The Pennsylvania State University
The Graduate School

MANY OBJECTIVE WATER RESOURCES PLANNING AND MANAGEMENT GIVEN DEEP UNCERTAINTIES, POPULATION PRESSURES, AND ENVIRONMENTAL CHANGE

A Dissertation in
Civil Engineering
by
Joseph Robert Kasprzyk

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2013

The thesis of Joseph Robert Kasprzyk was reviewed and approved* by the following:

Patrick M. Reed
Associate Professor of Civil Engineering
Thesis Advisor, Chair of Committee

Michael N. Gooseff
Associate Professor of Civil Engineering

Gregory W. Characklis
Professor of Environmental Sciences and Engineering, University of North Carolina Chapel Hill

Peggy A. Johnson
Professor of Civil Engineering, Department Head

Seth Blumsack
Assistant Professor of Energy and Mineral Engineering

*Signatures are on file in the Graduate School.

Abstract

Climate change and population growth require adaptation strategies that can ensure a sufficient amount of water supply over long planning horizons. In the past, water resources planning has been done using single-objective benefit-cost analysis, where a single estimate of a project's costs and benefits is calculated to select funded projects. The calculation of monetary benefit functions, however, is heavily dependent on several critical assumptions. For example, the analysis must assume that the preferences of diverse stakeholder groups will not change in the future system states. Moreover, operational design and implementation of engineered water resources systems must consider a broad suite of risk-based performance objectives. This dissertation research advances water resources planning and decision support techniques that can confront the limiting challenges associated with classical approaches. We specifically advance a *many objective* approach using multiobjective evolutionary algorithms (MOEAs) that allows planners to generate and evaluate planning alternatives that can balance diverse planning goals and objectives.

This dissertation contributes two new many objective planning frameworks, collections of techniques that use many objective analysis to further our understanding of how to improve planning under uncertainty. The first framework, termed *de Novo* Planning, incorporates the concept of "learning" into many objective planning formulations. *De Novo* Planning addresses the fact that planning formulations themselves change as decision makers solve problems and analyze results. Global sensitivity analysis using Sobol' variance decomposition is used to determine an appropriate level of complexity for decision variables in the system. Multiple problem formulations are then constructed and solved using a MOEA to test the insights learned through the sensitivity analysis.

The second planning innovation is termed Many Objective Robust Decision Making (MORDM). MORDM addresses deep uncertainty, a situation in which stakeholders do not know or cannot agree on the full suite of risks that are posed

to their system. Deep uncertainty can severely impact the expected performance of planning alternatives in ways that are difficult to predict. This issue is especially relevant since most system planning under uncertainty is evaluated using a single best estimate of the distributions of data. Estimates from historical system information and their associated likelihoods, though, could be incorrect. For example, climate change can alter the magnitude and timing of streamflow availability, which makes the historical data an unreliable indicator of future events. Robust Decision Making (RDM) has been advocated as a way to address this issue, by evaluating a wide array of plausible futures to show future system vulnerabilities. The MORDM framework introduced in this thesis bridges many objective analysis with RDM, by evaluating solutions in the many objective tradeoff with an ensemble of alternative futures that investigate key assumptions and uncertainties, quantifying the solutions' robustness, and facilitating choice of robust solutions for a final negotiated decision.

The dissertation's planning innovations are demonstrated using two test cases with differing hydrologic characteristics and regulatory structures. The first test case explores how to improve the supply reliability of a single city in the Lower Rio Grande Valley (LRGV) of Texas. The LRGV case study uses risk-based planning triggers to control a city's use of a water market, with transfers between agricultural use and municipal supply. The goal is to highlight how non-structural adaptation such as water marketing can aid water management in the arid western U.S. Problems of water availability are also becoming more apparent in the eastern U.S., where water planning was traditionally focused on flood management and droughts were not often considered a serious issue. The second test case explores multi-sector long-term supply planning for the Lower Susquehanna portion of the Susquehanna River Basin in Pennsylvania and Maryland. Two many objective problem formulations for the Lower Susquehanna expose biases and challenges of classical planning formulations. Subsequent exploration of deep uncertainty suggests critical modeling assumptions for the test case. Insights from the Susquehanna test case have the goal of assisting reservoir planning for infrastructure systems in the eastern U.S.

Table of Contents

List of Figures

List of Tables

Acknowledgments

First of all, I would like to thank my advisor Dr. Patrick Reed. My interaction with Dr. Reed began when he was my honors advisor at Penn State in Spring 2004. Dr. Reed has taught me a lot over the years: how to succeed in graduate school, develop my skills as a teacher, and perform cutting edge research. He's always included me in research presentations, workshops, and discussions, and through his support I have traveled to the UK, Austria, Germany, and across the US to present my research. In addition to his helpful advice, he has led by example and also provided friendship and mentorship. I look forward to continuing to interact with Dr. Reed in my career in the long term.

I would also like to acknowledge and thank my thesis committee: Dr. Michael Gooseff and Dr. Peggy Johnson from Penn State's department of Civil and Environmental Engineering, Dr. Seth Blumsack from Penn State's Energy and Mineral Engineering department, and Dr. Greg Characklis from the University of North Carolina at Chapel Hill's Environmental Sciences and Engineering department. Each of you has brought a unique perspective that has aided the research presented here and my own personal development.

I sincerely thank the co-authors of the studies in this thesis. Dr. Reed has been an invaluable aid in the publishing process. Dr. Brian Kirsch (now at the Colorado School of Mines) and Greg Characklis helped write the study comprising chapter 3. Dr. Shanthi Nataraj and Dr. Robert Lempert from the RAND corporation were valuable co-authors on the study comprising chapter 4. I am collaborating on the research in Chapter 5 with Evgenii Matrosov from University College London (UCL).

I'd also like to acknowledge the large amount of assistance I've received with software and methods. Brian Kirsch was instrumental in helping with the LRGV simulation in chapters 3 and 4. Evgenii Matrosov provided critical assistance getting the IRAS-2010 translation up and running, and Ivana Huskova (UCL) wrote some of the code used to link IRAS-2010 to the optimization algorithm. I also acknowledge the kind assistance of Dr. Julien Harou at UCL; it was a pleasure

working on the London project with him over the last few months. I've appreciated spending time with the UCL folks in Leipzig, Vienna, San Francisco, and State College (maybe we can add a few more cities to this list!). Dr. Joshua Kollat, now heading his own company DecisionVis, provided the AeroVis visualization code as well as hours of assistance and discussion about all sorts of things. Josh taught me most of what I know about programming, which has definitely paid dividends over the years. The KNN stochastic simulations in chapter 5 were aided by source code from Dr. Balaji Rajagopalan's group at the University of Colorado Boulder. I also acknowledge the SRBC's OASIS data and discussions with Dr. Daniel Sheer that were a helpful aid in Chapter 5. Scenario Discovery scripts in the R programming language were originally coded by Benjamin Bryant, formerly at RAND Corporation. Shanthi Nataraj also helped run the analysis and provide valuable feedback. Last but certainly not least, this dissertation makes extensive use of software and codes from the Pat Reed group including ε-NSGAII, the Borg MOEA, and MOEAframework. Thanks to David Hadka for answering what seems like hundreds of questions over the last few years. Jon Herman helped with simulations in Chapter 5 and has provided useful discussions over the past few months. Dr. Reed's group has been a great source of intellectual stimulation and I look forward to continuing interactions with group members for a long time to come.

I also acknowledge Keith Sawicz, Rachel Urban, Matthew Woodruff, and others on the fourth floor who have made graduate school a fun experience. Thanks to my horse trainer Dr. Suzanne Myers, my fellow riders (and also the horses!) for making the last few months of thesis writing tolerable. I also want to thank my colleagues on the climate change project for a nice working experience, including Josh Kollat, Wilbert Thomas, and Arthur and Mary Miller. This goes out to the loving memory of David Divoky, who also worked on that project and provided a great example of how to have fun and be intellectually curious while working on difficult projects.

Last but not least, I would like to thank my friends and family who have provided an immeasurable amount of support. My parents have always encouraged me to reach for the sky and do my best in everything that I attempt. They are always understanding of my troubles and help me when I'm in need. I wouldn't be where I am today if it wasn't for them. My best friend and staunchest advocate in graduate school has been my roommate Emil Laftchiev. He's shown me that graduate school teaches you as much about life as it does about research. Together with him I bought an aquarium, lost weight, learned how to ride a horse, and learned how to become a better, well-rounded person. I owe him a tremendous debt of gratitude and wish him the best in all his future endeavors. In the end, graduate school is a team sport as much as anything. I'm deeply honored to have played these innings with such a wonderful group of people.

Dedication

Dedicated to my friend Emil I. Laftchiev

Chapter 1

Introduction

Water resources planners and managers have long operated under the assumption that the historical record of water supply and demand provided a reliable estimation of future conditions. However, the validity of this approach is threatened by population pressures and environmental change, which can cause past records of streamflow availability to inadequately capture the future performance within these systems [1]. For example, climate change modifies the hydrologic cycle [2], changing surface water availability and increasing the likelihood of droughts. Water demand is highly uncertain [3], affected by climate change [4] and population growth [5]. Additionally, new water uses can emerge, such as water to support hydraulic fracturing for electricity generation [6]. These challenges can be considered deeply uncertain [7–9], meaning that decision makers cannot fully conceptualize or agree upon the range of possible risks in their system. This dissertation will utilize multi-objective tools and improved planning frameworks for identifying robust planning alternatives under conditions of deep uncertainty.

Water projects are typically evaluated using cost-benefit analysis [10], where project benefits are commensurated or transformed into their expected monetary value and compared to the project's costs to determine if benefits outweigh the costs. The classical approach of expressing benefits in terms of their dollar values has several significant limitations. In order to perform the transformation of benefits into dollar amounts, the analyst's assumptions (such as the discount rate) must maintain valid in the future system states [11]. Also, the recipients of the benefits (such as the residents of a river basin) may have preferences that change in

the future. Simply providing a proper baseline for a project's cost-benefit analysis is a significant challenge (i.e., what the state of the world would be like without the project [12]), especially when multiple policy alternatives are considered [13]. Some have argued that under cost-benefit analysis, political pressure is the only determinant of whether or not a project gets implemented [14].

Cost-benefit analysis is equivalent to a single-objective maximization, which finds a single optimal design that maximizes benefits for a system. The optimal solution to this problem, though, may prove to be inferior [15] when new objectives such as preservation of environmental quality or meeting conflicting water uses are considered. Banzhaf [16] outlines a historic disagreement between economists that advocated the single benefit-cost ratio and Arthur Maass and others at the Harvard Water Program that believed that a set of multiple objectives could be explicitly used in the design and funding negotiations for federal projects[1]. While multi-objective planning could have helped address some fundamental flaws of single-objective welfare economics [16], traditional benefit-cost ratios were adopted in U.S. Water Resources Council standards.

This dissertation demonstrates a *many objective* approach, in which planners can generate and evaluate alternatives with respect to four or more objectives simultaneously, avoiding lower-dimensional, myopic problem formulations that seek to confirm decision-makers' preconceptions of the best alternative for their system. Our approach generates high quality approximations to the Pareto optimal set of solutions, solutions that are better than all other feasible solutions in at least one objective. Considering multiple objectives lends problem insight; we demonstrate in Chapter 2, for example, that commonly-employed decision maker heuristics (rules of thumb based on intuition) that attempt to minimize water surplus and wasted water transfers inadvertently lower the reliability of water supplies in a modeled system. To generate our planning alternatives, we use multi-objective

[1] As cited by Banzhaf [16], the exchange between the two groups occurred in public [17–22]. Maass argued that the role of public investment in the U.S. was not merely to improve economic efficiency but also to redistribute income. He advocated the political process to determine appropriate tradeoffs between these objectives [17]. In his reply Haveman [18] essentially argued that different types of projects should be designed to either meet economic efficiency or income redistribution, but not both. Maass' position was that the political system is inherently multi-objective; without considering multiple objectives explicitly, "the tradeoffs are implicit, generally inefficient, and sometimes internally contradictory" [21].

evolutionary algorithms (MOEAs) [23], modern solution tools that use selection of good solutions and variation on those solutions to generate high-quality approximations to the Pareto optimal set in a single algorithm run. Use of MOEAs does not require many simplifications of the planning problem, since the analyst can integrate a full-complexity simulation model into the optimization routine. Although the benefits of combining simulation models were clearly recognized in early water resources planning and management efforts [24–26], very recent breakthroughs in computational power and MOEA-based search are now enabling studies to fully realize the benefits of these tools. This computation approach, coupled with advanced three-dimensional visualizations, allows us to visualize the tradeoffs between planning objectives in a manner originally intended by the Harvard Water Program (see Maass et al., page 308 [24]) but previously unavailable due to computational and conceptual limitations.

While improved planning frameworks and MOEAs strengthen water resources systems analysis, analysts must also consider the fact that planning formulations themselves change as decision makers solve problems and analyze results. In this sense, the problem formulation itself is "nonstationary", and decision makers form new hypotheses and pursue modified decision-making objectives as they solve their problems [27, 28]. Modern decision support uses the paradigm of "constructive decision aiding" [29], allowing a collaborative process for discovering problem formulations that capture evolving decision-making goals. In chapter 3 this dissertation proposes and demonstrates a *de Novo* Planning framework that will incorporate this new problem learning into many objective planning formulations.

A key limitation of planning studies is that they often use a single best estimate of the distributions of future inflows, supply costs, and demand trajectories to evaluate alternatives. Recent studies evaluating risks as broad as climate change planning [30] and terrorism risk insurance [31] have shown that decision makers should be cognizant of the ramifications when estimates of problem information and likelihoods are wrong. Robust Decision Making (RDM) [9] has been advocated as a way to address this issue, by evaluating a wide array of plausible futures and plan for system vulnerabilities. RDM approaches, though, have not emphasized the role of generating planning alternatives, as mentioned above. Chapter 4 contributes a framework termed Many Objective Robust Decision Making (MORDM) that

bridges many objective analysis with RDM to facilitate decision-makers' choice of robust solutions for planning.

The planning innovations are demonstrated using two test cases: a single city's municipal supply using a water market in the Lower Rio Grande Valley (LRGV) of Texas (chapters 3-4), and multi-sector long-term supply planning for the Lower Susquehanna basin on the border of Pennsylvania and Maryland (chapter 5). Each basin has a unique set of challenges and exhibits different hydrologic characteristics and regulatory structures. The following section outlines the remainder of the dissertation and provides a brief summary of each chapter.

1.1 Overview of Chapters

1.1.1 Chapter 2: Background

Chapter 2 provides the reader with background information on the core topics covered in this dissertation. The chapter defines many objective analysis and provides a motivating example of how it can enhance environmental decision-making. Then, the challenge of deep uncertainty is introduced, motivating the planning innovations that will be introduced in this work. Technical details on the quantitative methods used are then given. Finally, the chapter concludes with information about the motivations for the regional case studies used to test the frameworks.

1.1.2 Chapter 3: de Novo Planning

The first planning innovation in this dissertation moves beyond the idea of using a single, static problem formulation (i.e., decision variables, objectives, and constraints) for decision support. Chapter 3 presents *de Novo* planning, which uses multiple problem formulations to help incorporate learning into the decision support process. We use global sensitivity analysis using Sobol' variance decomposition to determine an appropriate level of complexity for decision variables and aid in choosing an appropriate set of objectives and constraints. A suite of multiple problem formulations is then constructed and solved using a MOEA to explore the insights learned through sensitivity analysis. The chapter uses the LRGV risk-based water supply management problem to demonstrate the frame-

work. Use of the framework illustrates how to adaptively improve the value and robustness of our problem formulations by evolving our definition of optimality while discovering key tradeoffs. Chapter 3 was adapted from a study published in *Environmental Modelling and Software* [32] co-authored with Patrick M. Reed, Gregory W. Characklis, and Brian R. Kirsch.

1.1.3 Chapter 4: Many Objective Robust Decision Making

Chapter 3 presents an analysis of a water supply planning problem under uncertainty, where inflows, demands, and prices are characterized by an estimated probability distribution, which is used to estimate a supply portfolio's expected performance. A key issue with this approach is that we cannot ascertain whether the selected alternative portfolios are robust to core assumptions or changes in likelihoods across the system's uncertainties. Decision support strategies that use the concept of robustness such as Robust Decision Making (RDM) can help address this issue. RDM seeks to evaluate the performance of policy strategies over an ensemble of deeply uncertain trajectories of the future. It then helps decision makers choose robust alternatives and characterize which deeply uncertain factors are most important in causing performance vulnerabilities. Chapter 4 seeks to overcome an important limitation in the prior RDM literature by bridging MOEA search to generate alternatives with the RDM analysis. This framework represents an innovation termed Many Objective Robust Decision Making (MORDM). MORDM extends the work of chapter 3 by improving the *ad hoc* method of choosing alternatives by relying solely on the expected values of a single set of assumptions and likelihoods in the Monte Carlo simulation. We use the LRGV test case to demonstrate the framework, with the goal of identifying parsimonious and robust rules for the city to operationally exploit the water market in their supply portfolio. Chapter 4 was adapted from a journal article in press at *Environmental Modelling and Software* [33] co-authored by Shanthi Nataraj, Patrick M. Reed, and Robert J. Lempert.

Chapters 3 and 4 present a sequence of planning innovations with the goal of helping decision makers identify appropriate problem formulations and choose robust planning alternatives, using the LRGV test case. Although the LRGV

test case has significant value, it focuses on a single city's municipal supply, and augmenting this supply with water marketing transfers is limited to regions in which water marketing is allowed, such as the western United States. In the eastern U.S., water marketing is often prohibited, and water managers must face challenges in managing supply for multiple sectors without the ability to build new infrastructure. This different regulatory context is addressed with a new test case in chapter 5.

1.1.4 Chapter 5: Formulation Biases for the Lower Susquehanna

Chapter 5 builds a new Lower Susquehanna test case to demonstrate how MORDM can be used to test the robustness of management alternatives for water supply infrastructure in the eastern U.S. Two many objective problem formulations are developed using the Interactive River Aquifer Simulation (IRAS)-2010 water resource system simulator and a recently introduced MOEA termed the Borg MOEA. The first formulation is reflective of the classic history-based deterministic water resources systems planning approach that still dominates practice. The second formulation moves beyond the historical record by considering streamflow and evaporation uncertainties modeled using a K-Nearest Neighbor stochastic simulation strategy. MORDM is then used to contrast the robustness of the two formulations' Pareto approximate solutions' robustness to deep uncertainties, associated with water demand targets, prolonged droughts, and increased interannual variability of streamflow and evaporation. The optimization-simulation and MORDM components of this study posed severe computational demands in excess 1,000,000 hours of computing. Thus in addition to introducing the test case implementation, chapter 5 also discusses application of High Performance Computing technology to carry out the experiment.

Beyond the case study-specific insights, a key goal in this study is to expose the potential negative consequences or formulation biases that result from the use of deterministic planning based historical hydrology. Although our deterministic many objective formulation contributes a more comprehensive characterization of the Lower Susquehanna test case's multi-sector demand tradeoffs relative to

classic single objective planning [34–39], deeply uncertain changes in the SRB have the potential to exacerbate drought risks. Moreover, this study explores if the sole consideration of hydrological uncertainties (i.e., stochastic generation of streamflow and evaporation) is sufficient to discover robust alternatives for the system while simultaneously identifying the key deep uncertainties that control its vulnerabilities. A journal article based on this chapter, co-authored by Patrick M. Reed and Evgenii Matrosov, will be submitted in Spring 2013.

1.1.5 Chapter 6: Concluding Remarks

Chapter 6 suggests the conclusions of this research, its contributions, and provides a guide for future work to expand on the contributions herein.

Chapter 2

Background

This chapter provides background for topics covered in this dissertation. Section 2.1 defines many objective analysis and presents an illustrative example. Section 2.2 first introduces a challenge called deep uncertainty. The section then introduces two new planning frameworks, which are collections of techniques that use many objective analysis to further our understanding of how to address deep uncertainty. Subsequently, section 2.3 provides background on quantitative tools that will be used throughout this work. Finally, section 2.4 outlines the two regional test cases that will be used to demonstrate the approaches used in this thesis.

2.1 Motivation

Chapter 1 introduced the challenges of providing a sufficient supply of water in the face of conflicting demands and environmental change. The field of *water supply planning* seeks to improve the reliability of water supplies, or their ability to sustainably meet their demands in the long term. This includes designing new water infrastructure, improving management of existing systems and implementing non-structural adaptation such as water marketing (transfers of water between use sectors or regions). The historical literature in the field [24, 25] has recognized that there are multiple criteria to consider when designing and evaluating these systems, including reliability [40], economic benefits [41] and environmental conditions [42]. These objectives often conflict and force water managers to make seemingly arbitrary choices between the different planning goals. For example, in

a reservoir, managers may have to release water from storage to maintain environmental quality downstream, especially in low flow conditions [43]. Such releases meet environmental objectives, but they conflict with other water demands such as municipal supply.

This dissertation uses water supply planning as a means to develop new planning frameworks using many objective analysis. The rest of this section proceeds as follows. Section 2.1.1 defines many objective analysis. Then, section 2.1.2 illustrates the use of many objective analysis within the Lower Rio Grande Valley (LRGV) test case. In section 2.1.3, we show the effect of changing our modeling assumptions on solutions' performance.

2.1.1 Many Objective Analysis

A planning problem with multiple conflicting objectives is often transformed into a single objective problem by "aggregating" the objectives together [44][1]. The process of aggregation requires a decision maker to determine the relative importance of each objective *a priori* (before the analysis begins)[2]. For example, a decision maker could place a high value on municipal supply, at the expense of providing environmental services. As discussed in Chapter 1, there are critical issues with the *a priori* approach, since aggregating the objectives may favor some of them over others in unpredictable ways [45]. In contrast, this thesis solves multiple objective problems using an *a posteriori* approach [46]. By not requiring an *a priori* description of preferences, the approach allows a decision maker to consider multiple objectives explicitly and simultaneously.

The following text rigorously formulates a generic multiple objective optimization, similar to the problems solved throughout this dissertation. Consider several different objectives for water planning. The cost of a system may be quantified by an objective f_1, while the reliability (the probability of meeting system demands) could be expressed by a second objective, f_2. A planner desires to simultaneously minimize the cost (f_1) and maximize the reliability (f_2)[3]. Constraints are desired

[1]Consider two objectives, x and y, where one wants to minimize x and y. One possible aggregation of the objectives is $z = ax + by$, where a and b are weights on the objectives and z is minimized.

[2]In other words, the decision maker must choose the weights a and b in advance.

[3]In equation 2.1, each objective is said to be minimized. To transform a maximization objec-

performance targets, beyond which the system's performance is considered unacceptable. An example constraint, c_1 could impose a restriction that the reliability of the system must be greater than or equal to 98%. Design levers (or decision variables[4]) are aspects of the system that are under the manager's control and can change. Design levers are expressed by a vector $\mathbf{l}$. These components are the fundamental aspects of a multi objective problem formulation[5] shown in equation 2.1.

$$\begin{aligned} \text{minimize } \mathbf{F}(\mathbf{l}) &= (f_1, f_2, \ldots f_P) && (2.1)\\ \forall \mathbf{l} &\in \Omega \\ \text{Subject to : } c_i(\mathbf{l}) &= 0 \; \forall i \in [1, q] && (2.2)\\ c_j(\mathbf{l}) &\leq 0 \; \forall j \in [1, r] && (2.3) \end{aligned}$$

In the equation, the P planning objectives are combined into a vector-valued objective function $\mathbf{F}(\mathbf{l})$. Decision variables (design levers) can be expressed as real-valued, integer, or binary variables. The equations show two different types of constraints (limits on acceptable performance): equality constraints in equation 2.2 and r inequality constraints in equation 2.3. In the context of constraints, *feasible* solutions are defined as those that can meet all the imposed constraints, and the set of all feasible decision variable values is termed Ω.

