems*, 2 ed., Springer, New York, NY.
- Cook, S. A. (1971), The complexity of theorem proving procedures, paper presented at Proceedings of the Third Annual ACM Symposium on the Theory of Computing, ACM, New York.
- Davis, D. R., L. Duckstein, and R. Krzysztofowicz (1979), Worth of Hydrologic Data for Nonoptimal Decision-Making, *Water Resources Research*, 15(6), 1733-1742.
- di Pierro, F. (2006), Many-objective evolutionary algorithms and applications to water resources engineering, University of Exeter, Exeter, UK.
- Fleming, P. J., R. C. Purshouse, and R. J. Lygoe (2005), Many-Objective Optimization: An Engineering Design Perspective, paper presented at Evolutionary Multi-Criterion Optimization, Springer Lecture Notes in Computer Science, Guanajuato, Mexico.
- Garey, M., and D. Johnson (1979), *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman, New York.
- Hadka, D., and P. M. Reed (In-Review), Diagnostic Assessment of Search Controls and Failure Modes in Many-Objective Evolutionary Optimization, *IEEE Transactions on Evolutionary Computation*.
- Kasprzyk, J. R., P. M. Reed, B. Kirsch, and G. Characklis (In-Press), Managing population and drought risks using many-objective water portfolio planning under uncertainty, *Water Resources Research*, doi:10.1029/2009WR008121.
- Keim, D., F. Mansmann, J. Schmeidewind, and H. Ziegler (2006), Challenges in Visual Data Analysis, paper presented at Proceedings of Information Visualization, IEEE Computer Society, London, UK.
- Kiefer, J. (1959), Optimum experimental designs (with discussion), *Journal of the Royal Statistical Society B*, 21, 272-319.

- Kollat, J. B., and P. Reed (2007), A Framework for Visually Interactive Decision-making and Design using Evolutionary Multiobjective Optimization (VIDEO), *Environmental Modelling & Software*, 22(12), 1691-1704.
- Kollat, J. B., P. M. Reed, and J. R. Kasprzyk (2008), A New Epsilon-Dominance Hierarchical Bayesian Optimization Algorithm for Large Multi-Objective Monitoring Network Design Problems, *Advances in Water Resources*, 31(5), 828-845.
- Langbein, W. B. (1979), Overview of Conference on Hydrologic Data Networks, *Water Resources Research*, 15(6), 1867-1871.
- Martello, S., and P. Toth (1990), *Knapsack Problems, Algorithms and Computer Implementations*, edited, John Wiley & Sons, LTD., New York, NY.
- Moss, M. E. (1979), Some Basic Considerations in the Design of Hydrologic Data Networks, *Water Resources Research*, 15(6), 1673-1676.
- Pareto, V. (1896), *Cours D'Economie Politique*, Rouge, Lausanne.
- Pelikan, M. (2002), *Bayesian Optimization Algorithm: From Single Level to Hierarchy*, University of Illinois, Urbana-Champaign.
- Pelikan, M., D. E. Goldberg, and F. Lobo (2002), A Survey of Optimization by Building and Using Probabilistic Models, University of Illinois at Urbana-Champaign, Urbana.
- Pelikan, M., and D. E. Goldberg (2003), A Hierarchy Machine: Learning to Optimize from Nature and Humans, *Complexity*, 8(5), 36-45.
- Pisinger, D. (2005), Where are the hard knapsack problems?, *Computers & Operations Research*, 32(9), 2271-2284.
- Reed, P., and B. S. Minsker (2004), Striking the Balance: Long-Term Groundwater Monitoring Design for Conflicting Objectives, *Journal of Water Resources Planning and Management*, 130(2), 140-149.
- Reed, P., R. Brooks, K. Davis, D. R. DeWalle, K. A. Dressler, C. J. Duffy, H. S. Lin, D. Miller, R. Najjar, K. M. Salvage, T. Wagener, and B. Yarnal (2006), Bridging River Basin Scales and Processes to Assess Human-Climate Impacts and the Terrestrial Hydrologic System, *Water Resources Research*, 42(W07418), doi:10.1029/2005WR004153).
- Russell, D., M. Stefik, P. Pirolli, and S. Card (1993), Sources of Structure in Sensemaking, paper presented at Proceedings of the SIGCHI Conference on Human factors in Computing Systems, Amsterdam, The Netherlands.
- Sanfilippo, A., A. Cowell, E. Malone, R. Riensche, J. Thomas, S. Unwin, P. Whitney, and P. Wong (2009), Technosocial Predictive Analytics in Support of Naturalistic Decision Making, in *The 9th bi-annual International Conference on Naturalistic Decision Making Jointly*, edited, Jointly organized with the British Computer Society INTERACTION Specialist Group, Covent Garden, London, UK.
- Shah, R., and P. M. Reed (In-Review), Comparative Analysis of Multiobjective Evolutionary Algorithms for Random and Correlated Instances of the Multiobjective d-Dimensional Knapsack Problems, *European Journal of Operational Research*.
- Shannon, C. E. (1948), A mathematical theory of communication, *Bell System Tech. Journal*, 7, 379-423.
- Simon, H. (1968), *The Sciences of the Artificial*, MIT Press, Cambridge.
- United States Geological Survey (1999), *Streamflow Information for the Nation*, edited by D. o. Interior.
- Zeleny, M. (2005), *The Evolution of Optimality: De Novo Programming*, paper presented at Evolutionary Multi Criterion Optimization: Third International Conference (EMO 2005), Lecture Notes in Computer Science Springer-Verlag, Guanajuato, Mexico.
- Zhang, Z., T. Wagener, P. Reed, and R. Bhushan (2008), Reducing uncertainty in predictions in ungauged basins by combining hydrologic indices regionalization and multiobjective optimization, *Water Resources Research*, 44(W00B04), doi:10.1029/2008WR006833.