The solution to the multiple objective problem in equation 2.1 is not a single optimal solution (design for the system). Rather, the solution is a tradeoff set of *multiple* solutions, using the concept of Pareto optimality or non-domination to define the tradeoff. Assuming minimization of all objectives, consider two solutions $\mathbf{l}_1$ and $\mathbf{l}_2$ with associated objective function values (i.e., $F(\mathbf{l}_1) = [f_1(\mathbf{l}_1) \ldots f_p(\mathbf{l}_1)]$). The solution $\mathbf{l}_1$ dominates $\mathbf{l}_2$ if and only if $\forall i, f_i(\mathbf{l}_1) \leq f_i(\mathbf{l}_2)$ and there exists j such that $f_j(\mathbf{l}_1) < f_j(\mathbf{l}_2)$. In other words, $\mathbf{l}_1$ has at least equal objective function performance to $\mathbf{l}_2$ in all objectives, and it is better in at least one objective. Extending

tive into one that is minimized, the value is multiplied by -1.

[4]Chapters 4 and 5 use the XLRM terminology from policy analysis (e.g., [47]), which stands for Uncertainties, Levers, Relationships, and Measures. Levers are equivalent to decision variables for optimization, and quantitative Measures refer both to objectives and also to constraints.

[5]See Cohon and Marks [48] for one of the first descriptions of multiple objective optimization in the water literature.

this definition to the entire decision space Ω, a solution is termed Pareto optimal if it is non-dominated with respect to all other feasible solutions[6].

This dissertation builds on a recent body of research called *many objective analysis* [50–54]. Many objective analysis refers to multiple objective problems of four or more objectives. These optimization problems can be difficult to solve, so one aspect of many objective analysis is developing competent solution tools that can find solutions on or near the true Pareto optimal front (covered later in section 2.3.1). After finding a tradeoff set, the analyst needs a way to observe how the conflicting objectives compare with one another and choose a negotiated solution or design. To do this, our approach uses interactive tradeoff curves (section 2.3.3) that help decision makers find a "balance" between objectives [55].

2.1.2 Many Objective Analysis of the LRGV

This section illustrates example results of the many objective analysis process. The results will also help demonstrate how many objective analysis can help planners confront cognitive challenges in their planning problems.

The Lower Rio Grande Valley (LRGV) test case is adapted from Kasprzyk et al. [56]. The LRGV faces rising municipal demands and a variable amount of water supply. The planning problem seeks to help a single city use a portfolio of water supply instruments to increase the reliability of its water supply. Initially, the problem is defined with two objectives and a single decision (how many reservoir rights to obtain). Managers minimize the cost of their reservoir rights and maximize the system reliability. Figure 2.1 is an example of a tradeoff plot for these two objectives. Each point as a different design for the system (a volume of reservoir rights), and the objectives are plotted such that the lower left corner represents the preferred direction for both objectives. Solution 1 "dominates" the gray region, labeled region i. This is because solutions that would fall in region i have higher

[6]This thesis is concerned with simulation-optimization problems where it is impossible to obtain the derivatives necessary to prove optimality. In the absence of derivatives, true Pareto optimality can only be shown by enumerating all possible feasible alternatives and calculating non-domination with respect to all feasible alternatives. Solution sets in this thesis are instead referred to as non-dominated sets or Pareto-approximate sets. They are the best known approximation to the Pareto optimal front, obtained by performing many multi-objective evolutionary algorithm optimization runs and sorting the results together. For more details on multiple objective optimization, see Deb [49] and Coello Coello [46].

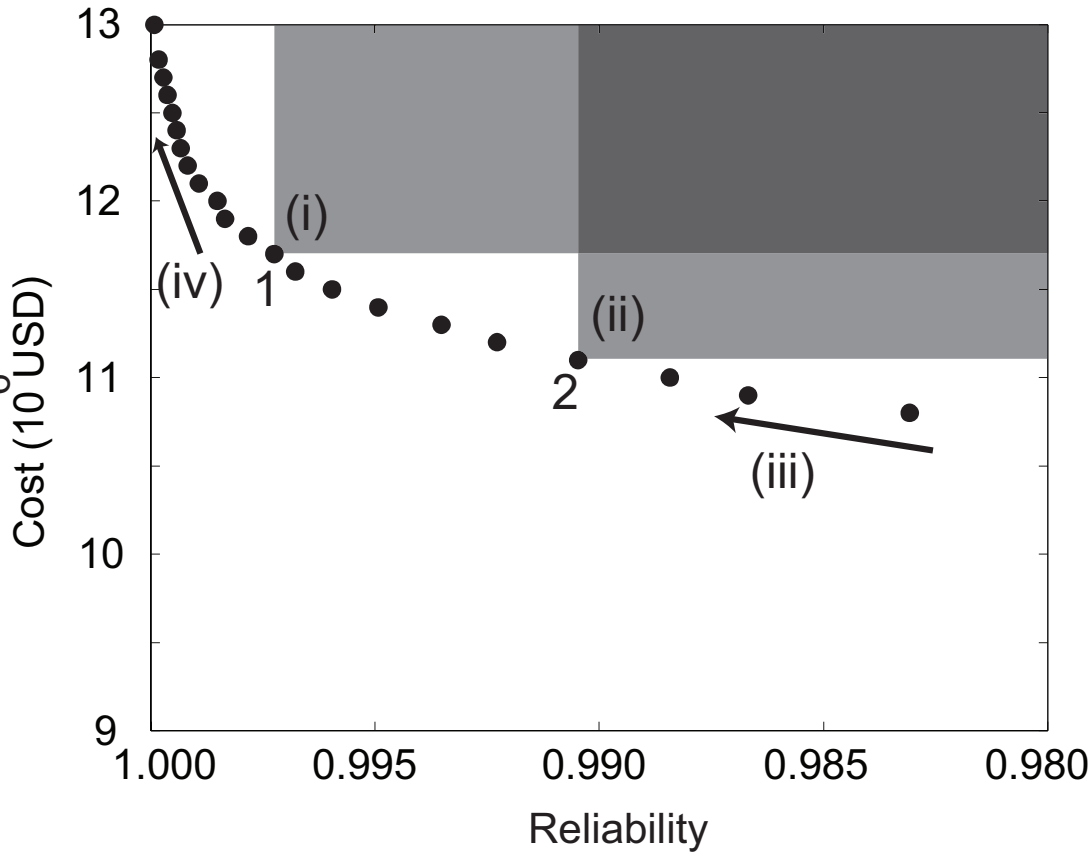

Figure 2.1. An example of a non-dominated tradeoff, in which a city buys excess water to meet reliability using only permanent rights. Highlights (i) and (ii) illustrate the concept of non-domination. For example, solutions in region (i) have higher cost or lower reliability than solution 1. Similarly, solution 2 dominates region (ii). Highlights (iii) and (iv) show the increasing marginal cost of adding reliability; the same increase of 0.005 reliability costs more at arrow (iv) than at arrow (iii). Adapted from [56].

cost and lower reliability than solution 1. If a decision maker values both cost and reliability, he or she would not choose a solution in this region. Similarly, solution 2 dominates region ii. Solutions 1 and 2 are incomparable, though, in the sense that while the first has better reliability, it also has higher cost. Extending this comparison to all solutions in the tradeoff, we have developed a set of solutions that is non-dominated with respect to other feasible solutions for the problem. The figure shows that an increase in the permanent water supply rights is necessary to maintain supply reliability. A 0.5% increase in reliability from 98% costs less (arrow iii) than the same increase from 99.5% to 100% (arrow iv).

Two key questions arise from the analysis of figure 2.1. First, can the city have

high reliability but also lower its costs? Second, are there other planning objectives that are important for this problem? In general, these questions motivate two *cognitive challenges* that are aided by many objective analysis, borrowing concepts from the cognitive psychology literature.

The first phenomenon is called "cognitive myopia" by Hogarth [57]. Hogarth's work showed that decision makers often lack creativity when solving problems, focusing on a small number of aspects of the problem. In the context of many objective analysis, cognitive myopia is when decision makers lack a full understanding of their alternatives by excluding key planning objectives. In water planning, there are several objectives beyond cost and reliability that are important, such as preserving flow for environmental services. Brill et al. [15] provide a theoretical explanation of what happens when a new objective is added to the problem. An analyst may find an "optimal" solution that minimizes cost. When the second objective is added, the optimal solution will likely be inferior with respect to the new objective. Many objective analysis helps confront cognitive myopia by solving problems with a large number of planning objectives and identifying solutions that balance multiple objectives simultaneously.

The second cognitive challenge is called "cognitive hysteresis" by Gettys and Fisher [58]. Cognitive hysteresis is when the initial conceptualization of a problem biases a decision maker's ability to find new solutions to the problem later on. The first formulation of the LRGV (figure 2.1) only uses reservoir rights. Managers assume that including transfers of water into their water supply could "waste" water resources if the transfers are not used. In doing so, the decision makers are actually underestimating the actual system risk [59], since rights may not be fully allocated if inflow to a reservoir decreases. Although the first formulation in figure 2.1 confirms that increasing the reservoir rights is the only way to increase reliability, subsequent formulations show otherwise. Kasprzyk et al. [56] added market-based instruments, which transfer water from the agricultural sector to municipal supply. Market-based planning solutions identified in our analysis could maintain high reliability and reduce costs. In other words, we can confront cognitive hysteresis by introducing new policy levers that force decision makers to go against their preconceptions (i.e. performance is only maintained using a costly surplus of permanent rights to reservoir water).

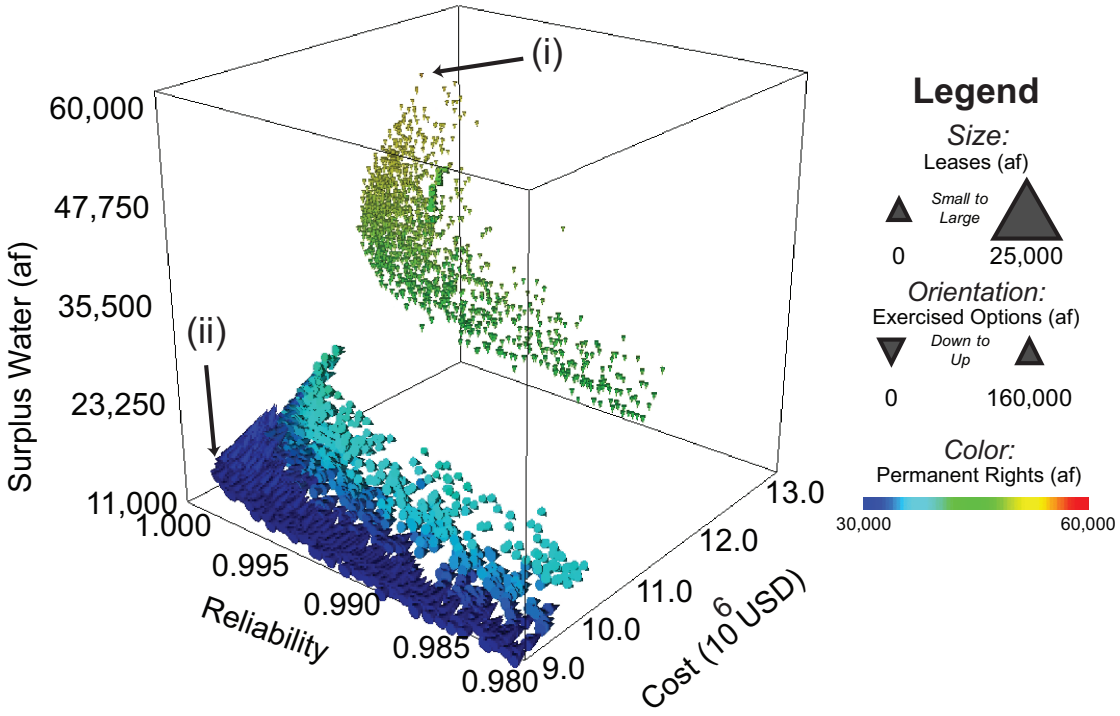

Figure 2.2. A glyph plot that shows several objectives from the full 6-objective tradeoff set. The spatial axes show cost, reliability, and surplus water, while the other plotting mechanisms visualize instruments in the portfolio. The figure shows that reduced water surplus can be easily achieved by considering the market. Moving from solution (i) to solution (ii) lowers surplus water and cost while maintaining high reliability. Adapted from [56].

Figure 2.2 is an example of a tradeoff plot for the many objective version of the LRGV planning problem[7]. In figure 2.2, the spatial axes show the cost and reliability with the surplus water shown in the vertical axes[8]. Surplus water is

[7]For the full definition of the problem, please consult Kasprzyk et al [56], and section 3.2.3 later in this thesis. A brief description is included here. We use 8 decision variables that include market-based supply instruments in addition to permanent rights. The variables control a portfolio of 3 supply instruments: permanent rights, options (transfers of water at a fixed price and fixed time) and leases (immediate transfers with a variable price). Six planning objectives are used. Cost and reliability are retained from figure 2.1. We minimize cost variability, a measure of the worst-case costs from market transfers. Surplus water, the volume of water carried over from year to year, is minimized. Dropped transfers, market transfers that expire from nonuse, are minimized. Finally, we minimize the number of leases as a proxy for transactions cost. Constraints on reliability and cost variability ensure high performance of each tradeoff solution.

[8]Our tradeoff plots are made using interactive software (see Kollat and Reed [60]). Within the software, the decision maker can choose different variables for each plotting axis. The plotting axes can show objectives, decisions, and other variables of interest (i.e., not all objectives must

included to show how the municipal supply could conflict with other regional uses. By minimizing the surplus water, we are showing how portfolios could free up storage water that could be used for other purposes, such as other demands or environmental releases. In the figure, the green and yellow solutions (highlight i) use approaches similar to the approach from figure 2.1 (that is, they have high permanent rights usage). While these solutions achieve high reliability, this comes with an associated high value for surplus water and potentially highly negative impacts on the availability of water for ecological services. Alternatively, highlight ii in figure 2.2 shows a solution that has higher market use[9], with lower cost, high reliability, and reduced surplus water (i.e., more water for ecological services).

2.1.3 Impact of Changing Assumptions: Drought Analysis for the LRGV

The results shown in the previous figures used a ten-year monthly planning scenario to evaluate solutions' performance. How do solutions' performance change, if we modify the assumptions behind the analysis itself? Kasprzyk et al. [56] coupled the driest year in the historical record with maximal demands from the empirical historical demand distributions to create a new drought scenario. The drought scenario was then used as a severe stress test for selected solutions in the LRGV tradeoff.

Figure 2.3 presents these results, where four selected solutions labeled solutions 3 - 6 in the prior study were evaluated in the drought scenario[10]. In January, the city begins with the modeled initial condition of S_{Jan} equal to 30% of the portfolio's permanent rights. Any water not used to meet demands is called "carry over" water in the following months. In February through December, the hashed lines represent water allocated to the city's permanent rights. Similarly, the light gray and white boxes represent leased and optioned water purchased by the city in the

be shown simultaneously).

[9]The cones are larger (they use more leases) and pointing toward the top of the figure (they use more options). The blue color indicates that their permanent rights usage is lower than the solution in highlight i.

[10]The full explanation of selected solutions is presented in Kasprzyk et al. [56]. In brief, there are two types of solutions considered. Solution 3 typifies "permanent rights dominated" solutions, with high values for rights (N_R) and lower values for the rest of the variables. Solutions 4-6 use the market to a higher degree.

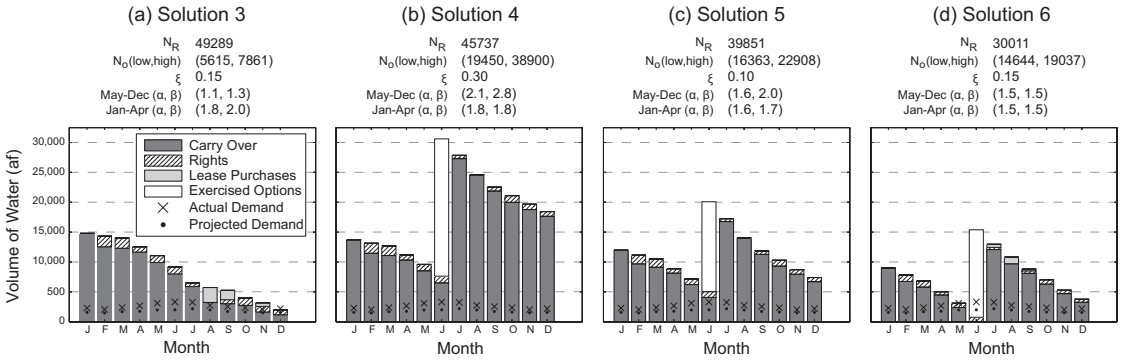

Figure 2.3. Monthly water balance for selected portfolios evaluated during the drought scenario in Kasprzyk et al. [56]. Supply from all three instruments is shown with the height of bars, while the required demand is shown with an x. When demand exceeds supply, a failure occurs. Decision variable values are shown above the graph, representing permanent rights (N_R), the adaptive options contract (high and low N_O, with the ξ threshold), and strategy for exercising leases and options (α and β for two parts of the year).

previous month. Therefore, the height of the bar is the total amount of supply available in each month. Demanded water is plotted with an x symbol, and in the drought scenario the demands are set to be the maximal demands in the historical distribution. A failure occurs when the demand is higher than the supply.

The drought scenario is difficult for our identified solutions because the solutions use fixed supply-demand ratios to control the city's use of the water market. The supply-demand ratios use expected historical supply, shown with a dot in the figure. The drought scenario's demands are much higher than the expected supply. A key question in this section is: can solutions identified using the ten year scenario still exhibit high performance in an additional severe drought?

All solutions shown in the figure were optimized to have very high performance in the 10-year scenario. Figure 2.3a shows that reservoir rights-dominated solutions such as solution 3 are less adaptive and more prone to very severe water supply failures. Under the reservoir-dominant planning of solution 3, the unexpected drought conditions in figure 2.3a caused the city to lose much of its water supply toward the end of the drought year. Not only did this approach inadvertently raise costs and surplus water in the 10-year simulations, but the solution also had more failures in the drought. Solutions 4-6, however, were able to use the water market to obtain transfers that were able to ensure reliability in the drought.

2.2 Planning Innovations

2.2.1 The Challenge of Deep Uncertainty

In the previous section, high-quality planning alternatives exhibited severe failures under a new drought scenario. This was interesting because each selected solution had very high reliability in the original 10-year simulation. Our changing modeling assumptions are captured in a concept called deep uncertainty. Deep uncertainty refers to a condition in which decision makers do not know (or cannot agree upon) the appropriate problem formulation that governs their system or the full suite of risks for their system[11].

This remainder of section 2.2 briefly introduces new planning innovations that use and extend many objective analysis to help confront deep uncertainty. Section 2.2.2 introduces *de Novo* Planning. *De Novo* Planning helps address uncertainty in our conceptualization of the problem formulation (objectives, decision variables, and constraints). The *de Novo* Planning framework can help a decision maker determine an effective formulation to use. We also require a method to help understand how this choice of formulation is affected by deep uncertainty in modeling assumptions, such as the input hydrologic data that is used within a simulation model. Many Objective Robust Decision Making explores this type of deep uncertainty and is introduced in section 2.2.3.

2.2.2 de Novo Planning

Milly et al. [1] outlines the challenge of "nonstationarity": the statistical properties that underly the natural environment are not constant, due to global climate change, land use change, and increasing population. This emerging focus on the nonstationarity of the data that describes our systems is certainly warranted.

[11]The term "deep uncertainty" is used in Lempert et al. [61] and refers to the concept of uncertainty discussed in Professor Frank Knight's 1921 book *Risk, Uncertainty and Profit.* There are conflicting definitions of what Knight by deep uncertainty [8,62–64]. Some researchers believe that Knight defined *risk* as events with a "knowable" distribution while *deeply uncertain* events have no possible probability distribution [8]. Our interpretation follows Langlois and Cosgel [8]. They show that Knight was concerned with planners' "ignorance of the future". To Knight, a deeply uncertain situation was one in which decision makers cannot classify the full suite of conditions that could cause risk to the system, in addition to the events' probabilities.

However, this dissertation seeks to address an additional type of nonstationarity in planning problems, which is a *structural* nonstationarity in the problem's definition. Recall that the problem formulation was defined in section 2.1.1 to be the decision variables, objectives, and constraints for planning. According to Rittel and Webber, problems such as water resources planning are considered "wicked" [65], because the problems lack a single static formulation. Instead, planning objectives are often discovered during the analysis itself [66]. Furthermore, decision makers often find new decision levers that could improve outcomes during the planning process. Liebman [67] builds on this discussion by suggesting that the purpose of quantitative simulation is to provide a formalized way to construct hypotheses about problems and learn more about our systems.

Constructive decision aiding [29] is an emerging focus that seeks to improve upon the traditional single-objective optimization approaches that were popular in the 1950's and 60's. This approach uses a collaborative process for discovering problem formulations that capture evolving decision-making goals. Planners can propose multiple hypotheses about the most important aspects of their systems, in order to discover new properties about their system during the planning process. The planners' evolving knowledge then drives changes in models, decisions, and objectives [68].

De Novo Planning in this dissertation builds on the work of Zeleny, who introduced a method called *de Novo* Programming. He challenges operations researchers to: "design optimal systems, do not merely 'optimize' given systems." [69] This statement underscores that when the problem formulation is kept constant, decision makers may miss key insights into their system. Zeleny shows how using a single Linear Program (LP) formulation for resource allocation wastes resources that were not binding constraints in the LP. *De Novo* Programming progressively changes the LP formulation to improve revenue, avoid squandering any resources, and gain new insights into the best production mix. Zeleny later wrote that objectives, decisions, and model structure are all in "continuous flux" and change in order to create cognitive equilibrium with decision makers [68]. In this dissertation, we build on Zeleny's work by introducing *de Novo* Planning: a many objective analysis framework in which iterative changes in the definitions of planning objectives, decisions, and constraints are used to incorporate knowledge gained when

solving the problem into decision support. This effort supports the suggestion by the National Research Council for an iterative "deliberation with analysis" framework [28]. The framework incorporates sensitivity analysis using Sobol' variance decomposition, which will be described in section 2.3.2.1. It will then be fully defined and demonstrated in chapter 3.

2.2.3 Many Objective Robust Decision Making

The previous section showed how *de Novo* Planning confronts the deep uncertainty in the problem definition itself. A further challenge is determining how the chosen problem formulation performs under a second type of deep uncertainty: not knowing the full suite of risks to a system or their probabilities. Land use change, the depletion of resources and climate change are three examples of human-induced changes that introduce deep uncertainties into planning problems [1, 70, 71]. Decision makers often use scenario analysis [72] to help plan for such challenges. Scenario analysis uses a small number of plausible values for key planning variables (such as economic growth) to create storylines for future conditions in a system. As highlighted by Groves et al. [73], there are several important limitations associated with using scenario analysis for deeply uncertain factors within a planning problem. First, a large number of different factors can shape the future, and a small number of scenarios cannot adequately cover all interactions between the different factors. Additionally, there is little guidance on how scenarios should inform decision making. For example, in climate change planning, a single value for global population growth is typically used in each scenario [74, 75], causing subsequent decision making to be contingent on the assumed value.

Decision support strategies that use the concept of *robustness* can help address these challenges, by identifying strategies that perform well across many different assumptions regarding the deeply uncertain factors [9, 30, 76, 77]. Robust Decision Making (RDM) [9, 47, 78, 79] is one such method to characterize robustness. RDM evaluates the performance of policy strategies over an ensemble of deeply uncertain trajectories of the future. Decision makers select plausible ranges for each deeply uncertain factor, but they do not *a priori* select scalar values for a small number of scenarios. Instead, RDM employs statistical algorithms to "discover" scenarios,

which are ranges of the exogenous factors[12] that in combination cause poor performance [78]. RDM therefore provides a tool for decision makers to determine how changes in their assumptions about exogenous factors affect the performance of their planning strategies. For complex environmental systems, this is especially useful because it can help planners determine the impacts of their predictions and assumptions on the decision making process. Drought forecasting, for example, requires forecasts of runoff, water storage, and return flow. Mitigation strategies are based on the assumptions behind these forecasts, but there are severe political ramifications when these estimates are wrong [80]. By avoiding *a priori* estimates of values of the deeply uncertain factors, RDM avoids a common decision maker bias of being overly confident about current management strategies and timid in adopting planning innovations that reduce risk [59]. RDM has been successfully demonstrated on a wide range of problems, including evaluating terrorism risk insurance [31] and in water planning [73]. Section 2.3.2.2 gives details of the RDM computations and an example of its use from the literature.

Most prior work in RDM focused on using a single evaluation metric to define performance and has not strongly emphasized how decision alternatives should be generated in support of the decision-making process. To overcome these limitations, the second major planning innovation in this dissertation seeks to expand the RDM and SD methodology to incorporate many objective analysis. The new framework is termed Many Objective Robust Decision Making, and it is demonstrated in chapters 4 and 5 of the dissertation.

2.3 Quantitative Tools

2.3.1 Multi-Objective Evolutionary Algorithms

The main task of many objective analysis is to find solutions on the non-dominated tradeoff between multiple objectives. This dissertation explores many objective problems in water supply planning, which is appropriate since the field has made important contributions to the broader literature of multiple objective planning.

[12]Exogenous factors are those in the modeled system considered not under the direct control of the decision makers, in contrast to the modeled factors considered part of the decision makers' choice set.

In his review of early multiple objective optimization, Coello Coello [46] cites the Harvard Water Program (HWP). The HWP was one of the first efforts to use quantitative models to evaluate water resource systems for different objectives[13]. The HWP 1962 monograph [24] even shows an "edges" visualization in which flood control, irrigation, and power generation are shown on three axes of a cube, mirroring the many objective plots shown in this dissertation (e.g., figure 2.2). Cohon and Marks [82] later solved an optimization problem explicitly for two objectives: maximizing net benefit and maximizing environmental quality. These early studies [48, 82, 83] solved two objective linear programming problems by solving the problem multiple times and varying weights on the objectives or transforming objectives into constraints. There are two issues with this classical approach. First, the approach requires simplifications in the system representation (i.e., the problems must be linear or convex). Non-linear relationships[14] cannot easily be accommodated in such an approach. Second, multiple computations are needed to achieve a full spread of solutions across the tradeoff set[15].

Multi-objective Evolutionary Algorithms (MOEAs) are a modern class of solution techniques that can provide high-quality approximations to the true Pareto optimal set for difficult problems. They use a full-complexity simulation model to calculate objective function values (see figure 2.4), allowing the analyst to solve problems that are non-linear, stochastic, discrete, and non-convex [54]. They also meet the second challenge outlined above; using a population-based approach allows MOEAs to develop approximations to the full set in a single algorithm run.

[13]The HWP was an important contribution to the design of water systems because it acknowledged that systems have different goals such as flood control, irrigation, and power generation. The authors advocated for multiple objective thinking, but they also used a single objective benefit function in some of their work. As quoted in Reuss [81], "the program assumed that the economy could be manipulated to maximize national welfare, but the reality is such *dirigste* manipulation is difficult in a democratic society". This quote suggests that there are additional objectives, not included in the economic benefit function, that should be considered in the analysis. The work presented in this dissertation, though, extends some of the ideas originally presented in the HWP, such as combining computer simulations and optimization approaches to better design water systems.

[14]For example, if the reservoir level is above a feet, restrict the flow to b from January 31 - May 1 and c otherwise.

[15]Recall that Pareto optimal solutions balance two conflicting objectives (see figure 2.1). It is difficult to find all solutions on the tradeoff set using the weighting approach, because many different combinations of weights are needed to ensure "full coverage" of the space and avoid gaps.

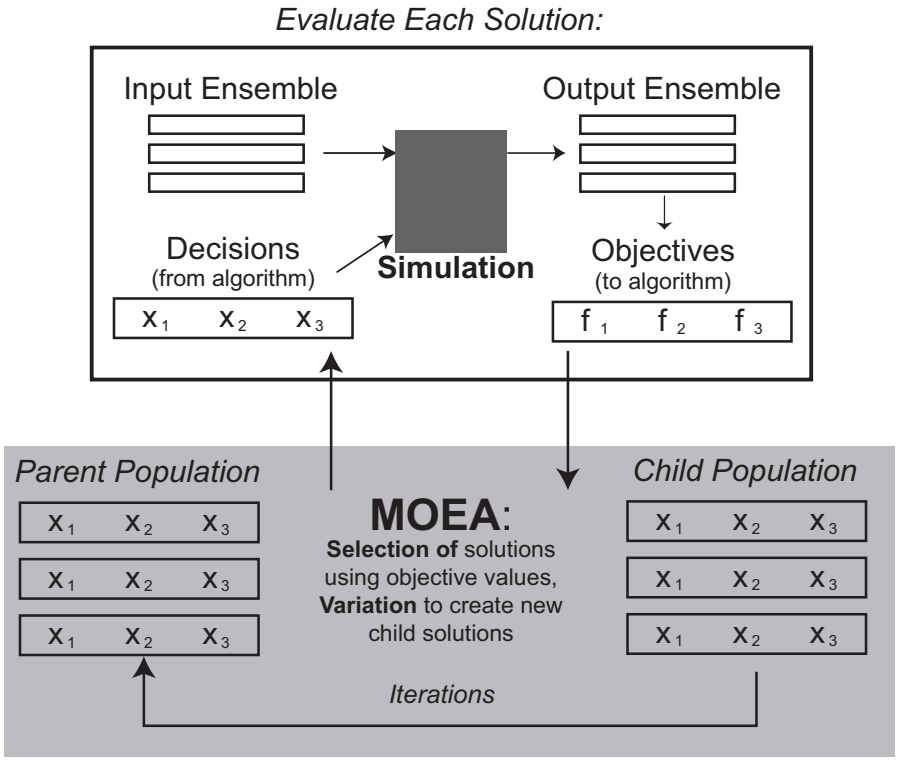

Figure 2.4. Illustration of a MOEA linked to a generic simulation model. A population of candidate solutions (system designs) is evaluated using a simulation model. The simulation model can accept an ensemble of input data, and output an ensemble of different performance metrics. The output ensemble can be used to calculate objective function values that are passed back to the algorithm. The MOEA uses an iterative process of selection of good solutions and variation on those solutions to evolve better populations.

MOEAs are part of a type of heuristic search algorithm called Evolutionary Algorithms (EAs). In general, EAs use an iterative process of selection of good candidate solutions and variation on those solutions to evolve solutions with better *fitness*[16]. A random population of solutions is first created each with different values for their decision variables. Each population member has their fitness evaluated. Selection maintains the solutions in the population with the best fitness.

[16]For a single-objective EA, the fitness function is usually equivalent to the performance of the single objective function. For solving multiple objective problems (MOEAs), fitness determines whether solutions are non-dominated with respect to each other. Solutions that have the best "rank" or best fitness function, are non-dominated on the first front. The solutions on the first front are removed, and non-domination is calculated with respect to the remaining solutions. A solutions' fitness is calculated as whether or not it is on the first front, and also whether it has many solutions in its vicinity [84].

Those solutions are allowed to "reproduce" in the next generation. Variation operators such as crossover, which switches information between two or more population members, and mutation, which locally perturbs one member, are performed on the solutions that survived the selection process. Each subsequent population's fitness improves iteratively. The process repeats until a user-defined stopping criterion. At the end of the process, the EA either finds the best approximation to the single optimal solution (in the case of single objective EAs) or an approximation to the set of Pareto optimal solutions (in the case of MOEAs).

For background information on MOEAs, please consult the following resources. Coello Coello [46] and Deb [49] review MOEAs and their application in many fields. Nicklow et al. [23] provides a comprehensive review of single and multi-objective evolutionary algorithms in the water resources domain. Several comparative studies are available, which show MOEA performance across an array of problems [54, 85, 86]. The remainder of this section introduces epsilon dominance, or user-defined solution precision (section 2.3.1.1) and the two MOEAs used in this dissertation (sections 2.3.1.2 and 2.3.1.3).

2.3.1.1 Epsilon-Dominance

Design of successful MOEAs requires that they can find non-dominated sets that are close to the true Pareto optimal set (convergence) and well-spread across the full range of values of the different objectives (diversity). Laumanns [87] shows that a process called ε-dominance archiving, where an ε grid is used to sort and maintain a set of promising solutions, can provide a proof of convergence and diversity[17]. In the concept of ε-dominance, non-dominance is not calculated using objective function values only, since doing so could possibly discover large numbers of solutions in a region not of interest to the decision maker. Instead, the ε-domination calculations consider a user-defined grid which represents the decision maker's significant precision on objective function values. Figure 2.5 illustrates

[17]The proof is summarized in Reed et al. [54]. There are two conditions required. The first is that the recombination operators have diagonal-positive transition matrices, achieved by using tournament selection with replacement and mutation. The second condition is elite preservation, which is guaranteed in epsilon archiving. The idea behind elite preservation is that non-dominated solutions are not lost as evolution progresses. If these two conditions are present, there is "a non-zero probability of generating a well-spread set of Pareto optimal solutions" [54].

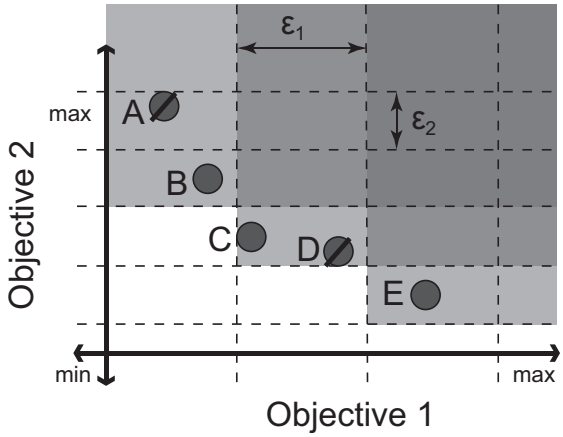

Figure 2.5. A two-objective minimization problem illustrating the concept of epsilon dominance. Since both objectives are minimized, preferred performance is toward the lower left corner of the figure. The user-specified epsilon resolution is shown with dashed grid lines. Within each grid cell only a single non-dominated solution is allowed to exist and all other solutions are eliminated (solution C dominates solution D). Overall epsilon non-dominance is assessed by performing sorting with respect to whole grid blocks (solution B block-dominates A). Note that solutions B, C, and E block-dominate the intersecting gray shaded areas.

this process with a two-objective minimization. First, the user specifies values for ε_1 and ε_2 that represent an objective resolution to be used in evaluations. Next, redundant solutions in each grid block are eliminated (i.e., solution D). Finally, epsilon nondominated sorting is performed with respect to entire grid blocks. Solution A is eliminated (or epsilon block-dominated) since the grid block containing solution B dominates the entire shaded area. The final epsilon non-dominated set is solutions B, C and E.

2.3.1.2 ε-NSGAII

Chapters 3 and 4 of the dissertation use a MOEA termed the epsilon Non-Dominated Sorting Genetic Algorithm, abbreviated ε-NSGAII [85, 88, 89]. The ε-NSGAII extends the original NSGA-II [84], a benchmark MOEA used in the prior literature, by adding epsilon-dominance archiving [87] (section 2.3.1.1) and adaptive population sizing [90].

Use of epsilon-dominance archiving in the ε-NSGAII creates "time continuation" [91], in that search can continue indefinitely due to new solutions being

presented within the search process. It does this by using a series of connected runs to evolve the approximation to the Pareto optimal set. The algorithm usually begins with a small number of candidate solutions. The iterative search progresses for a fixed number of generations, with each generation contributing solutions to an epsilon-dominance archive using a term in the MOEA literature known as elitism, or preservation of high-quality solutions. After each connected run is finished, the solutions within the archive are injected into a new, larger population comprised of 25% archive solutions and 75% randomly generated solutions. Thus the new population size is four times the archive size, with the high-quality solutions from the archive subsequently guiding further search. This adaptive population sizing is an effective way to change population size commensurate with problem difficulty and to avoid the difficulty of selecting an appropriate population size when choosing parameter settings for a MOEA.

The ε-NSGAII is a top performing MOEA, as evidenced in comparative studies [54, 85, 92]. One benefit of the ε-NSGAII is the above-mentioned flexibility of using epsilon dominance allows the algorithm to generate reasonably-sized Pareto approximations even for large, difficult problems. Kollat and Reed [88] used a long-term groundwater monitoring (LTM) test case with increasing problem dimension to explore how the ε-NSGAII's epsilon settings could be used to confront problem scaling, or the computational difficulty of solving problems of increasing size. The LTM problem tries to find the best sampling scheme that minimizes cost and measures of error in characterizing a contamination plume in the subsurface using a set of pre-installed monitoring wells. Binary decision variables are used (i.e., "1" if a well is sampled, "0" if it is not). The problem size is the total number of wells in the sampling field. As the problem size increases from 18 to 25 wells, a typical ε-NSGAII parameterization with fine-resolution epsilon settings has a quadratic growth in the number of function evaluations required to meet a target performance. Making the epsilon settings coarser for larger problem sizes, though, yielded a linear scaling performance, which drastically reduced the computational demand required for solving the problems. Furthermore, section 3.2.1 in this dissertation will contribute a novel use of epsilon-dominance for problems under uncertainty.

2.3.1.3 The Borg MOEA

Chapter 5 employs a massively parallel version of the recently introduced Borg MOEA [86, 93]. The Borg MOEA is used for two reasons, discussed as significant challenges in a recent review of the state of the art MOEAs [54]. The first is that new auto-adaptive algorithm technologies [93, 94] can help solve challenging problems by adapting their search properties without requiring input from the user. The infrastructure simulation in chapter 5 is non-linear, with "if-then" rules for operating a reservoir. Objective function calculations in the chapter are also subjected to noise (that is, the objective function calculation for a given solution is different each time the calculation is performed). The Borg MOEA's performance has been shown to be resilient to these types of mathematical challenges [54]. The second challenge is that emerging use of parallel computing can allow users to solve problems that would otherwise be intractable due to long function evaluation times. The function evaluation time for the new case study is several orders of magnitude longer than the LRGV simulation used in chapters 3 and 4, motivating the use of parallel computing to solve the problem in a reasonable amount of time.

The Borg MOEA uses a steady state [95], epsilon-dominance archiving (see section 2.3.1.1 and [87]), adaptive population sizing [85], time continuation [91] and adaptive operators [94]. Several recent applications have diagnosed its superior performance relative to other state-of-the-art MOEAs [54, 86, 96].

2.3.2 Sensitivity Analysis

Sensitivity analysis refers to methods that quantify how an output variable changes due to perturbations in input variables or parameters. A variety of sensitivity analysis approaches are available [97], ranging from local methods that vary a single parameter at a time to global methods that sample the entire parameter space and are more appropriate for complex non-linear models [98, 99]. The first method used in this dissertation is termed Sobol' variance decomposition [100], described in section 2.3.2.1. This method is used as a way to diagnose proper levels of complexity for decision variables. The second sensitivity approach is called the Patient Rule Induction Method (PRIM) [101]. PRIM seeks to find simple descriptions of parameter values that cause performance failures. PRIM is

introduced within the broader context of computational scenario discovery [78] in section 2.3.2.2.

2.3.2.1 Sobol' Sensitivity Analysis

Sobol' variance decomposition was chosen for the *de Novo* Planning analysis in chapter 3 due to prior work that has rigorously evaluated its effectiveness compared to other global sensitivity analysis approaches [98]. Use of Sobol' decomposition in chapter 3 focuses total-order sensitivity indices, which measure the impact of a variable X_i acting individually and as a result of all of its interactions with other variables. The total sensitivity index S_i^T is calculated using equation 2.4.

$$S_i^T = 1 - \frac{V[E[Y|X_{\sim i}]]}{V[Y]} \tag{2.4}$$

where $X_{\sim i}$ denotes the set of all variable inputs not including X_i. The upper term, $V[E[Y|X_{\sim i}]]$ represents the amount of variance in the output Y that would be reduced if one set all other variables in the analysis constant, allowing only X_i to vary over its range [102].

The computational technique for calculating the total-order sensitivity indices is fully described in [103] and demonstrated in [98]. Sobol' quasi-random sequence sampling [104], with a sample size denoted by q, is used to generate random parameter values as the input to the Sobol process. Our computations utilize the recommendations of [105] to calculate the first-order (individual effects), second-order, and total order indices using $q \times (2p+2)$ runs. The sample size was explored using values of $q = 2^k$, where k is an integer [104]. An additional diagnostic tool in our analysis was the assessment of the convergence and uncertainty of the Sobol' indices using the moment method for bootstrapping the sensitivity index estimates [98, 106] to attain the 95% confidence intervals on the magnitudes of the sensitivity indices.

2.3.2.2 Scenario Discovery and PRIM

Scenario Discovery (SD) identifies simple, easy-to-understand descriptions of uncertain model inputs that best predict when robust planning solutions perform

poorly [107]. Following Lempert et al. [47], the Uncertainties, Levers, Relationships, Measures (XLRM) paradigm is used to describe the notation for this process. A solution from optimization or another design procedure is defined by a set of decision variables or levers ("L") that define decision maker actions. Deep uncertainties ("X") are represented by a vector of values that quantify possible values for model parameters or descriptions of input data. The system is described by a quantitative simulation of the solution that maps relationships ("R") between levers and uncertainties. The result is a matrix of measures ("M"), otherwise called objective function values, that define the solution's performance.

Contrasting to Sobol' sensitivity analysis, the emphasis is not on finding *how* sensitive output metrics are to uncertain parameters or variables. Instead, the uncertain variables are sampled to see *which values* of the variables cause the outputs to be above or below a certain threshold. The goal of SD is to develop a short list of the key factors that will strongly shape the robustness of decision alternatives.

Figure 2.6 shows a hypothetical example of the SD process. In step 1, thresholds are set for each of the performance measures, in accordance with plausible stakeholder preferences. Uncertainty ensemble members that violate these thresholds are hereafter termed "vulnerable". In our example, the threshold is set at the 90th percentile of the distribution for measure m_1. In step 2, the points with vulnerable performance are colored in black, with the other points shown without a fill color. The horizontal and vertical axes in step 2 are values for uncertainties, x_1 and x_2.

As is shown in the figure, it can be difficult to describe the cluster of black points (i.e., the values that cause vulnerabilities) in simple terms. SD therefore employs the Patient Rule Induction Method (PRIM [101][18]) to automatically calculate "scenario boxes" such as the gray shading in the figure. Each scenario represents ranges of exogenous parameter values over which the candidate solution performs poorly. PRIM's boxes are similar to traditional scenarios because they provide simple descriptions of future trajectories. An important difference between our scenario boxes and traditional point-valued scenarios, though, is that

[18]Lempert et al. [108] compared PRIM to scenario discovery using Classification and Regression Tree (CART) analysis, favoring PRIM due to its high level of user interactivity. SD using PRIM is implemented in R, and is freely available [109].

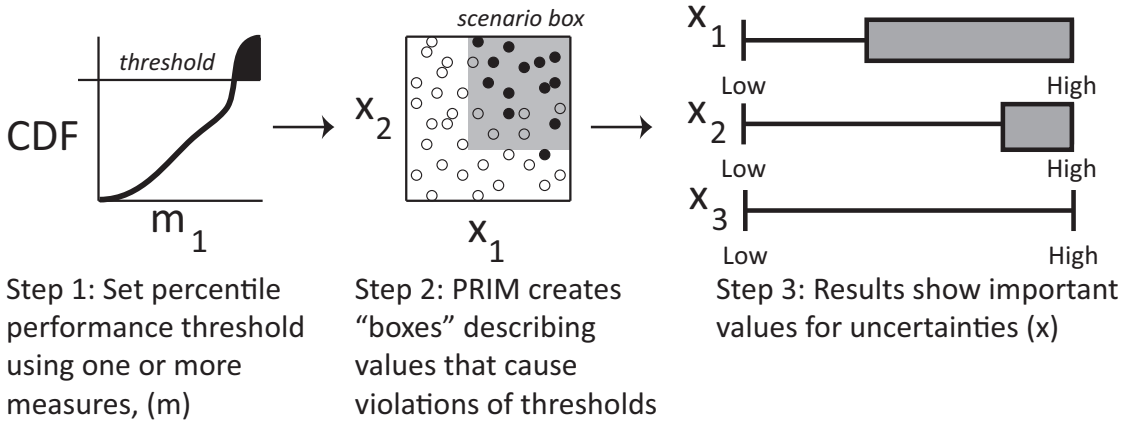

Figure 2.6. The three main steps of scenario discovery using the patient rule induction method (PRIM). Each step uses interactive visualizations in the R statistical package, with the user setting performance thresholds and levels of coverage and density for PRIM scenario boxes.

the boxes are developed using quantitative descriptions of system performance.

PRIM tries to find scenario boxes: simple descriptions of the values of uncertainties, x, that cause vulnerable performance. The box, B, comes in the form $B = \{a \leq x_j \leq b, j \in L\}$. In other words, a subset of dimensions of uncertainty x_j are constrained to be between lower and upper ranges a and b. There are two main iterative steps in finding the box B. PRIM "peels" away thin layers of the uncertainty space to change the box. It also increases or decreases the *number* of restricted dimensions, based on how well the box captures the set of scenarios of interest.

In an interactive process, PRIM suggests alternative candidate boxes from which the users choose. The interactive procedure allows the user to select boxes that balance three measures of box quality. *Coverage* quantifies how many of the vulnerable points are captured within a scenario. *Density* indicates how many of the captured points within a scenario are actually in the vulnerable set. *Interpretability*, which indicates how easily users can understand the information, is considered to decrease with the number of parameters used to define the box [78].

Note that the box B only constrains a subset of the total number of dimensions of uncertainty. In this manner, SD can suggest which uncertain parameters are less important in describing the vulnerable cases. Step 3 in Figure 2.6 is an example of how the constrained dimensions in the scenario box are visualized. In the example

only 2 dimensions out of 3 are constrained, indicating that the first two dimensions are the most important uncertainties for the system. Moreover, the grey shading shows the specific ranges of variables where vulnerabilities emerge.

In Figure 2.6, 18 points are captured in the scenario box. Of the points captured in the box, only 12 are vulnerable, but 13 points in total are vulnerable. Therefore, the coverage of our candidate box is $12/13 = 92\%$, whereas the density of the box is $12/18 = 67\%$. Boxes that have high density may "miss" many vulnerable points because the vulnerable points are spread through the uncertainty space. Conversely, scenarios with high coverage may not constrain the uncertainty dimensions very much, leading to low density and a poor amount of scenario precision. For a more detailed discussion please refer to Bryant et al. [78].

Bryant and Lempert [78] demonstrate SD using a quantitative model that estimates the benefits and costs of a policy requiring 25 percent of fuels to come from renewable sources by the year 2025. The vulnerability of the policy was defined using the measure, M, of high economic costs to taxpayers. Nine uncertain parameters, X, were explored, including the biofuel production cost, the biofuel yield, and demand elasticities for oil supply and transportation demand. Their results identified a scenario that constrained four of the nine parameters – high transportation demand elasticity and biomass backstop price coupled with low-cost biomass supply and a biofuel production cost larger than \$72.65/unit would lead to high costs of the policy. Once the scenario is discovered, the decision makers can then decide how likely this discovered scenario is, and use assumptions about that likelihood to determine planning priorities. As another example, Groves and Lempert [73] supported the California Department of Water Resources by identifying key uncertainties for the state's water management using SD. Their results distilled a set of 16 uncertain model inputs into three important policy-relevant dimensions: the population growth rate, the amount of naturally occurring conservation, and the cost of increasing water efficiency.

2.3.3 Interactive Visual Analytics

Many objective analysis provides a large number of alternatives from which to choose. It can be difficult to choose a proper alternative from the large number

of solutions that comprise the set. Therefore, many objective analysis uses an approach called interactive visual analytics [110–113]. Interactive visual analytics refers to interacting with multiple, linked views of data to gain a better understanding of a problem. The system combines the powers of human innovation and pattern recognition with quantitative tools (i.e., it creates a "joint cognitive system" [114, 115]). The interactive process allows the decision maker to see which objectives conflict with each other and what the impact of the decision levers is on system performance.

The approach used in this dissertation follows the method of Kollat and Reed [60] by visualizing tradeoff solutions both in the objective space (i.e., the quantitative values for performance metrics) and the decision space (the values that define a water supply portfolio or management regulations). The visualizations use scatter and glyph plots (figures 2.1 and 2.2) using interactive software that allows us to quickly change which variables are plotted. The software also allows "brushing" ranges of variables that are of interest. The analyst can then mark and highlight interesting solutions across different linked views. Parallel coordinates [116] are also used to visualize the objectives simultaneously and identify conflicts between objectives [51]. By selecting promising solutions and modifying the model assumptions and parameters, we also interrogate the effectiveness of our search on unmodeled objectives [117] and further inform problem modifications for future work. An additional use of interactive visual analytics comes from the analyst interacting with the PRIM algorithm for determining the coverage and density of discovered scenarios (see section 2.3.2.2).

2.4 Regional Problem Motivation

2.4.1 Lower Rio Grande Valley

The first case study in this dissertation focuses on water marketing [118] in the Lower Rio Grande Valley (LRGV) of Texas, USA [5, 119–121]. Due to limited regional groundwater storage, the primary sources of water in the LRGV are the Falcon and Amistad reservoirs, in which the water supply is shared between the United States and Mexico [119]. The reservoirs have an estimated combined stor-

age of 7.2 cubic km (5.8 million acre feet), with 2.6 cubic km (2.1 million acre feet) reserved for flood protection [122].

Our initial many objective exploration of the LRGV has shown that this type of analysis can improve the efficiency of urban water supplies and help maintain adaptability to risks posed by population growth and droughts [56]. In the LRGV, the presence of a water market does not imply that urban water planners can easily determine how to maintain high levels of reliability while also seeking to minimize a city's water surplus. This complex risk management problem requires flexible planning frameworks that can incorporate new knowledge (such as which innovative market instruments to use) and evaluate strategies rigorously for their complexity and effectiveness given uncertainty about long-term conditions within the LRGV.

The LRGV test case models a hypothetical city based on Brownsville, Texas, USA [122]. The modeled city has an average annual water use of 26 million cubic meters (21,000 acre feet) and participates in a water market in which water is transferred from the agricultural sector. Regional agricultural use contributes mostly to irrigation of low-valued crops such as cotton and accounts for 85 percent of regional water use. Rapidly increasing population demands motivate economic incentives for the irrigators to transfer water using the market from irrigation to municipal supply [122][19]. Water portfolio planning strategies are evaluated using both a ten year expected performance evaluation as well as a single year severe drought scenario, both using a monthly time step. This section provides a brief introduction to the quantitative simulation model used to simulate the city's use of the market, and the reader is encouraged to refer to [56] and [123] for more details.

Three types of water supply instruments are considered in our case study. The first is permanent rights, where users are allocated a percentage of reservoir inflows every month[20]. The second instrument is the options contract [124], in which users pay a small up-front fee at the beginning of each year for the right to exercise all

[19]The volume of water needed for the city's municipal supply is relatively small compared to the other demands from the reservoir. The assumption is that water will always be available from the irrigators to municipal supply.

[20]Following prior work [122], our test case assumes that permanent rights are fulfilled as a percentage of reservoir inflows, so the amount allocated to rights is variable in each month.

or some of the options contract at a fixed price. Options contracts for the case study are similar to a European call stock option, where the contract can only be exercised at a set month (June) every year. The third instrument is spot leasing [5], in which the user can purchase a variable amount of water in any month at the market price. Note that the options contract provides some reduction in cost volatility compared to the spot leasing market, but the restriction of exercising only at one time reduces the options' flexibility.

A Monte Carlo simulation model samples historical hydrology, water pricing, and projections for the region's rapidly growing population demands and is used to evaluate the performance of the city's water supply portfolio [56, 122, 123]. The portfolio is comprised of anticipatory risk-based water purchasing rules that control how the city acquires spot leases and exercises its options contracts. The rules are formulated to trigger alternative strategies for purchasing water transfers given uncertain forecasts of the city's supply and demands. Chapters 3 and 4 formulate simple risk-based rules that guide the city's use of the LRGV water market, while providing decisions that are robust to potentially severe droughts.

The city's water supply portfolio is represented using a suite of decision variables for each of the supply instruments. The city's permanent rights are specified by a volume, N_R. In each month, water is allocated to the permanent rights on a *pro rata* basis using the ratio of N_R to the volume of total regional water rights. If the city had 10 percent of the regional rights, therefore, it would be allocated 10 percent of the available inflows in each month. The rights have an annualized price, p_R, set to \$1.82 per 100 cubic meters (\$22.50 per acre-foot) [123]. The city's options contract dictates a volume of water N_O that represents the maximum amount that can be exercised in the exercise month. In this study, a single exercise month is used, with decisions made at the end of May (with water available for use starting in June). This volume of water also influences the up-front cost for holding the contract, set to the volume of options multiplied by the fixed options price p_O of \$0.43 per 100 cubic meters (\$5.30 per acre foot). An adaptive options contract introduced in recent work [56, 123] provides more flexibility to the city's use of the options contract. The adaptive contract allows the city to reduce its contract volume when it anticipates that it has a sufficient volume of water in its supply account. In each simulation year, the city chooses between a high ($N_{O_{\text{high}}}$)

and low-volume ($N_{O_{\text{low}}}$) options alternative based on the ratio of its current supply at the beginning of the year (N_{r_o}) to its permanent rights. An additional decision variable, ξ, sets this threshold (see equation 2.5).

$$N_O = \begin{cases} N_{O_{\text{low}}} & \text{if } \frac{N_{r_o}}{N_R} \geq \xi \\ N_{O_{\text{high}}} & \text{if } \frac{N_{r_o}}{N_R} < \xi \end{cases} \tag{2.5}$$

The third instrument, leases, is denoted by N_l. The city is charged a price per cubic foot for acquired leases, $\hat{p}_l$, drawn randomly from the lease price distribution. Following [122], there are two distributions of historical lease pricing based on the reservoir level, with higher prices corresponding to a lower reservoir level.

Each model run for either the ten-year scenario or the drought begins with a volume of water expressed as a percentage of the city's permanent rights, such that the initial volume of water N_{r_o} relates to the rights according to equation 2.6:

$$N_{r_o} = i_{fr} N_R \tag{2.6}$$

where i_{fr} is fraction of the initial rights given to the city to begin the simulation. Additionally, regional demand growth is modeled in the study as exponential growth with a compounded percentage change, dm, in each year. Both the initial fraction i_{fr} and the demand growth dm are explored in the thesis to diagnose their effect on the water supply portfolios' performance.

The volume of water acquired from the options contract and leases is determined by two types of anticipatory thresholds. In time period k, the α_k variable determines "when" the city purchases water, by comparing the ratio of expected supply to expected demand and comparing it to the alpha value. If the amount of water is less than the threshold, the city must purchase water on the market. The actual volume of water purchased is dictated by β_k. For example, if α_k is 1.5 and β_k is 2.0, the city will purchase water if their expected supply to demand ratio is lower than 1.5. When the city purchases this water on the market, it purchases a sufficient volume such that their supply is 2.0 times their expected demand forecast. In all months except the options exercise month, these market purchases are made using spot leases. In the options exercise month, the city first meets its required market acquisition volume by exercising their options at the fixed price

p_X ($1.22 per 100 cubic meters) if this price is lower than that month's sampled lease price. Conversely, if the leases have a lower price, leases are used instead of options. Additionally, the city can augment its options purchase in the exercise month with extra leases if the volume of their options contract is not large enough to meet the amount of water that the β_k value requires.

Two sets of α and β strategy variables were used in the first many objective LRGV study (covered in this chapter, and presented fully in Kasprzyk et al. [56]), and in other prior work [122,123]. The first set covers from from January to April, and the second from May to December (including the options exercise month)[21]. The goal of the *de Novo* analysis in Chapter 3 will be to ascertain the appropriate level of complexity for these strategy variables while still maintaining efficient and cost-effective market use. This result will be further explored in the robustness analysis of Chapter 4.

2.4.2 Lower Susquehanna

Chapter 5 develops a test case in the Lower Susquehanna portion of the Susquehanna River Basin (SRB, figure 2.7). The SRB spans 27,000 square miles across large portions of New York, Pennsylvania, and Maryland. Its water resources have a high economic value, providing approximately 51 percent of the freshwater input to the Chesapeake Bay [125]. The Susquehanna River Basin Commission (SRBC) is tasked with managing the SRB's vast water resources. Our Lower Susquehanna test case focuses on alternative demand management problem formulations that are similar in nature to how the SRBC manages the system in low flow conditions [126]. The main source of supply is the Conowingo pond in Maryland [43], near the Pennsylvania border (see figure 5.1 in chapter 5). The man-made pond was created when a dam across the Susquehanna River was completed in 1928, creating a 9,000 acre pond with a maximum depth of 98 feet and 35 miles of shoreline. Although the dam was primarily built for hydroelectric generation, the pond is also used for municipal water supply, industrial cooling water, and fishing and

[21] In each period k, β_k was always constrained to be greater than or equal to α_k. Note that the use of the market could be controlled by fewer variables (i.e., one threshold for both "when" and "how much", regardless of time of year), or more variables (as many as one unique value per month, with distinct α_k and β_k).

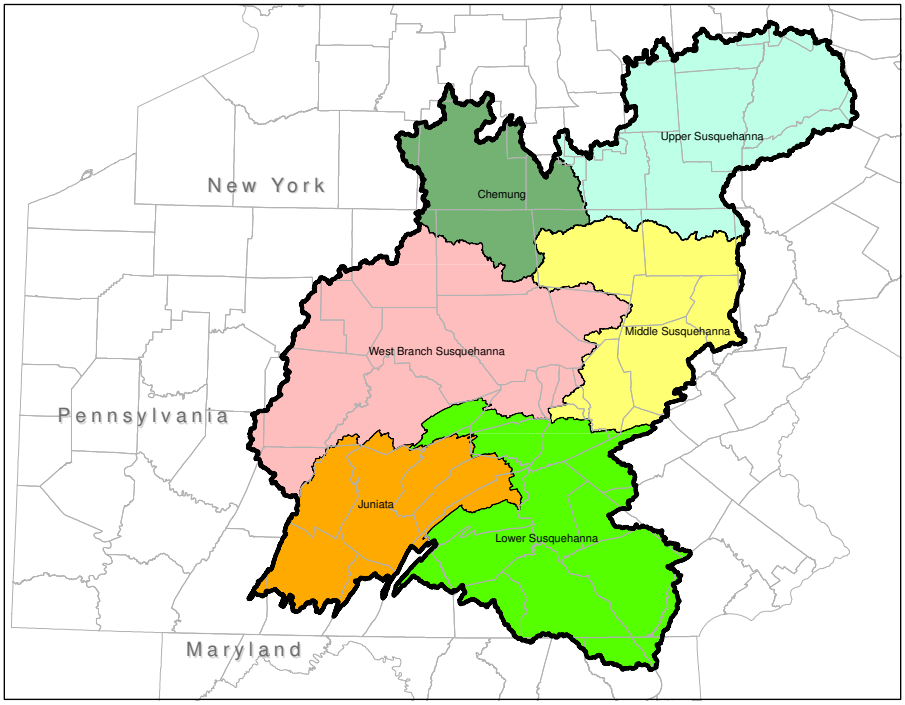

Figure 2.7. The Susquehanna River Basin, with sub-basins shown.

other recreation. To manage operations of the Conowingo and the larger basin upstream, the SRBC uses the OASIS management model [127–129]. OASIS is a general purpose water resources model that uses a linear program (LP) solver to allocate water to meet multi-sector demands. The model solves a new LP at each time step, using a series of goals and constraints to optimally allocate water and meet continuity restrictions. Therefore OASIS does not use operational "rule-curve" based water operations (e.g., releasing a certain quantity of water based on the season) and instead shows how the system would respond if all sites are perfectly coordinated across the system. The OASIS model is used operationally by the SRBC as means of computer-aided decision support across the multi-sector stakeholders impacted by regulation decisions. The motivation for our case study stems from an existing SRBC report detailing management of the pond [43]. In the report, the SRBC modified water demands, downstream flow requirements,

and storage releases in order to manage Conowingo's performance in low flow conditions [43]. Recommendations from the report were used in the real system to improve its management.

The use of OASIS as a means to guide adaptive planning within the SRB represents a major contribution to the field of water resources planning and management. However, the system's deterministic use of historical hydrology is a strong limitation given the potential for significant changes in the basin's hydrology and human demands [1, 125, 130–134]. All of the decisions about management alternatives in OASIS are based on the historical record of inflow and evaporation within the SRB. The SRBC is concerned about the increasing incidence of droughts in the basin, motivating a stochastic approach that can generate new plausible sequences to test assumptions about future droughts. A second challenge that emerges with the existing OASIS based planning is the requirement of trial-and-error exploration of alternatives. Although the GUI allows users to manually modify operating procedures, there is no mechanism for discovering optimal tradeoffs between the sectoral concerns. Moreover, the deterministic use of the historical record has the potential to underestimate drought risks and the consequent system conflicts. Chapter 5 seeks to more comprehensively address the 2009 Pennsylvania Department of Environmental Protection's [135, 136] call for improved "assessment of practical alternatives" for water supply.

Chapter 3

Many Objective de Novo Planning Under Deep Uncertainty

This chapter proposes a *de Novo* planning framework and demonstrates it on the Lower Rio Grande Valley (LRGV) risk-based water management problem (section 2.4.1), in which a city seeks the most efficient use of a water market. The framework couples global sensitivity analysis using Sobol variance decomposition with multi-objective evolutionary algorithms (MOEAs) to generate planning alternatives and test their robustness to new modeling assumptions and scenarios. The LRGV case study uses a suite of 6-objective problem formulations that have increasing decision complexity for both a 10-year planning horizon and an extreme single-year drought scenario. The *de Novo* planning framework demonstrated illustrates how to adaptively improve the value and robustness of our problem formulations by evolving our definition of optimality while discovering key tradeoffs. This chapter was adapted from a study published in *Environmental Modelling and Software* [32].

3.1 Introduction

Climate change, population growth, and increased urbanization pose serious challenges to urban water supply management [1, 137–140]. These changes lead to increased water demands and amplified hydrologic variability, subsequently leading to higher risks for water supply failures [2]. The severe costs associated with structural adaptations such as building new reservoirs motivate the need for in-

novative nonstructural adaptation techniques such as water marketing [118]. A water market seeks to allocate water to its "highest value use" through transfers between regions [141] or different user sectors [142]. The presence of a water market, though, does not imply that the city knows how to develop the most efficient water supply portfolio that fulfills all of its planning goals [56]. In fact, there is a significant amount of deep uncertainty (see section 2.2.1) in choosing an appropriate problem formulation for water portfolio planning (i.e. the decision variables, objectives and constraints used). Specifically, this chapter rigorously tests alternative formulations of a city's decision strategies and carefully explores the effect of modeling assumptions by constructing challenging planning scenarios. The goal of this analysis is to develop robust solutions that have good performance under many different modeling conditions [9] and aid decision makers in understanding the effects of their use of estimated probabilities [143] on the planning process.

The framework is presented in figure 3.1. Step 1 begins with an *a priori* problem formulation that represents planners' initial conception of the problem through a quantitative model, decision variables that control strategies within the model, and objectives and constraints that measure strategies' performance. In step 2, the framework diagnoses the effect of decision variables and model parameters using Sobol' variance decomposition [100]. The illustration in figure 3.1 shows that different variables can have a wide range of sensitivity performance across different evaluative metrics. Step 3 uses insights from the sensitivity analysis to construct a new many objective planning problem. Objectives and constraints can be removed or added depending on their sensitivity structure or insights learned from previous iterations of the framework. Additionally, a suite of decision variable formulations of increasing complexity are used to explore the implications of the sensitivity analysis results. This framework seeks a balance between the complexity and effectiveness of a planning formulation.

Step 4 solves the *de Novo* formulations using a MOEA. After a quantitative tradeoff comparing performance across decision variable formulations is developed, step 5 uses interactive visual analytics [110–113] to view the tradeoffs interactively when evaluating the competing decision variable formulations. Exploration of decision variables' impact on many objective tradeoffs has been successfully demonstrated in prior work [144]. Use of interactive visual analytics represents *a posteori*

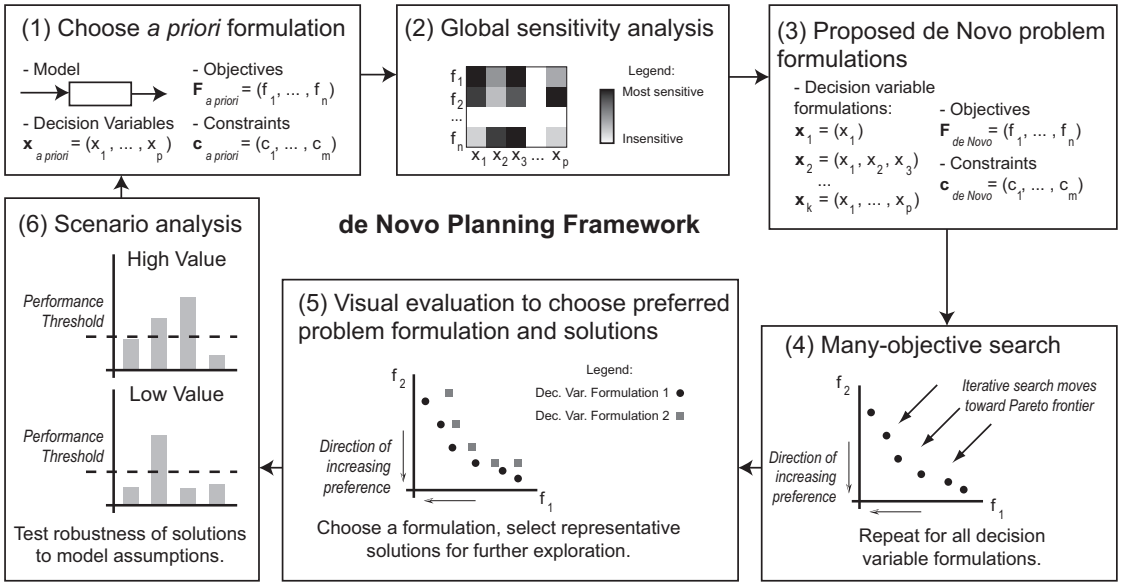

Figure 3.1. Illustrative example of the de Novo planning framework.

decision making, where decision makers explore our approximate Pareto optimal sets to negotiate a choice of alternative as a final decision [60, 145–147]. A major benefit to this approach is that it allows the decision makers to modify their preferences and perform experiments through setting thresholds on objective function values and adding unmodeled objectives [117] to their analysis. Within step 5, the planners can choose the decision variable formulation that provides preferred performance compared to the other formulations. In this manner, formulations themselves can be considered non-dominated with respect to each other if they provide non-dominated solutions of interest to the decision maker and/or increase the diversity of alternatives that can be considered [15]. This focus on finding the non-dominated problem formulation (as compared to the classical focus on non-dominated solutions within a single static formulation) is a unique contribution of this work. Selected solutions within this preferred formulation are also used to further interrogate the effect of deeply uncertain model assumptions on the solutions' performance. Step 6 shows how deviations in model assumptions can change the performance of the selected solutions. Note that this process is iterative, and further improvements can be made to the problem formulation after this scenario analysis (i.e., the formulation from step 6 becomes a new *a priori* formulation for the next investigation). Overall, the *de Novo* planning framework seeks to

facilitate learning and innovation in decision making problems solved under deep uncertainty.

3.2 Methods

Several techniques used in this chapter have been introduced in previous sections. The chapter uses MOEAs (section 2.3.1), Sobol' global sensitivity analysis (section 2.3.2.1), interactive visual analytics (section 2.3.3) and the Lower Rio Grande Valley (LRGV) case study (section 2.4.1). The remainder of this section proceeds as follows. First, we introduce a new technique for handling uncertainty in MOEAs (section 3.2.1). Next, we fully define the planning objectives used here (section 3.2.2). Our de Novo Planning demonstration utilizes an *a priori* problem formulation introduced in section 3.2.3, and the drought scenario used to test solutions is discussed in section 3.2.4.

3.2.1 Handling Uncertainty in MOEAs

The illustrative example in chapter 2 showed how MOEAs are effective at developing many objective tradeoffs to complicated problems. Uncertainty in objective function calculations provides an additional challenge to using MOEAs. The major sources of uncertainty in the objective function calculations for the LRGV case study result from sampling distributions of hydrologic conditions, water demands, and lease pricing. The uncertainty is problematic since the nondomination ranking procedure may preserve solutions in which a different ensemble of sampled data in the Monte Carlo simulation would have caused them to be dominated by another solution. We address this issue by using a novel application of epsilon dominance; the epsilon resolution is set to minimize the likelihood that solutions are mistakenly added to the epsilon nondominance archive due to uncertainty or noise in their mean-based objective rankings. Figure 3.2 illustrates the approach on a two objective problem. An epsilon non-dominated approximation to the Pareto set is illustrated with four solutions labeled A-D. The solutions' error bars indicate the noise or range of each solution's mean values for objectives 1 and 2, and the circle illustrates an example instance of the mean. Epsilon values control the ranking

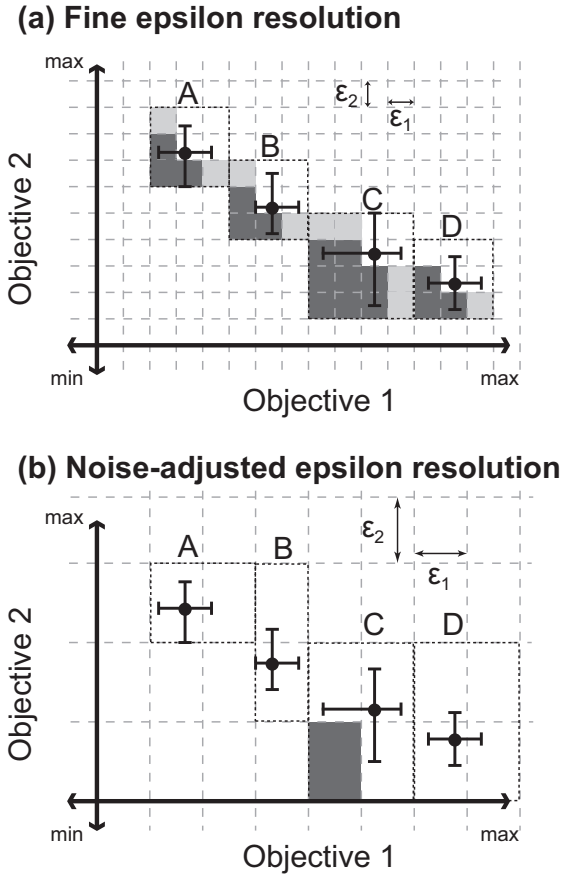

Figure 3.2. Consequences of epsilon resolution settings for an example two-objective minimization under uncertainty. The Pareto set consists of 4 solutions in each panel with the circle indicating mean objective value performance and the error bars showing the range of objective function values. Panel A has a fine epsilon resolution while panel B's resolution is more coarse. The black dotted line shows the range of epsilon blocks into which a draw of each solution could fall. Dark gray blocks indicate that a draw in this block would dominate the mean objective value, lighter gray blocks indicate that a draw in this block would be non-dominated with respect to the mean value. Draws in the gray blocks would result in copies of the same decision vector maintained in the Pareto set with different objective values. However, the reduced resolution of panel B reduces the likelihood of those copies being made.

calculation and determine the solutions that will remain in the set; the values are set to a fine resolution in figure 3.2a and adjusted for noise in figure 3.2b. The dotted lines indicate the range of epsilon blocks into which draws of each solution would fall.

Within the MOEA's search process, multiple copies of decision variable vectors are made and evaluated with a new ensemble sample of input data. Copies of

solutions that fall in the light grey blocks would be falsely classified as being epsilon non-dominated with respect to the solution's uncertain mean value and would remain in the Pareto approximation set. Copies that fall in the darker grey boxes would dominate the mean and replace it in the set, and copies that fall elsewhere would be dominated by the existing instance of the mean. Using the fine resolution in figure 3.2a yields a high likelihood that copies of already existing solutions (i.e. with identical values of the decision variables) may erroneously survive in the reference set. In figure 3.2b, we size the epsilon blocks to be equal to the largest sampled range or uncertainty in the means of any solution on a given objective. This reduces the chance that a solution's copies will be replicated in the Pareto approximation set, because draws of the solutions are more likely to be made within the same epsilon block and would be eliminated within the sorting procedure. In figure 3.2b, only solution C has the possibility of its copy dominating the mean, and situations such as this would become less frequent as the number of dimensions increases. The approach summarized in figure 3.2 increases the competition among solutions to stay epsilon non-dominated within each large block. In a many objective problem as in [56], it is very unlikely that the noise would bias a solution to be falsely non-dominated in all objectives. Robust solutions will have to repeatedly beat other candidates by stably controlling their epsilon box (i.e., the means for all objectives are non-dominated).

The approach presented in figure 3.2 is advantageous because it requires no modification of the algorithm itself and it is valid within high-dimensional problems with no closed-form analytical representations of the probability distributions of objectives. The approach is similar to the ranking methodology of [148] but requires no *a priori* assumption of the shape of the uncertainty distribution. To obtain the noise-adjusted epsilon values for search, a representative set of solutions is evaluated for several independent identically distributed random draws to develop an uncertainty estimate for each objective's Monte Carlo based estimate of its mean. The epsilon blocks are then sized to be equal to the largest sampled range of uncertainty across the set of selected solutions. Sampling different types of solutions is helpful in determining if they exhibit varied probabilistic behavior depending on the decisions. While there is some computational cost in performing the initial Monte Carlo sampling necessary to create these settings, the result-

Table 3.1. Planning Metrics

Name	Symbol	Description
	Efficiency	
Cost	f_{cost}	Cost of rights, options, leases
Surplus Water	f_{surplus}	Water held at end of year
Cost Variability	f_{costvar}	High tail of cost distribution
Dropped Transfers	f_{dropped}	Volume of expired transfers
	Risk Indicators	
Reliability	f_{rel}	Probability of successfully meeting demands
Critical Reliability	$f_{\text{crit. rel}}$	Probability of avoiding critical failures (supply less than 60% of demand)
Resilience	f_{resil}	How quickly the system recovers after a failure
Vulnerability	f_{vuln}	Volume of most severe failure
	Market Use	
Number of Leases	$f_{\text{num. leases}}$	Number of leases regardless of volume
Volume of Leases	$f_{\text{vol. leases}}$	Volume of purchased leases
Volume of Exercised Options	$f_{\text{vol. ex. options}}$	Volume of exercised options
Cost Percentage Leases	$f_{\text{cost per. leases}}$	Contribution of leases to total cost
Cost Percentage Options	$f_{\text{cost per. options}}$	Contribution of options to total cost

ing search problem that uses the noise-adjusted epsilon values yields a smaller Pareto approximation set size than with the fine resolution epsilon settings, maintains diversity throughout the objective space, and reduces overall computational time [88]. The noise-adjusted epsilon values increase the robustness of solutions, since the uncertain solutions face increased competition to survive in successive generations in larger grid blocks [149].

3.2.2 Performance Metrics

Each performance metric considered in this study uses the quantitative Monte Carlo simulation described in section 2.4.1. This section will describe these metrics

in detail, representing three categories as presented in table 3.1. *Efficiency* metrics evaluate costs and the volumes of water carried over or not used for supply, *risk indicators* focus on water portfolios' modes of failure and recovery, and *market use* metrics quantify the extent to which portfolios rely on the water market to provide supply.

The following notational definitions refer to our Monte Carlo simulation ensemble, with the variable M representing the number of Monte Carlo samples taken. The simulation is run for T planning years where $T = 10$ for our long-term analysis and $T = 1$ in the drought. The index i is used below to indicate the planning year, and the index j denotes the month within that year. The notation $E[]_i$ denotes an expectation over the M Monte Carlo samples of the ith year. The notation $\mathbf{x}_k$ represents the vector of decision variables that describes the city's water supply portfolio. The subscript k denotes the form of the decision variable vector used.

3.2.2.1 Efficiency Metrics

COST The cost of each portfolio comprises costs from permanent rights, an up-front fee for the options contract, the exercised options, and purchased leases. The cost metric (equation 3.2) is a sum of the annual costs (equation 3.1) calculated as follows,

$$\begin{aligned} f_{\text{annual cost}}(\mathbf{x}_k)_i &= \mathrm{E}\Bigg[N_R p_R + N_{O_i} p_O \\ &\quad + N_{x_i} p_x + \sum_{j=1}^{12} \left(N_{l_{i,j}} \hat{p}_{l_{i,j}} \right) \Bigg]_i \qquad (3.1) \\ f_{\text{cost}}(\mathbf{x}_k) &= \sum_{i=1}^{T} f_{\text{annual cost}_i} \qquad (3.2) \end{aligned}$$

with terms as defined previously. The permanent rights volume is constant for all T simulation years. The decision to exercise options is made once per year, denoted by the subscript i. If the adaptive options contract is used, the volume of options available (N_{O_i}) is determined by the city's initial water supply in each year according to equation 2.5. Leases incur cost when purchased in each of the 12 months of the T simulation years, with each lease purchase subject to a uniquely

sampled price, $\hat{p}_{i,j}$, in the jth month of the ith year.

SURPLUS WATER The surplus water metric quantifies the water held by the city at the end of each simulation year. This volume of water, which includes volumes of water from permanent rights, options, and leases, has been minimized in previous work [56] to free water for other uses (such as ecological flows). Formally, the metric calculates an average of the annual expected surplus water volumes:

$$f_{\text{surplus}}(\mathbf{x}_k) = \sum_{i=1}^{T} \frac{1}{T} \Big(\mathrm{E}[S_j]_i \Big), \quad j = 12 \tag{3.3}$$

where the variable S_j denotes the city's water supply (comprising rights, options, and leases) in month j.

COST VARIABILITY Variance in the distribution of costs for each portfolio is introduced by the anticipatory rules for exercising options and purchasing leases, since each draw in the Monte Carlo ensemble has different volumes of options and leases acquired. The Contingent Value of Risk (CVAR) captures this variability, defined as the mean of costs falling above the 95th percentile [123]. Following our previous work [56], the metric captures the year with the highest cost variability, to ensure that the rest of the simulation years have a lower amount of variability (see equation 3.4),

$$f_{\text{costvar}}(\mathbf{x}_k) = \frac{\max_{i \in [1,T]} \mathrm{CVAR}_i}{f_{\text{annual cost}_i}} \tag{3.4}$$

where the index i denotes the year with the highest CVAR cost, and $f_{\text{annual cost}_i}$ is expected annual cost in that year.

DROPPED TRANSFERS The dropped transfers metric stems from the fact that leases and exercised options in this study expire after a year of non-use. The variable a describes the "age" of the water in the city's supply account, so that when $a > 12$, the water has not been used for 12 months after its acquisition and therefore expires (it is no longer considered available to meet demand). This metric is important since water managers would prefer portfolios that avoid transfers that result in large volumes of water being dropped. The dropped transfers objective is computed as the sum of the annual expected volume of dropped transfers (see

equation 3.5).

$$f_{\text{dropped}}(\mathbf{x}_k) = \sum_{i=1}^{T} \left(\mathrm{E}\left[\{N_{x_i} : a > 12\} + \sum_{j=1}^{12} \{N_{l_{i,j}} : a > 12\} \right] \right) \tag{3.5}$$

Note that the metric measures two components: a volume of water from exercised options (one value in the ith year, N_{x_i}) and leased water (acquired in the jth month of year i, $N_{l_{i,j}}$). In calculating the metric, both entire lease acquisitions and portions of lease acquisitions that are unused are considered. For example, if the city purchases 1,000 cubic meters and only 300 cubic meters are used, 700 cubic meters are said to be dropped.

3.2.2.2 Risk Indicator Metrics

Drawing from [40], the concepts of resilience, reliability, and vulnerability are used to quantify the risk-based performance of portfolio planning strategies.

RELIABILITY The reliability of a portfolio captures the probability of successfully supplying the city's water demands (i.e. how often the city avoids failure). A portfolio's reliability $r(\mathbf{x})_i$ is initially defined following the formulation of [122].

$$r(\mathbf{x}_k)_i = 1 - \frac{E[n_{\text{fail}}]_i}{12} \tag{3.6}$$

where $E[n_{\text{fail}}]_i$ represents the expected number of monthly failures in the year i. These failure events are defined as the city's supply (S_j) falling strictly short of the simulated demand (d_j) regardless of the shortfall volume, according to equation 3.7.

$$S_j < d_j \tag{3.7}$$

The aggregate reliability metric f_{rel} calculates the lowest expected reliability of any year in the planning period.

$$f_{\text{rel}}(\mathbf{x}_k) = \min_{i \in [1,T]} (r_i) \tag{3.8}$$

When used in the optimization, f_{rel} is maximized to ensure that each planning year has performance at least as high as this lower bound.

CRITICAL RELIABILITY Critical reliability, $f_{\text{crit. rel}}$, is calculated in the same manner as reliability (see equations 3.6 and 3.8), but the definition of failure is set to the city not being able to meet at least 60% of its simulated demand [122], given in equation 3.9.

$$S_j < 0.6 d_j \tag{3.9}$$

RESILIENCE Resilience measures "... how quickly [a water resource system] returns to a satisfactory state once a failure has occurred" [40]. Following the notation of [150], the set of "satisfactory" states represents the condition where the monthly supply is strictly greater than the simulated demand (see equation 3.7). An "unsatisfactory" state consequently represents a failure in that month. In this study, resilience is function of a portfolio's performance during the whole simulation regardless of simulation year. The index t is used such that the simulation begins at $t = 1$ and ends at $t = 12T$, where T is the number of years in the simulation. This convention accounts for failure periods that span more than one calendar year, from December of one simulation year through January of the next simulation year, for example. Let Z_t equal 1 if there is no failure in a month t. This relationship is expressed using the notation for the supply in month t (S_t) and the demand in month t (d_t) as follows:

$$\begin{aligned} \text{if} \quad S_t > d_t \quad \text{then} \quad & Z_t = 1 \\ \text{else} \quad & Z_t = 0 \end{aligned} \tag{3.10}$$

Since resilience considers transitions between satisfactory and unsatisfactory states, the variable W_t is introduced to indicate a transition into failure in timestep $t+1$ after an observed failure at time step t (equation 3.11).

$$W_t = \begin{cases} 1 & \text{if } Z_t = 1 \text{ and } Z_{t+1} = 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.11}$$

Using the above definitions, equation 3.12 defines resilience for the mth Monte

Carlo realization, denoted by rs_m.

$$rs_m = \begin{cases} 1.0 & \text{if } Z_t = 0 \; \forall \; t \in [1, T] \\ \frac{\sum_{t=1}^{12T} W_t}{12T - \sum_{t=1}^{12T} Z_t} & \text{otherwise} \end{cases} \tag{3.12}$$

In equation 3.12, the resilience is 1.0 if there are no failures. In the presence of failures, resilience is the ratio of the number of transitions into a failure state to the number of failure time steps. If there is the same number of failures as the number of transitions into the unsatisfactory state, this means that the system quickly "rebounds" from a failure and does not spend more than one step in the unsatisfactory state. Across all M Monte Carlo realizations, the resilience metric is defined using an expected value in equation 3.13:

$$f_{\text{resil}}(\mathbf{x}_k) = \sum_{m=1}^{M} \frac{1}{M} rs_m \tag{3.13}$$

VULNERABILITY Vulnerability measures the most severe failure period, defined as a set of failures with the largest differential between the volume of supply, S_t, relative to the demand, d_t. Using the notation of [150], let the number of failure periods of one or more timestep be denoted by G. Also, let J_g be the set of timesteps representing the gth transition from the satisfactory state, to the unsatisfactory state, and back. For example, if the 3rd and 4th months represented the first time the system went into a failure, then $J_1 = \{3, 4\}$. If the 5th month was satisfactory, the 6th month contained a failure, and the 7th month was satisfactory, then $J_2 = \{6\}$. Equation vn_m denotes a measure of vulnerability for the mth realization (i.e. a single time series similar to equation 3.12).

$$vn_m = \max_{g \in [1,G]} \Big(\sum_{t \in J_g} (d_t - S_t) \Big) \tag{3.14}$$

Each period of failures is examined in the time series, with the period with the most severe failure recorded as a volume, vn_m. Similar to the resilience metric shown above, the vulnerability metric calculates an expectation across all M Monte Carlo

realizations (equation 3.15).

$$f_{\text{vuln}}(\mathbf{x}_k)i = \sum_{m=1}^{M} \frac{1}{M} vn_m \tag{3.15}$$

3.2.2.3 Market Use Metrics

NUMBER OF LEASES Since leases can be purchased in any month, some portfolios may specify a large number of leases to be purchased, which may imply a large transactions cost for the city. The number of leases metric attempts to quantify this by explicitly counting the number of leases regardless of volume, according to equations 3.17 and 3.16 below:

$$\phi_{i,j} = \begin{cases} 1 & \text{if } N_{l_{i,j}} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.16}$$

$$f_{\text{num. leases}}(\mathbf{x}_k) = \sum_{i=1}^{T} \left(\mathrm{E}\Big[\sum_{j=1}^{12} \phi_{i,j} \Big] \right) \tag{3.17}$$

where ϕ accounts for whether or not a lease was acquired, regardless of its volume.

VOLUME OF LEASES The volume of leases metric calculates the average annual expected volume of leases for a given portfolio. This metric was included in the analysis since the anticipatory risk-based rules that govern options and leases do not provide a deterministic amount of market use but rather provide a strategy for the city to use the market based on the amount of supply and demand that is forecasted in future decision periods. Equation 3.18 calculates a sum of lease volumes across the 12 months of the ith year for each Monte Carlo simulation, and take the expected value across all Monte Carlo simulations. Using the T values for the expected annual volume of leases, a planning period average is calculated from each of these annual values.

$$f_{\text{vol. leases}}(\mathbf{x}_k) = \sum_{i=1}^{T} \frac{1}{T} E[\sum_{j=1}^{12} (N_{l_{i,j}})]_i \tag{3.18}$$

VOLUME OF EXERCISED OPTIONS Similar to the volume of leases metric, this metric provides a calculation of the average volume of exercised options for a

given portfolio. Equation 3.19 first uses the Monte Carlo simulation to calculate an expected value of the exercised options (N_{x_i}) in year i. An average volume of exercised options is then calculated by averaging these T annual values.

$$f_{\text{vol. ex. options}}(\mathbf{x}_k)_i = \sum_{i=1}^{T} \frac{1}{T} E[N_{x_i}]_i \tag{3.19}$$

COST PERCENTAGE LEASES The following cost percentage metrics seek to quantify the contribution of market use to the total cost. To calculate the cost percentages of leases metric, the portion of the total cost that was due to lease acquisitions is divided by the total cost.

$$f_{\text{cost per. leases}}(\mathbf{x}_k) = \frac{\sum_{i=1}^{T} \mathrm{E}\left[\sum_{j=1}^{12} \left(N_{l_{i,j}} \hat{p}_{l_{i,j}}\right)\right]_i}{f_{\text{cost}}} \tag{3.20}$$

where f_{cost} is the cost metric presented in equation 3.2.

COST PERCENTAGE OPTIONS Similar to the previous metric, the cost percentage of options metric quantifies the contribution of the up-front options and exercised options to the portfolio's total cost. The metric is calculated using equation 3.21:

$$f_{\text{cost per. options}}(\mathbf{x}_k) = \frac{\sum_{i=1}^{T} \mathrm{E}\left[N_{O_i} p_O + N_{x_i} p_x\right]_i}{f_{\text{cost}}} \tag{3.21}$$

where the up-front options cost is calculated in the first term and the exercised options cost is calculated in the second term in the numerator. With the adaptive options contract, each year i has a distinct N_{O_i} (see equation 2.5); if the adaptive contract is not used, every year has the same value for N_{O_i}.

3.2.3 A Priori Problem Formulation

In the first many objective analysis of the LRGV [56], successively higher-dimensional objective formulations distinguished the impact of adding market-based water supply instruments for urban water portfolio planning within the LRGV. This chapter builds on those results and shifts our focus to evaluating

the appropriate level of decision variable complexity to yield simple and effective risk-based rules to guide the city's market use. The *a priori* decision variable formulation was first defined in [56] and is presented in equation 3.22.

$$\begin{aligned} \mathbf{x}_{a\ priori} &= (N_R, N_{O_{\text{low}}}, N_{O_{\text{high}}}, \xi, \\ &\quad \alpha_{\text{Jan-Apr}}, \beta_{\text{Jan-Apr}}, \\ &\quad \alpha_{\text{May-Dec}}, \beta_{\text{May-Dec}}) \end{aligned} \tag{3.22}$$

Equation 3.22 states that the city's water supply portfolio is determined by volumetric variables for permanent rights (N_R) and the adaptive options contract (low volume alternative $N_{O_{\text{low}}}$ and high volume alternative $N_{O_{\text{high}}}$). The risk-based thresholds are ξ, which the city uses to decide between its high and low options alternatives in every year, and the alpha and beta variables that control options exercising and lease acquisitions. Equations 3.23 through 3.26 present the objectives and constraints used in the *a priori* problem formulation.

$$\begin{aligned} \mathbf{F}_{a\ priori}(\mathbf{x}_{a\ priori}) &= (f_{\text{cost}}, f_{\text{rel}}, f_{\text{surplus}}, f_{\text{costvar}}, \\ &\quad f_{\text{dropped}}, f_{\text{leases}}) \\ &\quad \forall \mathbf{x} \in \Omega \end{aligned} \tag{3.23}$$

$$\text{Subject to} : c_{\text{rel}} : \quad f_{\text{rel}} \geq 0.98 \tag{3.24}$$

$$c_{\text{costvar}} : \quad f_{\text{costvar}} \leq 1.1 \tag{3.25}$$

$$c_{\text{critrel}} : \quad f_{\text{crit. rel}} = 1.0 \tag{3.26}$$

In the above equations, $\mathbf{F}_{a\ priori}(\mathbf{x}_{a\ priori})$ represents the vector-valued objective function for the *a priori* formulation with components as described previously in section 3.2.2. The Ω represents the space of feasible values of $\mathbf{x}_{a\ priori}$, and c variables represent constraints. In this formulation, each of the objectives in the vector-valued objective function are minimized except for reliability, f_{rel}, which is maximized. Also note that each term in the formulation refers to the ten year scenario, with the metrics previously defined. Constraint values for equations 3.24 and 3.25 represent preferred ranges for the objectives and reflect values set in prior work. The critical reliability constraint set to 1.0 reflects a highly risk-averse planning problem developed in [56] in which no simulations have a critical failure

where supply is less than 60% of demand.

3.2.4 Drought Scenarios

The drought scenario tests how alternative LRGV water supply portfolios perform in a single year with statistically unlikely low allocation volumes and maximally high demands. Inclusion of this drought scenario reflects deep uncertainty in planning for multi-year severe droughts. Following the treatment of Knightian or deep uncertainty in [8], extreme droughts are an example condition where a planner cannot fully conceptualize the possible risks to their system. The city begins the scenario with a percentage of their permanent rights, i_{fr}, similar to the beginning of the ten year scenario. In each month, the city must meet the maximum value of the Gaussian distribution of demand adjusted for the tenth year of the multi-year simulation. These maximally high demands are coupled with monthly inflows and allocations from the driest calendar year identified from our input hydrologic data. The 19th year in the record was chosen because it represents the lowest observed inflows and allocation volumes in almost every month. The drought scenario tests the portfolios' performance when the expected value estimates of water allocations are violated (i.e. low inflows and high demands in every month). Since the anticipatory thresholds used for market transfers are based on expected values of the historical inflows and demands, the drought scenario tests how robust the thresholds are to changing assumptions. This shift between the assumed conditions of the ten year planning scenario to a single dry year focuses on supporting robust decision making by exposing vulnerabilities of planning strategies that arise due to model assumptions [73]. Moreover, if failures occur from a historically-observed drought, then the system vulnerability is potentially more severe than assumed in the Monte Carlo probabilistic model due to deep uncertainties associated with the potential climate change impacts and population demands that could impact the LRGV [1, 9]. The drought scenario is used in this study as an exploratory tool for evaluating the robustness of solutions in the many objective tradeoff set.

Table 3.2. Variables for Sobol' Analysis

Decision Variable	Range	Description
N_R	37 - 74	Volume of Permanent Rights [10^6 cubic m]
$N_{O_{\text{low}}}$	0 - 24.7	Low-Volume Options Contract Alternative [10^6 cubic m]
$N_{O_{\text{high}}}$	$1.1N_{O_{\text{low}}}$ - $2.0N_{O_{\text{low}}}$	High-Volume Options Contract Alternative [10^6 cubic m]
ξ	0.1 - 0.4	Low to High Options Threshold
$\alpha_{\text{May-Dec}}$	0.0 - 3.0	Lease/Options Strategy for May-Dec. ("when to acquire?")
$\beta_{\text{May-Dec}}$	$\alpha_{\text{May-Dec}}$ - 3.0	Lease/Options Strategy for May-Dec. ("how much to acquire?")
$\alpha_{\text{Jan-Apr}}$	0.0 - 3.0	Lease Strategy for Jan.-Apr. ("when to acquire?")
$\beta_{\text{Jan-Apr}}$	$\alpha_{\text{Jan-Apr}}$ - 3.0	Lease Strategy for Jan.-Apr. ("how much to acquire?")
dm	0.011 - 0.023	Demand growth (exponential rate)
i_{fr}	0.0 - 1.0	Fraction of initial rights

3.3 Computational Experiment

3.3.1 Sensitivity Analysis

Our analysis first performs global sensitivity analysis of the *a priori* decision variable formulation (see section 3.2.3) using the variance decomposition method of Sobol'. Table 3.2 presents the variables explored in the Sobol' analysis with their respective ranges. The ranges for the decision variables in table 3.2 follow [56], and the reader can consult this reference for more information. The Sobol' analysis also includes the yearly demand growth dm and the initial rights i_{fr} to further explore the effect of model parameters in addition to the decision variables. The range for dm was chosen based on reports of national water use released every five years from the U.S. Geological Survey (U.S.G.S.) [151], focusing on Cameron County, Texas. This county in the LRGV includes Brownsville, the city upon which our hypothetical case study is based. Using the published data from 1985-2005, the

Table 3.3. MOEA Search Parameters

Symbol	Value	Description
M	5,000	Monte Carlo Sample Size
T	10	Planning Period [years]
p_m	1/p	Probability of mutation (p: num. of model param.)
p_c	1.0	Probability of crossover
η_c	15	Distribution index (crossover)
η_m	20	Distribution index (mutation)
NFE	500,000	Number of function evaluations

five-year trends in water use suggest a yearly exponential growth rate of between 1.1% and 2.3%, which is the range for dm used in this study. For the i_{fr} parameter, the Sobol' analysis uses the parameter's full range, from 0.0 representing a failure before the planning period began, to 1.0 representing water supply portfolios that keep a large volume of water in their supply account and consequently end the ten year simulation with a volume of supply approximately equal to their entire permanent rights volume.

Two runs of the Sobol' analysis were performed, the first using the ten-year planning horizon and the second run using the one-year drought scenario. All metrics described in section 3.2.2 are valid for the drought except for cost variability and dropped transfers (the number of years is $T = 1$ and a single draw is performed with characteristics described in section 3.2.4). For both tests, a sample size of $q = 2^{13}$ was used, based on initial tests that showed convergence for this sample size, using an approach similar to [98].

The sensitivity analysis ensemble comprises $q \times (2p + 2)$ simulations of the LRGV management model, each with a randomly generated set of decision variables, demand growth factor, and initial rights. For each of these $q \times (2p + 2)$ simulations, an identical Monte Carlo simulation of 5,000 draws of the historical hydrology, demands, and lease pricing is used. This ensemble size was chosen to coincide with the number of Monte Carlo draws in each function evaluation of the optimization [56]. Furthermore, an identical sample of draws is used for each generated parameter set to reduce the likelihood that our reported variable sensitivity is artificially caused by differences in draws of the input data between runs.

3.3.2 Parameterizing Multi-Objective Search and Handling Uncertainty

The ε-NSGAII is used to create the Pareto approximate tradeoffs for each problem formulation considered. The algorithm's parameters used in this study are given in table 3.3 and were set following recommendations of previous work [56, 85, 88, 152].

The ε-NSGAII utilizes an adaptive population sizing approach that resizes the algorithm's population based on how many archive solutions have been found. Parameters for the adaptive population sizing were set according to [153]. Each MOEA run lasts for 500,000 function evaluations, with each function evaluation comprising M independent Monte Carlo simulations of the water management model. This termination criterion was deemed appropriate by consulting visualizations of preliminary tests that showed very few solutions being added to the archive at this number of function evaluations.

The reference set presented in this work is generated by sorting each model case's final approximated Pareto optimal set together to compare the case's performance against each other. Each case was evaluated using a set of 50 random seed trials per case, to reduce the influence of random effects within the probabilistic search operators and generation of initial random populations. The reference set generation process occurs as follows. In step 1, we perform a sort using the objective values obtained during search (each function call in the algorithm had a unique set of Monte Carlo draws). Step 2 evaluates each solution in this set with an identical large sample of 100,000 draws of the historical data. The rigorous sample used in Step 2 uses an identical ensemble of input data for each evaluation to minimize the likelihood that variance in objective function values for different solutions are an artifact of their uncertain input data. Step 3 enforces constraints with respect to the new objective values found in Step 2. This test of robustness eliminates solutions that were feasible in the algorithm's evaluation, but infeasible in the new larger sample. Step 4 sorts the remaining solutions with respect to the new objective values (from the large sample). The tradeoff presented in this paper reflects the results of this final sort in Step 4. All simulations for both the sensitivity analysis and many objective search were completed using Penn State's CyberSTAR infrastructure, a computing cluster with quad-core AMD Shanghai

Processors at 2.7 ghz and Intel Nehalem processors at 2.66 ghz.

3.4 Results

3.4.1 Sobol Sensitivity Indices

This section presents the results from our Sobol' sensitivity analysis of the *a priori* problem formulation. The goal of the sensitivity analysis is to identify the relative importance of decision variables controlling the city's water supply portfolio, the demand growth parameter, and the city's initial water supply on groups of output performance metrics. Figure 3.3 summarizes the total sensitivity for the ten-year scenario (figure 3.3a) and the drought (figure 3.3b). Each row represents a different evaluation metric, with the metrics grouped into *efficiency*, *risk indicators*, and *market use*. Each column represents a variable studied within the Sobol' Analysis, with the darkness of a block indicating the variable's value for the Sobol' total sensitivity index (equation 2.4), from white representing zero and black representing greater than or equal to 0.6. Recall that the total order indices represent the percentage of the ensemble variance controlled by a parameter's impacts by itself as well as all of its higher order interactions with other parameters.

3.4.1.1 Ten Year Sensitivity

Figure 3.3a presents sensitivity analysis results from the ten year planning period. To compile these results, each set of values for the ten variables in table 3.2 was run within an identical Monte Carlo sample of 5,000 draws for the planning period of ten years with a monthly time step. The columns in the figure show that permanent rights (N_R), the initial fraction of rights (i_{fr}), and the May-December alpha variable ($\alpha_{\text{May}-\text{Dec}}$) were sensitive for almost all metrics. The January-April alpha variable $\alpha_{\text{Jan}-\text{Apr}}$ was also sensitive across many metrics. In contrast, some variables are insensitive across almost all metrics, including the high-volume options alternative ($N_{O_{\text{high}}}$), the low-high options threshold (ξ), and the beta variables ($\beta_{\text{May}-\text{Dec}}$ and $\beta_{\text{Jan}-\text{Apr}}$). This result generally suggests that these variables are not as important in determining the values for the city's planning metrics. The sensitivity across metrics is most similar when the metrics are within the same

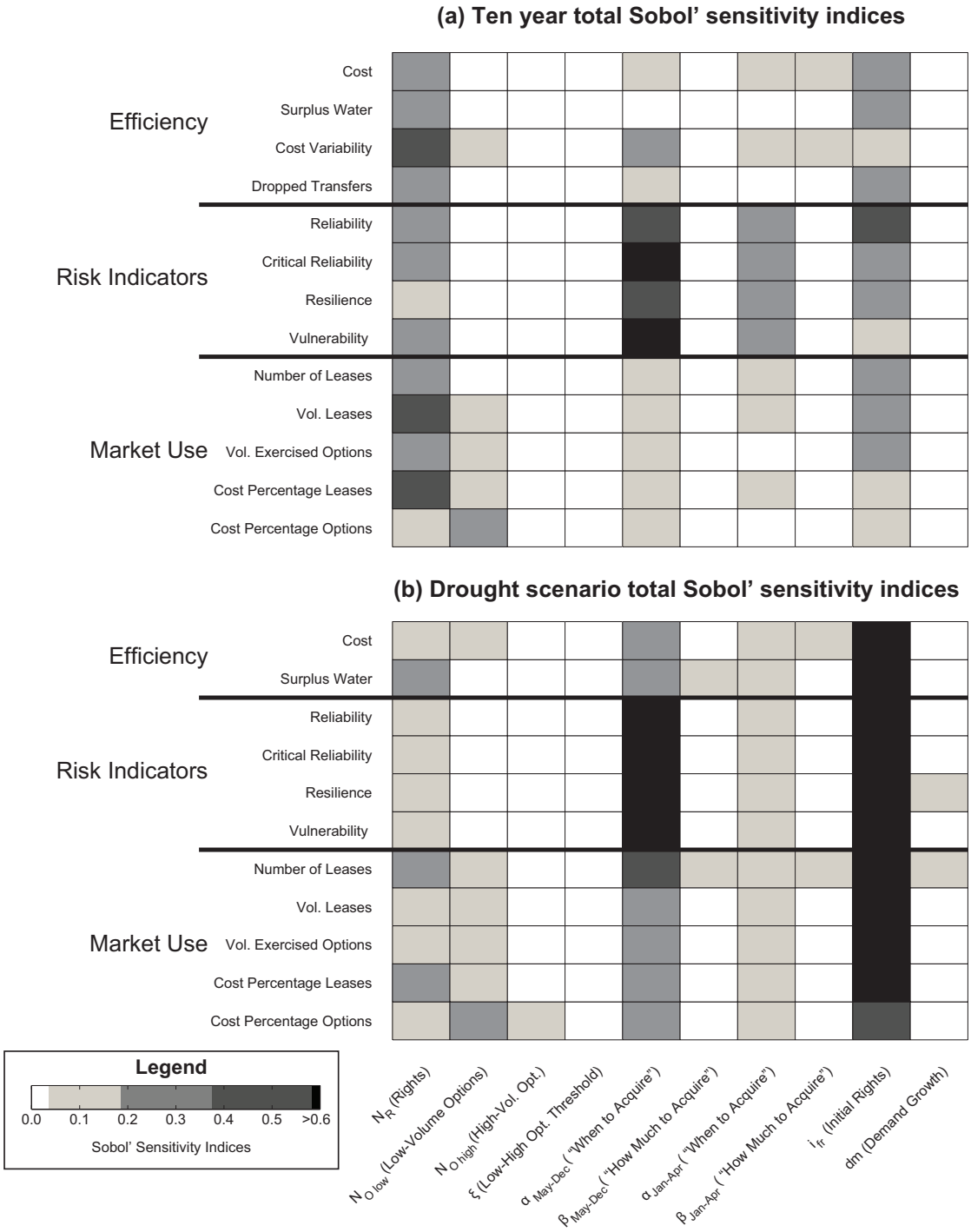

Figure 3.3. Sobol results for the two scenarios. The groups of rows represent groups of decision metrics, with each row a single metric. The columns represent variables, both decision variables and model parameters. The color of each block represents the magnitude of the total Sobol sensitivity index.

group. For example, the same set of variables (N_R, $\alpha_{\text{May}-\text{Dec}}$, $\alpha_{\text{Jan}-\text{Apr}}$, and i_{fr}) is sensitive for each metric in the risk indicator group. Therefore, permanent rights, the threshold for "when" the market is used, and the initial volume of water in the city's account are the largest determinants of its risk-based performance. While the low-volume options volume $N_{O_{\text{low}}}$ was not significant in the risk indicator group, it is a significant control on each many metrics in the market use group, with the rights, alpha variables, and initial rights also having an effect. When comparing the sensitivity structure across groups, however, we see slightly different behavior, such as the alpha variables having a larger influence on the risk indicators than on the efficiency and market use metrics. Similarly, the January-April beta variable $\beta_{\text{Jan}-\text{Apr}}$ is only sensitive in the efficiency group (cost and cost variability) and not in the others.

The sensitivity analysis results show some surprising trends that would run counter to typical problem preconceptions. The relative importance of the permanent rights (N_R) is surprising; the sensitivity analysis suggests that rights are more important than the market-use alpha and beta variables on controlling the volume of leases. Portfolios that have high permanent rights rely on allocations to these rights to fulfill their supply, rendering the market-use variables relatively unimportant for determining the aggregate volume of leases. Permanent rights are also important for the city's efficiency metrics; the surplus water is only sensitive to changes in the permanent rights and initial fraction of rights. Additionally, the insensitivity of the beta variables suggests that these variables do not contribute much to the city's use of the market, apart from influencing cost and cost variability. Using only the most sensitive variables from this section would yield a decision variable formulation that focuses mainly on rights, a single options contract, and the alpha variables for determining market use (omitting beta and the adaptive options contracting).

3.4.1.2 Drought Sensitivity

Figure 3.3b presents results from the Sobol' sensitivity analysis for the drought scenario. Note that each of the metrics in section 3.2.2 is valid for the drought except for the cost variability and the dropped transfers, which are omitted in the figure. The variables identified as being most sensitive across all the metrics in the

ten year analysis are also important in the drought (the rights, alpha variables, and initial rights). There are however some strong differences between the average 10 year planning sensitivities and those attained for the drought scenario. Unlike in the ten year sensitivity results, the high-volume options contract ($N_{O_{\text{high}}}$) influenced the cost percentage for options, and the beta variables influenced the number of leases, cost, and surplus water. Furthermore, the demand growth parameter dm had an influence on resilience and the number of leases. Some of these effects may be due to the single-year nature of the drought. In a single year, the adaptive options contract decision between the high and low volume options alternative is made only once, providing a clearer signal to the sensitivity analysis than in the ten year scenario.

Another contrast to the ten year analysis is in the relative importance of some variables. Every metric is very sensitive to the initial fraction of rights, since a high or low initial fraction of rights in the drought scenario drastically affects the city's supply decisions. If the city has a large volume of water in its supply account (due to a high initial rights), it may not need to perform as many mitigating actions in the drought as when it has little to no water (i.e. in a multi year drought occurring before our hypothetical single year drought begins). The May-Dec alpha variable $\alpha_{\text{May}-\text{Dec}}$ is also very important in the drought, since the city's risk-based thresholds need to be high to counteract the drought's low inflows relative to expected conditions. Permanent rights are less important than in the ten year scenario, since portfolios with high volumes of permanent rights become exposed to failures in the drought's statistically low permanent rights allocations in every month. The contrasts between the drought and the ten year analysis provide further evidence that its inclusion in the planning framework yields a different set of controls on supply portfolio performance that can improve the portfolio's robustness to new conditions.

3.4.2 Sensitivity-Informed Problem Modifications

The second phase of our analysis uses the results from the sensitivity analysis and insights from prior work [56] to inform our *de Novo* many objective analysis. In support of this analysis we have defined a suite of decision variable formulations of

increasing complexity, termed "model cases". The set of model cases is designed to begin with a model that uses only the most sensitive decision variables across all metrics, with subsequent decision variables being added to represent more complex risk-based rules for portfolios. The final, most complex formulation in the set tests the *a priori* decision variable formulation described in section 3.2.3. The analysis therefore attempts to distinguish when reductions in the complexity of our proposed decision rules is warranted within a many objective decision making framework. Furthermore, the multiple model cases allow us to explore sorting at the problem formulation level, to determine which formulations are "non-dominated" in the many objective analysis.

Table 3.4 presents the model cases explored in this study. Model case I builds a water supply portfolio using three decision variables: the volume of permanent rights (N_R), the volume of options (N_O) and the threshold that controls both "when" the city uses the market and also "how much" water it acquires, α. Case II retains the permanent rights and single-volume options contract of Case I but uses two variables to guide the city's market use. $\alpha_{\text{Jan}-\text{Apr}}$ controls the city's acquisition of leases in January through April, and $\alpha_{\text{May}-\text{Dec}}$ controls the city's use of options and lease acquisitions for the rest of the year. Case II was created to explore the effectiveness of the January-April market threshold; this variable was not as sensitive as the May-December threshold in the sensitivity analysis. Furthermore, this added flexibility relative to case I may be useful for the city, in that it allows the city to emphasize a more conservative strategy in one part of the year in order to anticipate the summer drought period. Case III also retains the single-volume options contract but uses both beta and alpha variables for the city's market use. Similar to [56], the beta variables are constrained to always be greater than or equal to alpha and control the volume of water acquired after the city decides to use the market. Setting beta greater than alpha could lead to the city using the market less frequently because it is purchasing larger volumes. While the beta variables were not sensitive for many metrics in the Sobol' analysis, we add them here to isolate the individual effects of groups of decision variables as effectively as possible. Adding betas at this step also serves as a multi-objective evaluation of our Sobol' results, which are limited to uniform, independent sampling of beta and single metric impacts (i.e., they do not inform the decision maker of a decision

Table 3.4. Model Cases

Case	Volumetric Decisions	Strategy Decisions	Notes
I	N_R, N_O	α	Single opt. contract, one alpha controls "when" and "how much"
II	N_R, N_O	$\alpha_{\text{May}-\text{Dec}}$, $\alpha_{\text{Jan}-\text{Apr}}$	Single opt. contract, two alphas control "when" and "how much"
III	N_R, N_O	$\alpha_{\text{May}-\text{Dec}}$, $\alpha_{\text{Jan}-\text{Apr}}$, $\beta_{\text{May}-\text{Dec}}$, $\beta_{\text{Jan}-\text{Apr}}$	Single opt. contract, alphas control "when", betas control "how much"
IV	N_R, $N_{O_{\text{low}}}$, $N_{O_{\text{high}}}$	ξ, $\alpha_{\text{May}-\text{Dec}}$, $\alpha_{\text{Jan}-\text{Apr}}$, $\beta_{\text{May}-\text{Dec}}$, $\beta_{\text{Jan}-\text{Apr}}$	Adaptive opt. contract, *a priori* formulation

variable's impact on a portfolio's nondomination across all objectives)

Model cases I-III exploit the most sensitive variables to explore whether or not the city could meet its water demands effectively with fewer decision variables than the most complex *a priori* formulation. To provide a complete test of how reductions in model complexity affect the decision making problem, the *a priori* formulation is included as model case IV. Model case IV introduces the adaptive options contract in which the city chooses between the low-volume ($N_{O_{\text{low}}}$) and high-volume ($N_{O_{\text{high}}}$) options alternatives based on a threshold decision variable (ξ, see equation 2.5). By comparing the *a priori* decision variable formulation with the reduced-complexity decision variable formulations, our analysis can determine when reductions in the complexity of decision variable formulations greatly change our many objective tradeoffs and inform choice of the simplest yet most effective rules to guide the city's use of the water market.

Cases I - IV are tested within a many objective problem formulation that was formulated to exploit the diagnostic information provided in the sensitivity analysis as well as insights from prior work. Equations 3.27 through 3.31 present the formulation for model case k using the metric definitions of section 3.2.2. The formulation combines analysis components from the ten year analysis (abbreviated "10 yr.") and the drought (abbreviated "dr.").

$$\begin{aligned} \mathbf{F}(\mathbf{x}_k) \quad = \quad & (f_{\text{10 yr. cost}}, f_{\text{10 yr. surplus}}, f_{\text{10 yr. crit. rel}}, \\ & f_{\text{10 yr dropped}}, f_{\text{10 yr num. leases}}, f_{\text{dr. trans. cost}}) \\ & \forall \mathbf{x}_k \in \Omega \end{aligned} \tag{3.27}$$

$$\text{Subject to} : c_{10\text{ yr. rel}} : \quad f_{10\text{ yr. rel}} \geq 0.98 \tag{3.28}$$

$$c_{10\text{ yr. costvar}} : \quad f_{10\text{ yr. costvar}} \leq 1.1 \tag{3.29}$$

$$c_{10\text{ yr. crit. rel}} : \quad f_{10\text{ yr. crit. rel}} \geq 0.99 \tag{3.30}$$

$$c_{\text{dr. vuln.}} : \quad f_{\text{dr. vuln}} = 0 \tag{3.31}$$

In this formulation, each objective is minimized except for critical reliability $f_{10\text{ yr. crit. rel.}}$, which is maximized. The drought transactions cost objective $f_{\text{dr. trans. cost}}$ is the portion of the cost resulting from acquisition of options and exercising of leases (i.e., the permanent rights cost and up-front options costs are omitted).

The objectives of cost, surplus water, dropped transfers, and number of leases were retained from the *a priori* formulation due to their interesting tradeoff structures that guided our analysis in prior work [56]. Our formulation transitions from the reliability objective to a critical reliability objective to more accurately capture the risk aversion in the system. Also, the cost variability objective was removed because it shared a similar sensitivity structure to the cost and it did not provide much information beyond its inclusion as a constraint ($c_{10\text{ yr. crit. rel.}}$). Finally, the drought transfers cost objective was added to the formulation due to the drought's differing sensitivity structure from then ten year scenario, to explore the tradeoff between long-term cost savings in the ten-year scenario and added costs during the drought.

Two constraints were retained from the *a priori* formulation: the reliability constraint ($c_{10\text{ yr. rel}}$) and the cost variability constraint ($c_{10\text{ yr. costvar}}$). The magnitude of the cost variability constraint was introduced in prior work [122] and ensures that the city's cost variability (a statistical measure of high magnitude, low probability costs) is not significantly higher than their average cost in each planning year. Cost variability is effective as a constraint, but adding it as an objective would not contribute to the analysis since our prior work found that its tradeoffs have limited planning power. The constraint on critical reliability $c_{10\text{ yr. crit. rel}}$ was modified to $f_{10\text{ yr. crit. rel}} \geq 0.99$ to capture a wide range of objective performance while still allowing a sufficient level of high performance for the portfolios. Recall that the reliability and critical reliability are calculated using the expected number of failures in the scenario. Levels of the reliability constraints in this work reflect

a very high reliability relative to typical planning by water utilities. Finally, the drought vulnerability constraint $c_{\text{dr. vuln.}}$ was added to ensure that no candidate portfolios had any failures in the drought to capture the severe risk aversion that characterizes urban water planning. Note that our alternative formulations all represent severely challenging, highly constrained explorations of stochastic decision spaces. It is a major advancement in this work to be able to approximate the many objective tradeoffs for these alternative problem formulations.

Data from the reference set for the most conservative formulation in [56] (equivalent to the *a priori* formulation here) were used to determine typical values for the initial rights for portfolios. After beginning the simulation with an i_{fr} of 0.3, the portfolios maintained a ratio of their surplus water to permanent rights at approximately 0.4. We fit a normal distribution to the first simulation year of these results with a mean of 0.4118 and a standard deviation of 0.0285. The initial rights i_{fr} in the ten year scenario were then sampled from this distribution in each Monte Carlo draw. The city also began the single year in the drought scenario with a i_{fr} of 0.4118 to maintain continuity with the ten year scenario. In both scenarios, the demand growth parameter dm was set to 0.023, the largest value in its feasible range based on our analysis of the USGS data.

The ε-NSGAII was used to generate solution sets using the sensitivity-informed problem formulations for the model cases. As introduced in section 3.3.2, the noise in calculations of the mean objective function values is used to parameterize the epsilon dominance settings within the algorithm. To capture a representative sample across model cases and solution types, a set of 20 non-dominated solutions was chosen from preliminary algorithm runs. For each solution, we performed 50 draws of the model, using a new set of 5,000 Monte Carlo draws of the input data in each sample. The CDFs in figure 3.4 reflect the deviation of each solution's objective function calculations about the mean, calculated across these 50 draws. The figure shows variations in the 50 replicate ensemble mean calculations using the ten year model for cost (figure 3.4a), surplus water (figure 3.4b), critical reliability (figure 3.4c), number of leases (figure 3.4d), and dropped transfers (figure 3.4e). Note that the drought transfers cost is not sampled because this objective is calculated deterministically within the single drought scenario. Each solution is shown with a light gray line, while the solution with the largest range is shown with a

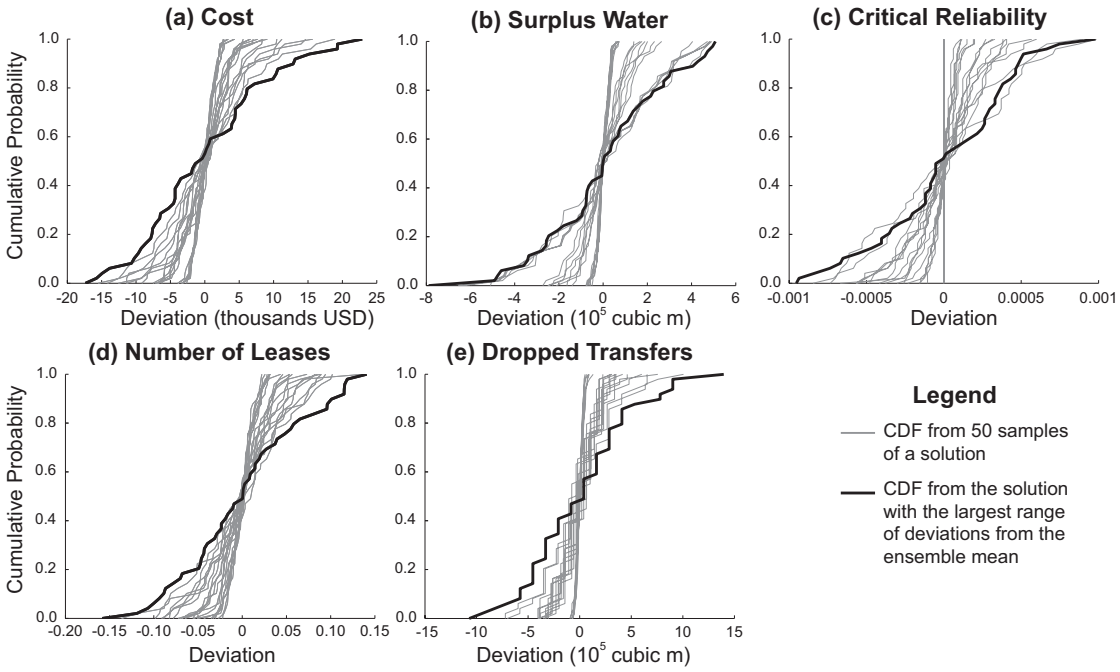

Figure 3.4. Cumulative distribution plots of objective function performance for 20 selected solutions. Each CDF was constructed from an ensemble of 50 calls of the noisy objective function.

bold black line. Table 3.5 summarizes the noise-filtering epsilon settings that result from this analysis. These larger epsilons provide two major benefits for the MOEA search: (1) they enhance the robustness of solutions identified in search and (2) they reduce the computational demands of search by reducing the size of the Pareto-approximate solution sets attained for each model case. The noise adjusted epsilons serve to coarsen the resolution of the non-dominated set and for formulations with small objective counts it is important to confirm that a sufficient representation of tradeoffs is attained. As the number of objectives increases, this issue reduces as a concern because of the rapid growth rates of Pareto optimal sets with objective count.

3.4.3 Multi-Objective Tradeoffs

This section presents the multi-objective solution set for model cases I-IV using the sensitivity-informed *de Novo* problem formulation developed in this work. Figure 3.5 shows the model cases sorted together, to directly compare the objective

Table 3.5. Objectives' Epsilon Settings

Objective	Value
10-yr Cost	\$30,000
10-yr Surplus Water	1,233 cubic m
10-yr Critical Reliability	0.002
10-yr Drops	2,467 cubic m
10-yr Number of Leases	0.3
Drought Trans. Cost	\$10,000

function performance of the cases relative to each other. In the figure, each cone represents one alternative supply portfolio, with the cone's coordinates plotting each portfolio's cost ($f_{10\text{ yr. cost}}$), number of leases ($f_{10\text{ yr. num. leases}}$), and surplus water ($f_{10.\text{ yr. surplus}}$). The orientation of each cone represents the percentage of cost due to market transfers (the sum of the cost percentages of options and leases). Cones oriented vertically along the surplus water axis and pointing toward the lowest value of surplus water (i.e. pointing downward) represent portfolios with limited market use, whereas cones pointed toward the highest value of surplus water have up to 31 percent of their cost in market transfers. The size of the cones plots the portfolios' critical reliability, with small cones indicating $f_{10\text{ yr. rel}} = 0.99$ and the largest cones indicating a value of 1.0 for this objective. The color of the cones indicates each model case: cones for case I are navy blue, case II is shown in cyan, case III's cones are yellow, and case IV is shown using red cones. The ε-NSGAII was run for 50 random seed trials for each model case. After these runs, the model cases were combined together to create a final reference set using the four step approach outlined in section 3.3.2.

Of the 447 solutions in the total set, model case I contributed 108 solutions (24 percent of the total), model case II contributed 18 solutions (4 percent), model case III contributed 196 solutions (44 percent), and model case IV contributed 125 solutions (28 percent). Two distinct groups of solutions emerge. The first set has high values for cost and surplus water with low percentages of their costs in the market and are hereafter referred to as "permanent rights-dominated solutions" (see highlight (i) in the figure). The rest of the solutions, termed market-dominated solutions and shown in area (ii) of the figure, are typified by lower costs and surplus water values and a range of values for the number of leases objective. The blue arrows in the figure indicate the preferred direction within the tradeoff (minimum

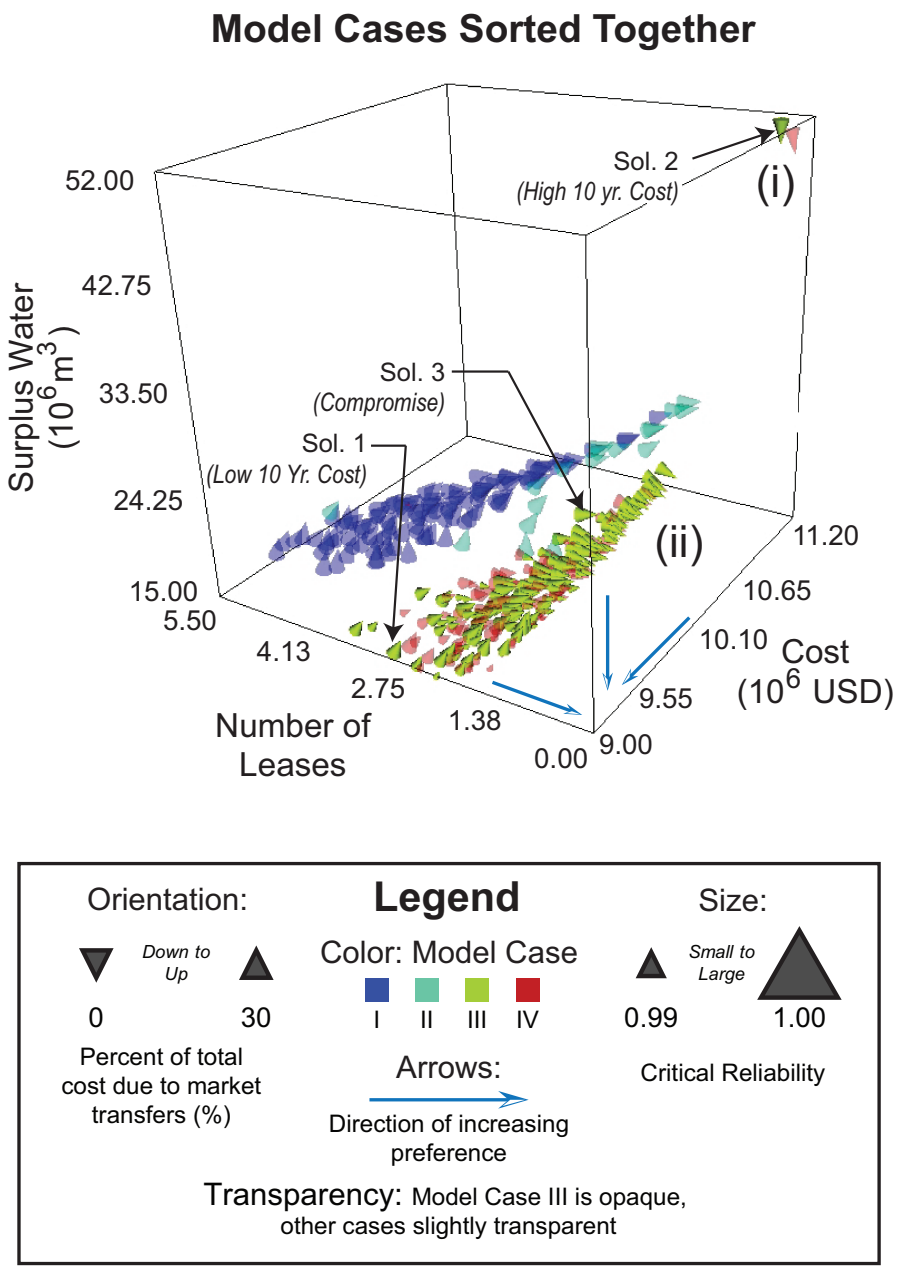

Figure 3.5. Approximations of the Pareto set for model cases sorted together. The spatial axes plot surplus water, number of leases, and cost objectives. The color of the cones indicates the model case. The orientation of the cones represents the percentage of total cost due to market transfers, whereas the size of the cones represents each portfolio's critical reliability. Model case III is shown opaque with other cases shown with transparency for the purpose of highlighting the contribution of case III.

values for the three spatial objectives).

Model case I is the simplest decision variable configuration, with the city's market use determined by three decision variables: N_R for rights, N_O for the volume of the options contract, and α controlling all lease acquisitions and options exercising. The market-dominated solutions from this case in figure 3.5 are able to achieve efficient performance relative to the permanent rights-dominated solutions, but the other cases have preferred performance in the spatial axes relative to case I. For example, the added flexibility in model case II (a different α in the beginning of the year than in the end) yields more diversity in the tradeoff (i.e. a larger objective value range). The large size of solutions in cases I and II shows that they exhibit good performance with respect to the reliability objective. The solutions from these cases have better costs, number of leases, and surplus water performance relative to I and II. Interestingly, these solutions also do so by using less of their cost in the market (i.e. they are oriented pointing downwards relative to solutions in cases I and II). These results suggest that solutions in these cases can achieve high reliability using fewer leases on average than I and II (due to the betas being higher than alphas). The solutions from Cases III and IV span a wide range of the spatial objectives in figure 3.5. Examining the spatial axes in the figure, it is difficult to ascertain the difference between the objective value performance of cases III and IV.

Based on results of figure 3.5, the final stage of our analysis focuses on the contributions of model case III to this final reference set, because it represents a simple and effective formulation of the planning problem. This case had the largest number of solutions in the reference set, and its contribution to the reference set featured solutions from both the permanent-rights dominated solutions and the market-dominated solutions. Figure 3.6 presents a parallel coordinate plot of case III's contribution to the total reference set. Each line in this plot represents the objective function values for a single solution. The values on each axis are plotted such that the position on the vertical numberlines represents the relative magnitude of each objective, and the direction of increasing preference for each objective is pointing downward. Because of this plotting convention, conflicts between objectives can be seen when the lines cross [51]. Color represents the percent of the total cost due to market use, similar to the reference sets figure,

ranging from 2.0 percent in blue to 27.5 percent in maroon.

Market-dominated solutions (yellow and red solutions that spend more of their supply costs in the market) exhibit low costs and surplus water, with higher values for the dropped transfers $f_{\text{10 yr. drop}}$ and number of leases $f_{\text{10 yr. num. leases}}$ objectives. Solutions with higher percentages of their cost in the market actually have lower values for their 10 year aggregate $f_{\text{10 yr. cost}}$ objective. A prominent conflict exists between the cost ($f_{\text{10 yr. cost}}$) and the drought market transfers cost ($f_{\text{dr. trans. cost}}$) objectives. Portfolios that exhibit high cost in the long-term scenario have cost savings in the drought. This result, though, is contingent on our assumption of constant initial rights in the drought. The assumption is equivalent to assuming that a drought would occur after a "typical" year with the city holding a sufficient amount of water in its supply account. As a further exploration of solution performance, the next section will explore the impact of this initial condition on several selected water supply portfolios.

Table 3.6 and the annotations in figures 3.5 and 3.6 show three solutions selected for further analysis. In table 3.6, the $\%N_x$ metric is defined as percentage of Monte Carlo draws in which the city exercises their options (calculated as an average across the 10 simulation years). All other metrics are as defined previously, and each measures performance in the ten year scenario except as noted. Solution 1 is termed "Low 10 Yr. Cost." It has preferred cost and surplus water objective performance, and exhibits the highest market use of any of the selected solutions. Note in table 3.6 that there is a distinct difference in this solution's α and β values; in May-December, the beta is 1.53 whereas the alpha is 1.29. This separation would not have been possible in the simpler decision variable formulations of cases I and II. For solution 1, 25.1 percent of its cost is from market use, with 77.1 percent of the Monte Carlo draws resulting in an exercised option. Solution 2 is termed "High 10 Yr. Cost." It has higher costs and surplus water relative to solution 1 but its low market use and tendency to store large volumes of surplus water yields preferred performance with respect to solution 1 in several other objectives. Solution 2 also has very low drought transfers costs, due to the fact that the high volume of water carried over at the beginning of the drought allowed the city to avoid high volumes of market transfers in the drought scenario. Solution 2's $\alpha_{\text{Jan}-\text{Apr}}$ and $\beta_{\text{Jan}-\text{Apr}}$ values are very low, indicating that this solution would never have had to buy

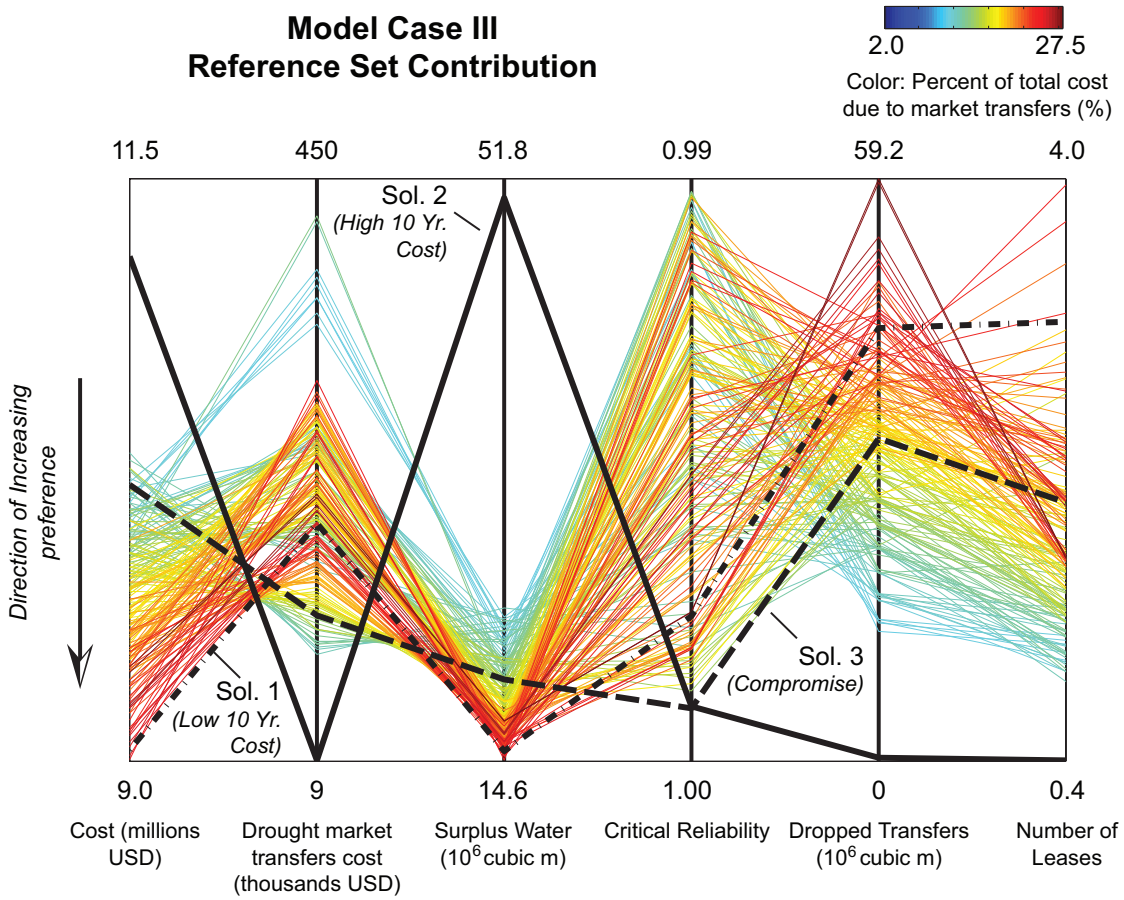

Figure 3.6. Parallel coordinate plot for model case III's contribution to the reference set. Each solution is represented by a line, with the color of the line representing the percent contribution of market transfers to the solution's planning period cost, and the vertical position of the line representing the relative objective function value for each respective objective. Solutions 1 - 3 are picked for further analysis and are shown with bold lines.

leases in these months. Solution 3 is termed the "Compromise" solution. Each of its objective function values falls between the magnitude of the other solutions' objective function values. The solution has relatively high volumes for both the permanent rights N_R and the options contract N_O, and it exhibits some separation between α and β values. Examining the market use metrics in the third part of table 3.6 shows that solution 3's market use is lower than solution 1's, and that solution 3 exercised its options fewer times than solution 1 did. Additionally, its critical reliability $f_{10\,\text{yr. critrel}}$ is higher than in solution 1. In general, the high cost solution 2 represents how utilities typically plan, whereas the solution 1 could be

Table 3.6. Selected Solutions' Performance

Solution	1	2	3
Name	Low 10 Yr. Cost	High 10 Yr. Cost	Compromise
$f_{\text{10 yr. cost}}(10^6\ \$)$	9.05	11.12	10.19
$f_{\text{10 yr. surplus}}\ (10^6\ \text{m}^3)$	15.2	50.6	19.8
$f_{\text{10 yr. critrel}}\ (\%)$	99.8	99.9	99.9
$f_{\text{10 yr. dropped}}\ (10^6\ \text{m}^3)$	44.0	0.365	32.8
$f_{\text{10 yr. numleases}}$	2.74	0.404	1.78
$f_{\text{dr. trans. cost}}\ (10^6\ \$)$	0.189	0.00975	0.119
$N_R\ (10^6\ \text{m}^3)$	37.0	60.8	46.0
$N_O\ (10^6\ \text{m}^3)$	17.2	0.0049	21.0
$\alpha_{\text{May}-\text{Dec}}$	1.29	1.29	1.24
$\beta_{\text{May}-\text{Dec}}$	1.53	1.34	1.60
$\alpha_{\text{Jan}-\text{Apr}}$	1.20	0.06	1.39
$\beta_{\text{Jan}-\text{Apr}}$	1.28	0.09	1.39
$f_{\text{cost per. rights}}(\%)$	74.9	99.8	82.7
$f_{\text{cost per. options}}(\%)$	22.7	0.00	16.0
$f_{\text{cost per. leases}}(\%)$	2.4	0.18	1.34
$\%\ N_x$	77.1	1.53	44.6
$N_x\ (10^6\ \text{m}^3)$	10.8	0.0	5.95
$N_l\ (10^6\ \text{m}^3)$	2.19	0.110	1.30

implemented by a utility that was highly tolerant of risk. Solution 3, though, is a reasonable amount of market use for a risk averse utility that has high potential to lower costs while maintaining high reliability through using the market.

3.4.3.1 Exploration of Solutions through the Drought Scenario

Recall that the drought scenario is a single statistically dry year coupled with the highest simulated demands. This section explores the effect of changing the initial volume of water available to the city at the beginning of the scenario, to show how deeply uncertain model assumptions about water availability or timing of drought can affect the city's water portfolio performance. The three solutions chosen for further analysis each had good objective function performance across all six objectives considered in the many objective problem formulation. The different values for the initial condition in this section serve as a proxy for simulating when a severe drought occurs. The first test, $i_{fr} = 0.4$, tests how the city's supply portfolio would perform if a drought occurred after a year with average supply on hand. The value used is slightly lower than what was used in the optimization, and this small perturbation is well within the uncertainty on how much the city would have to

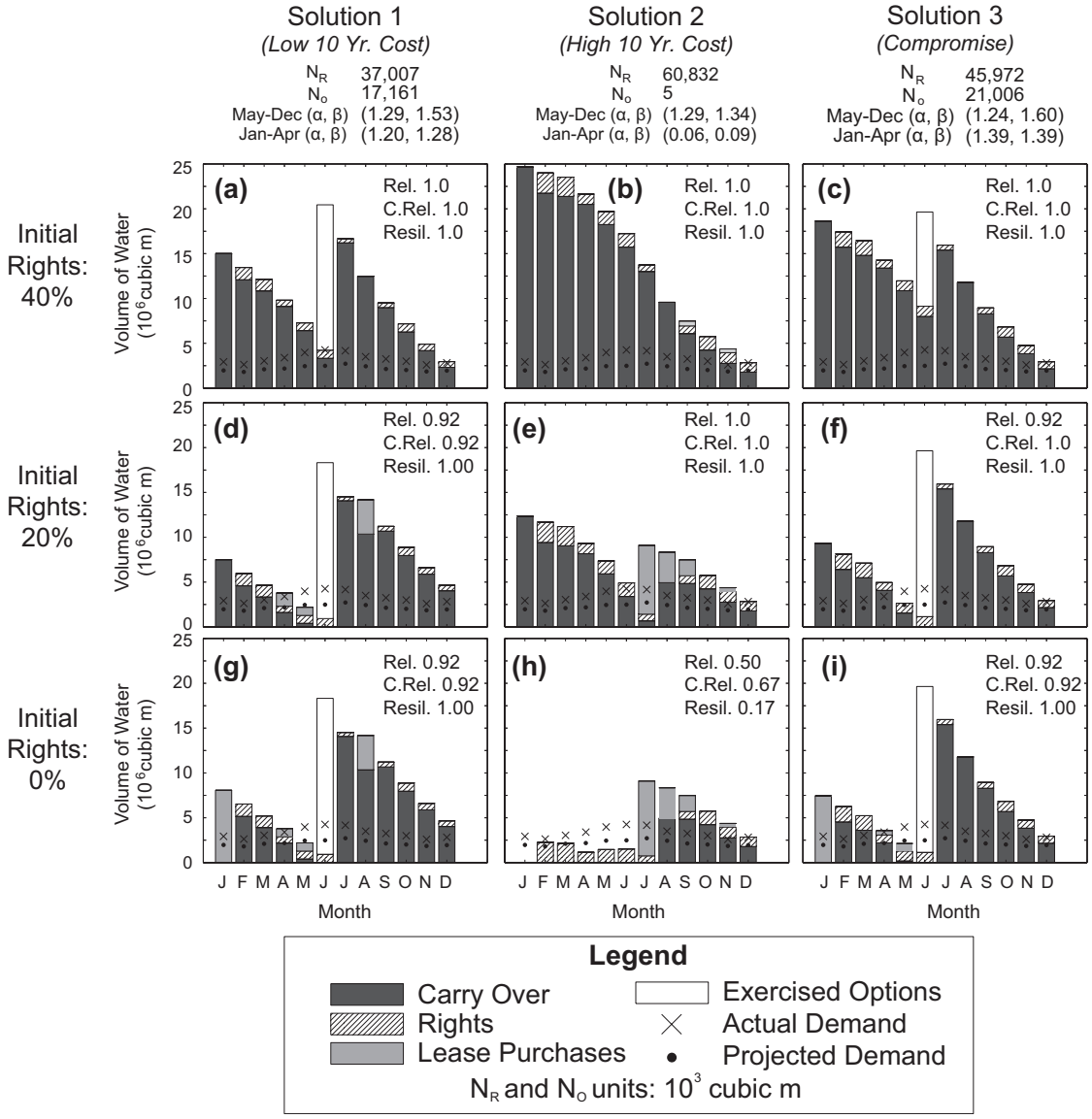

Figure 3.7. Drought scenario analysis results. The bars show the amount of water in the city's supply account in each month. The scenario demand is shown with an x, and the projected or mean demand is shown with a circle. Failures occur when the ordinate of the demand is higher than the city's supply. The columns represent different selected solutions, and the rows represent different assumptions of initial rights that begin the scenarios. Risk indicators are shown in the upper right hand corner, with "Rel." representing reliability, "C.Rel." representing critical reliability, and "Resil." representing resilience.

hold from year to year. The second test has $i_{fr} = 0.2$, which will greatly limit the amount of water in the beginning of the drought simulation and trigger different behavior for market utilization. Lower initial conditions will specifically test the January-April market utilization variables, since the city will be more vulnerable in those months than on average. The final test has $i_{fr} = 0$. An equivalent situation to this test would be a multi-year drought in which the city experiences a failure in the month before our drought scenario begins.

Figure 3.7 presents the results of these tests, with each solution's decision variable values shown above the bar charts, and each panel showing the city's equivalent reliability ("Rel."), critical reliability ("C. Rel.") and resilience ("Resil.") for that test. The first row is for the initial condition at 40 percent of the portfolio's permanent rights (i.e., $i_{fr} = 0.4$), the second row shows 20 percent, and the third row shows 0 percent (no water carried over). The three columns each represent the three selected solutions. The vertical axis on each bar chart shows a water balance for the city's water supply in each month of the simulation, with each bar representing one of the 12 months of the year starting in January. The initial conditions for each test are defined using equation 2.6, where the city is granted a starting volume of water commensurate with its volume of permanent rights N_R. Recall that this initial condition is considered a surrogate for the water that is often left in the city's water supply account from the previous year.

After January where the initial condition water begins the simulation, the city's water from the previous months that is not used is indicated as "carry over" water and shaded dark gray. For the months of February through December, allocations to the city's permanent rights are shown using hashed lines, purchased leases are shown in light gray, and exercised options are shown in white. The demand volume used in the anticipatory thresholds (i.e. the expected value demand) is shown with a circle. However, the drought scenario's maximal demand that the city must actually meet is shown with an x. Failures occur when the demand (shown with an x) falls higher than the supply (of carried over water, rights allocations, options, and leases) indicated by the bar. Note that for each of these tests, our drought scenario examines what would happen if the city used the market in the exact manner that the evolved portfolios dictated. Our goal here is to examine the solutions' performance as if the decision rules were codified such that it would be

exceedingly difficult to change the regulations even in an emergency situation.

Panels (a)-(c) of figure 3.7 show the tests when the initial condition was 40 percent, slightly lower than what was used in the optimization. All three solutions had no failures in this test, but each leaves December with much less water than its starting volume. Solutions 2 and 3, in particular, had barely enough supply to meet its December demand, and in the subsequent year they would start with a condition similar to the 0 percent initial condition shown in panels (g)-(i). Another consequence of this test is that the large volume of initial condition water in panel (b) for solution 2 quickly degrades with the low allocations and high demands of this drought scenario. This result highlights that while portfolios may be able to rely on carry-over water for supply in typical conditions, violating these assumptions may expose the city to severe risks for supply failures.

Transitioning to panels (d)-(f), the city has a lower percentage of its rights available as an initial condition. Market decisions triggered by lower volumes of water supply in the first months of the year before May will be determined by the January-April α and β variables. Even though solution 1 in panel (d) has fairly high values for these variables, the moderate volume of water available is enough to not trigger the thresholds until the end of April into May, when a small lease is not enough to prevent a failure in May. After the option exercise month, the city has sufficient volumes of water to avoid failure. Solution 3 exhibits a similar behavior in panel (f), with the same performance metric values without having to purchase a lease due to higher initial conditions and different strategy variables. Note also that one of the failures of solution 1 in panel (d) was critical (i.e. the city could provide less than 60 percent of its total demands), whereas the failure for solution 3 in panel (f) was not critical. A combination of the carry-over water and multiple leases allowed solution 2 to avoid any failures in this test. An interesting aspect of this test, though, was that the city is forced to rely on many leases to avoid failures here, whereas it only purchased 0.404 leases on average in the 10-year scenario.

Panels (g)-(i) show the test in which the city has no water in its account before the drought scenario begins. Recall that portfolios that have very low water (or a failure) in the last month of the other drought scenario tests would be forced to confront a test such as this in the subsequent year (i.e. there is no water available and the drought continues). The adaptivity of the market-dominated

solutions 1 and 3 in panels (g) and (i) allowed both portfolios purchase leases in January, April, and May, with an additional purchase by solution 1 in August. Both solutions had one critical failure, but exhibited resilient performance in that there were no failures in the month after the shortage. Solution 2 in panel (h), however, had failures in six out of the twelve months. The low α and β variables in the first five months paralyzed the city's water supply, with failures in January through June (many of which were critical failures). After the option exercise month, the city's market use variables allowed it to purchase several leases and it had reliable supply until the end of the year.

The results show the importance of having sufficient values for the volumetric variables (N_R and N_O) as well as risk-based triggers that are conservative enough such that the city uses the market at appropriate times. Portfolios that had their β values higher than α tended to use the market less frequently for larger volumes which can increase supply security. The comparative analysis done in this section also focuses our attention on the difference between early and late year patterns (i.e. in the months before and after the option exercise month); across different initial conditions, the pattern in the early or late months are the same. Thus, the initial conditions have less of an effect for some portfolios after the options exercise month. Another important result is that even the safest portfolios had a low amount of water at the end of the year, but the compromise solution has the highest volume left on average. These insights can inform further problem modifications in subsequent iterations of the problem. For example, the assumptions of initial conditions or the appropriate planning time scale (short term versus long term) can be critically important for developing effective water supply portfolios.

3.5 Conclusion

This work supports the view that decision variable and objective formulations are constantly changing and being improved by new learning or decision maker preferences [68]. Typical environmental planning problems are solved within a static formulation using a quantitative model with a fixed set of decision variables that determine the planning strategy. The sensitivity analysis using Sobol' variance decomposition, however, showed that several variables were insensitive across

many candidate planning metrics considered in the work. Our sensitivity-informed changes to our formulation ("*de Novo* planning") removed the variables that did not contribute to the variance of key water supply evaluation metrics. Changes to our planning problem formulation also included the addition of a drought cost objective because drought year sensitivities were significantly different than average year controls. The full diagnostic value of these changes was not seen, though, until interactive visual analytics were used to review the MOEA-generated tradeoffs. Visualization of the alternative formulations' many objective tradeoffs showed that a decision variable formulation of moderate complexity (model case III) was sufficient to develop portfolios that exhibited improved efficiency with respect to surplus water, balanced market use, and reduced water supply risks.

This study shows the effectiveness of combining multiple tools for increasing the robustness of environmental planning. The framework demonstrates the value of global sensitivity analysis, many objective optimization, and interactive visual analytics to promote problem understanding and incorporate stakeholder learning into the planning process. This study also contributes computational advances in solving noisy multi-objective problems by introducing the concept of noise-adjusted epsilon dominance settings to filter solutions that are highly uncertain and lack robustness with respect to planning objectives.

The robustness of planning solutions is further explored in this study using the concept of deep uncertainty, which is characterized by decision makers not being able to predict accurately the categories of risk to which their systems are vulnerable. In addition to the average year planning horizon, the extreme drought scenario showed how market use can help cities avoid catastrophic failures in times of low water availability. Our exploration of deep uncertainty in the drought scenario indicated that solutions that depended on large volumes of "carry-over" water for supply would be left exposed to critical water shortages after twelve months of low water supply (i.e., a drought that extends beyond a single calendar year). Resolving this challenge requires a compromise between high market use and high permanent rights use to minimize the impacts of sustained drought while simultaneously reducing a city's long term planning costs and surplus water capacity in non-drought years.

These insights have been contingent on a set of assumptions: a single best esti-

mate of the distribution of input data to the LRGV (i.e., hydrology, demands, and lease pricing) as well as estimates of key model parameters such as the population growth rate. Chapter 4 will explore these assumptions, and determine if our choice of model case and constituent candidate solutions is still valid if these assumptions are violated.

Chapter 4

Many Objective Robust Decision Making

This chapter contributes the many objective robust decision making (MORDM) framework. MORDM builds on the many objective evolutionary decision support work in previous chapters, adding concepts and methods from robust decision making (RDM) to facilitate the management of complex environmental systems. Many objective evolutionary search is used to generate alternatives for complex planning problems, enabling the discovery of the key tradeoffs among planning objectives. RDM then determines the robustness of planning alternatives to deeply uncertain future conditions and facilitates decision makers' selection of promising candidate solutions. MORDM tests each solution under the ensemble of future extreme states of the world (SOW). Interactive visual analytics are used to explore whether solutions of interest are robust to a wide range of plausible future conditions (i.e., assessment of their Pareto satisficing behavior in alternative SOW). Scenario discovery methods that use statistical data mining algorithms are then used to identify what assumptions and system conditions strongly influence the cost-effectiveness, efficiency, and reliability of the robust alternatives. The framework is demonstrated using the LRGV test case (section 2.4.1). This chapter is adapted from a study published in *Environmental Modelling and Software* [33].

4.1 Introduction

This paper contributes the many objective robust decision making (MORDM) framework, which combines many objective evolutionary optimization, robust decision making (RDM), and interactive visual analytics to facilitate the management of complex environmental systems. The MORDM framework seeks to address several key challenges for environmental systems undergoing change. The first is how to evaluate the performance of alternative planning and management strategies. To make these evaluations, planners have traditionally used cost benefit analysis, in which a project's benefits are commensurated to their expected monetary value and then compared to a project's costs to determine whether the project will be funded [10]. Aggregating these multiple performance measures into a single value can yield negative decision biases that result because different aspects of performance are rewarded and penalized in ways that cannot be predicted *a priori* [45]. The approach has also been shown to inadequately compensate for non-monetary benefits [154] especially when multiple policies are considered [13]. Multi-objective approaches instead seek to quantify the large number of conflicting objectives that characterize planning. In addition to cost, it has been recognized that complex planning efforts often reveal additional critical performance objectives [66], such as maximizing reliable performance, minimizing environmental damages, and improving system efficiency. The Harvard Water Program was one of the earliest efforts to advocate for the multi-objective planning approach by emphasizing the importance of both economic objectives and engineering performance objectives [16, 24, 81]. Considering "many" objectives explicitly and simultaneously can also aid planners in avoiding cognitive myopia [57]. Cognitive myopia arises when decision makers inadvertently ignore aspects of the problem (such as important decision alternatives or key planning objectives) by focusing on a limited number of *a priori* specified alternatives or a narrowly defined, highly aggregated definition of "optimality" [15].

As discussed in the prior chapters, a problem with multiple objectives can be posed as a vector minimization problem, which seeks to find the Pareto approximate set of solutions. In practice, there are several factors that make it difficult to find Pareto approximate solutions for complex problems. Severe performance con-

straints may arise due to regulatory requirements [155], resource limitations [156], or risk aversion [56, 122, 140]. The presence of constraints severely limits the number of feasible solutions. Also, the simulation models that quantify performance for complex environmental systems may include many interacting subsystems, such as agricultural, municipal, and environmental concerns, that introduce non-linearities and non-separable dependencies [23, 54, 157, 158]. These complexities have motivated a growing number of researchers to exploit MOEAs. The mathematical challenges that have motivated the growing use of MOEAs are very relevant to the planning and management of complex environmental systems. These systems require a risk-based, stochastic approach in which decision makers can evaluate the resilience of their system [40] as well as the adaptability and robustness of their decisions [159]. A growing number of studies have recently explored using MOEAs under uncertainty [32, 56, 149, 160–165]. However, such expected value calculations of the objectives and constraints may not sufficiently characterize risk under conditions of deep uncertainty.

The term *deep uncertainty* (see section 2.2.1) refers to components of a planning or management problem where decision makers cannot agree upon the full set of risks to a system or their associated probabilities [7–9, 47]. Land use change, the depletion of resources, and climate change are three examples of human-induced changes that introduce deep uncertainties into planning problems [1, 70, 71]. Decision makers often use scenario analysis [72] to help plan for such challenges. Scenario analysis uses a small number of plausible values for key planning variables (such as economic growth) to create storylines for future conditions in a system. Groves et al. [73] highlighted several important limitations associated with using scenario analysis for deeply uncertain factors within a planning problem. A large number of different factors can shape the future, and a small number of scenarios cannot adequately cover all interactions between the different factors. Moreover, there is little guidance on how scenarios should inform decision making. As introduced in section 2.2.3, decision support strategies that use the concept of *robustness* can help address these challenges, by identifying strategies that perform well across many different assumptions regarding the deeply uncertain factors [9, 30, 76, 77]. This chapter uses Robust Decision Making (RDM) [9, 47, 78, 79] to characterize robustness. RDM evaluates the performance of policy strategies over an ensemble of

deeply uncertain trajectories of the future. Decision makers select plausible ranges for each deeply uncertain factor, but they do not *a priori* select scalar values for a small number of scenarios. Instead, RDM employs statistical algorithms to "discover" scenarios, which are ranges of the exogenous factors[1] that in combination cause poor performance [78]. RDM therefore provides a tool for decision makers to determine how changes in their assumptions about exogenous factors affect the performance of their planning strategies.

The purpose of this chapter is to propose and demonstrate the MORDM framework, by combining the strengths of MOEA optimization and RDM. The MORDM framework makes two primary contributions. First, RDM has not previously incorporated global optimization techniques such as MOEAs to discover planning alternatives. MORDM exploits MOEAs to solve many objective problems of four or more objectives, thus providing a rich set of alternatives as inputs to RDM. Second, we address the issue of selecting a preferred solution from MOEA-generated tradeoffs. Solution robustness is a promising way to ensure acceptable performance even if system conditions strongly deviate from those used to evaluate the optimality of alternatives (i.e., Pareto satisficing[2] behavior in extreme states of the world). MORDM represents *a posteriori* decision support, in that it does not require assumptions about decision maker preferences before the analysis begins. This study demonstrates how interactive visual analytics [110–112] can support collaborative decision making and enhance planners' ability to effectively process the large amount of information generated by the MORDM framework. This research builds off the historical work in joint cognitive systems [114, 115] with the intent of maximizing the combined analytical strengths of humans and computers when addressing planning and management problems for complex environmental systems. Furthermore, the MORDM framework is designed to emphasize learning and stakeholder feedbacks as part of the decision making process. Our focus on learning and stakeholder feedbacks reflects the fact that public planning problems are rarely static, and have formulations that must change over time [65].

[1]Exogenous factors are those in the modeled system considered not under the direct control of the decision makers, in contrast to the modeled factors considered part of the decision makers' choice set.

[2]By Pareto satisficing, we mean that a solution's performance remains close to the Pareto optimal surfaces for each of many future states of the world.

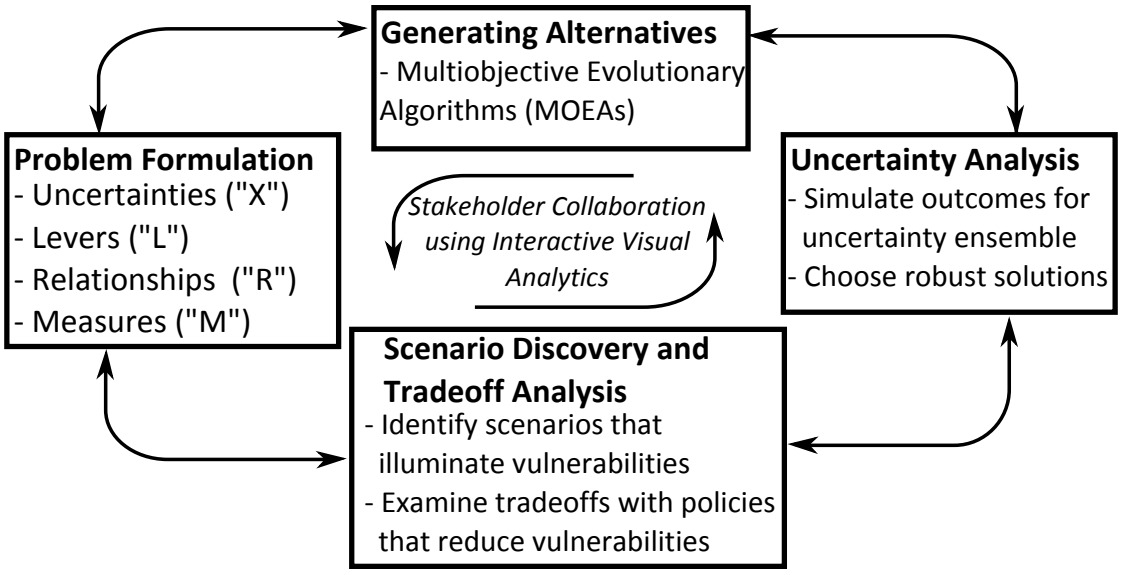

Figure 4.1. The four steps of the Many Objective Robust Decision Making (MORDM) framework. The process typically begins with problem formulation. Each step facilitates stakeholder involvement using interactive visual analytics.

Figure 4.1 presents the MORDM framework, consistent with the iterative "deliberation with analysis" decision support process recommended by [28]. Initially, one or more problem formulations begin the process. Each problem formulation is a formalized hypothesis about what the decision makers feel is most important for their problem, and it is continually updated as we learn more about the system [67, 69, 166]. For example, [32] introduce the *de Novo* planning where a variety of decision strategies, objectives, and constraints are explored simultaneously to carefully evaluate the strengths and weaknesses of alternative problem formulations. The acronym XLRM [47, 61] is used to describe the four problem formulation components. The formulation identifies uncertainties ("X"), factors beyond the decision maker's control, using two main categories. The first category of uncertainties have well-characterized distributions. The second category, termed deep uncertainties, occur when the model representing a system, key model parameters, or the probability distributions representing classical uncertainties are not known or cannot be agreed upon [61]. Decision levers represent actions the decision makers can take to modify their system ("L"). A quantitative relationship ("R") maps decision maker actions to outcomes, typically using a simulation model. Finally, performance measures ("M") are used to gauge success. After

defining an initial problem formulation, the second step of MORDM, termed Generating Alternatives, uses an MOEA to find a Pareto approximate set of solutions. This step incorporates the classical uncertainties by including them in the relationships governing the system. The third step, Uncertainty Analysis, interrogates solutions' performance under deep uncertainties (e.g., climate change effects, population changes, land-use changes, etc.). To do so, we globally sample deeply uncertain exogenous factors that strongly influence the plausible future states of the world (SOW), and evaluate the performance measures of each solution in every SOW. Highly interactive visual analytics are then exploited to understand Pareto approximate tradeoffs as well as the robustness of the component solutions that compose the tradeoffs. The decision maker can choose candidate solutions for further exploration after step 3. In Step 4, we subject the candidate solutions to a scenario discovery process (see section 2.3.2.2 and [78]), which seeks to identify the specific combinations of deeply uncertain exogenous factors that most strongly influence the ability of potentially robust solutions to meet their multi-measure goals. Concepts and methods from RDM contribute to multiple steps of the MORDM framework found in Figure 1. RDM informs the problem formulation, in particular identifying deep uncertainties and contributing the XLRM framework. RDM informs the Uncertainty Analysis, by evaluating the performance of candidate solutions over a wide range of plausible SOW. Finally, RDM's SD identifies the key drivers of those SOW in which the candidate solutions fail to perform well.

The remainder of this chapter proceeds as follows. Section 4.2 explains the rationale of each step in MORDM. The case study used to demonstrate the framework is introduced in Section 4.3. Results are presented in Section 4.4. Sections 4.5 and 4.6 provide general discussion and concluding remarks.

4.2 Methods

4.2.1 Problem Formulation

As illustrated in Figure 4.1, the problem formulation component of the MORDM framework represents decision makers' evolving hypotheses about the most important uncertainties, levers, relationships and measures (XLRM) for their system.

Decision levers quantify an action or policy that can be taken to influence the system. Creative formulation of decision levers is critical for discovering planning alternatives that can dramatically improve system performance. For example, risk-based contracts based on insurance instruments can dramatically improve how water utilities confront hydrologic extremes and growing planning uncertainties [167]. Measures quantify the performance of decision makers' actions with respect to multiple outcomes. A "relationship" represents a mapping from actions to outcomes as quantified using the output measures. The relationship can vary from a simple screening model [73] to an agent-based model that considers multiple actors interacting with each other [9, 168].

When formulating a planning problem in the modern context of environmental change, it is vital to very carefully consider its associated uncertainties. Literature on complex environmental systems often distinguishes natural variability from epistemic uncertainties that can be reduced through further observation or knowledge [169]. Epistemic uncertainty can include errors in estimated probabilities for extreme events [74] and errors associated with model structure [157]. MORDM focuses on epistemic uncertainty from several sources: assumptions for estimated probabilities, model structures, and alternative configurations of decision levers. We build on the concept of "deep" uncertainties that emerge from the suite of risks in a system or their associated probabilities that are not known or cannot be agreed upon [61]. The motivating idea is to characterize our uncertainties with currently available data while also acknowledging the nonstationarity of environmental systems [1]. Nonstationarity can be incorporated in planning by exploring how changing conditions can cause environmental systems to deviate from their expected behavior.

The feedback arrows in Figure 4.1 illustrate the importance of learning and feedbacks across all of the steps of the MORDM framework. Treating problem formulation as a learning process supports the idea that decision support should improve stakeholders' conceptual understanding of complex environmental systems [67,170]. For example, after initial trials of optimization, some measures may be removed and others added [32,55]. The decision makers may also condition their decision on some measures that are not considered during the optimization [117].

4.2.2 Generating Alternatives Using MOEAs

Many objective problem formulations aim to show decision makers critical tradeoffs between their performance measures [48, 56, 171]. MOEAs provide an effective way to discover such tradeoff solutions for complex environmental systems. Their population-based search yields approximations to the Pareto optimal front in a single algorithm run(see [23, 46, 49, 54, 172] for reviews).

For MOEAs to successfully attain high quality approximations to the Pareto frontier, they must converge to a diverse group of solutions that cover the full extent of an application's tradeoffs. The concept of convergence is an important measure of MOEA performance for a problem. It measures how close the MOEA's approximation set has come to the theoretical Pareto optimal front or a best-known approximation to the front. Diversity maintenance is the second critical component to successful MOEA search. It emphasizes that a MOEA must find points "well-spread" across the entire Pareto front. Although early MOEAs had difficulties in maintaining convergence and diversity for challenging problems [54], several modern MOEAs have theoretic proofs of convergence and diversity [54,87,173,174]. Recent diagnostic assessments of top-performing MOEAs on severely challenging many objective test problems [86] and on many objective water resources applications [54] emphasize that recently introduced auto-adaptive MOEAs are effective, efficient, and easier-to-use. These studies highlight that top performing MOEAs can maintain both convergent and diverse search for problems with up to 10-objectives that have a broad range of properties (i.e., nonlinearity, non-convexity, discreteness, stochasticity, and severe constraints).

4.2.3 Uncertainty Analysis

The previous step results in a set of Pareto approximate solutions contingent on the best-estimate values and probability distributions for input parameters to the simulation model. This step explores how the solutions in this set perform when these best-estimate assumptions are relaxed. The MORDM uncertainty analysis interrogates each solution in the Pareto approximate set with a large number of alternative SOWs to see how the solution performs under a range of assumptions regarding the exogenous factors. Performing the analysis on all points shows the

analyst which regions of the space are robust, which may be very different than the performance attained in the baseline SOW. The uncertainty analysis ultimately helps the decision maker choose a formulation and one or more constituent solutions that have acceptable (Pareto satisficing) performance across a wide range of plausible future scenarios.

There are three important design considerations for uncertainty analysis in MORDM. The first is how to sample deeply uncertain exogenous factors (e.g., future climate conditions, population growth, market pricing, etc.) to create an uncertainty ensemble[3]. Our demonstration in the current chapter first sets the upper and lower ranges of the deeply uncertain dimensions and then uses Latin Hypercube Sampling (LHS, [175])[4] to create an ensemble of possible values for the uncertainties. Each ensemble member, or SOW, represents a set of values for each exogenous factor whose impact on future conditions is being explored[5]. Further extensions of this approach could also use multiple simulation model structures or different governing equations across each SOW. A second consideration is how many measures to use for calculating robustness. Some previous RDM work has considered robustness over multiple attributes [31, 47, 176], but these studies have not systematically identified the full range of strategies that might satisfy different preferences among the objectives. In contrast, MORDM provides such a capability. Finally, the analyst should determine appropriate statistical thresholds for defining robustness. For its example application, this chapter selects potentially robust candidate solutions based on each solution's performance in the most extreme SOW samples. This does not mean that we believe the most extreme ensemble members will actually occur. Rather, we consider these extreme cases in order to understand the conditions where a proposed strategy may fail to meet its performance goals.

[3]Exogenous factors are those that can be considered external to the modeled system. Note that whether a factor is exogenous or endogenous depends on the specific model. For example, if a population growth depends on a factor that is explicitly included in a simulation (such as resource availability), the population growth would be endogenous. If no population model is included in a simulation and decision makers assume a rate, the assumed rate is considered exogenous.

[4]LHS was chosen as a sampling method following prior work [78]. LHS breaks the sampling space into evenly-spaced grid boxes and chooses a random value within each grid box. This method of striation can more evenly sample a space relative to uniform random sampling [97].

[5]Other factors not included in the SOW could also affect the policy's performance. Subsequent use of SOW in this paper refer only to changes in the dimensions chosen for the model, similar to a *ceteris paribus* assumption in economics.

4.2.4 Scenario Discovery

Scenario Discovery (SD) uses statistical cluster analysis on databases of simulation model runs to identify simple, easy-to-understand descriptions of the combinations of uncertain model inputs that best predict the future states of the world where the strategies identified as robust in the uncertainty analysis nonetheless perform poorly [107]. The method was fully introduced with an illustrative example in section 2.3.2.2, and a brief summary will be provided here. First, thresholds are set for each of the performance measures, according to plausible stakeholder preferences. Ensemble members that violate the thresholds are termed vulnerable. The Patient Rule Induction Method (PRIM) automatically calculates "scenario boxes" that provide simple descriptions of ranges of exogenous parameter values over which a candidate solution performs poorly. These boxes are expressed in the form $B = \{a \leq x_j \leq b, j \in L\}$. In other words, a subset of dimensions of uncertainty x_j are constrained to be between lower and upper ranges a and b.

PRIM uses an interactive process to suggest alternative candidate boxes from which the users choose, using statistical measures of the quality of boxes, or how well they can be interpreted by the user while still capturing the vulnerable set. Scenario Discovery also suggests which uncertain parameters are less important in describing the vulnerable cases. Some RDM analyses use such scenarios to help identify policies that might ameliorate these vulnerabilities and then present tradeoff curves to help participants decide whether these new policies are worth adopting [61, 79, 176, 177]. In these analyses, the potentially ameliorating policies are chosen from an already-existing set of options or handcrafted by analysts with input from stakeholders. A MORDM analysis might significantly improve the choice set by reapplying its evolutionary algorithms specifically to search for policies that ameliorate the vulnerabilities identified by SD with minimal tradeoffs with other goals. However this potentially promising but challenging step is left for future work and discussed further in chapter 7.

4.2.5 Interactive Visual Analytics

Since the MORDM framework represents a form of constructive *a posteriori* decision support [9, 29, 70, 170], a large amount of information is available to inform

decision makers' choices. A successful system therefore requires a method for processing and viewing the data that makes it tractable for decision makers to interpret and analyze. Interactive visual analytics, first introduced in section 2.3.3 (see also [110–113]), uses multiple linked views of decision-relevant information to facilitate the data processing procedure. In our proposed MORDM framework, visual analytics is critical for facilitating insights and the articulation of preferences (i.e., MOEA-generated tradeoffs, PRIM assessment of uncertainties, etc.).

As shown in the center of figure 4.1, visual analytics can enhance all four MORDM steps. The first major use of interactive visual analytics for MORDM is visualizations of decision levers, performance measures, and metrics of robustness for each solution. These components can be combined with other variables that were not included in optimization formulations [117] but may ultimately influence a decision maker's analysis[6]. Several plotting types can be used, with the main purpose of comparing solution properties and ultimately choosing a robust solution. Glyph plots use 3-dimensional cartesian coordinates, as well as points' size, color, rotation, and transparency to show trends in up to 7 dimensions simultaneously [60]. Prior work has used these visualizations to compare competing problem formulations [32, 56] and tradeoffs through time [178]. In addition to glyphs, parallel coordinate plots [51, 116, 179] can show many dimensions at once, but only allow pairwise comparison between subsequent dimensions. Both plotting types are interactive, with the analyst changing which variables appear on each axis in real-time. The analyst can also "brush" solutions that meet user-defined criteria [180], reducing the number of plotted solutions at any one time.

Interactive visual analytics are also used in uncertainty analysis and scenario discovery. The analyst can visualize how each solution responds to the sampled uncertainty ensemble using the multiple performance measures. Within SD, visuals for coverage and density are integral in allowing the user to guide PRIM in creating scenarios. This process can inform modifications to the performance thresholds, or perhaps even demonstrate that additional dimensions of uncertainty should be included in subsequent analyses.

[6]Decision makers often see decision variables, objectives, and other variables as interchangeable [29]. In other words, they would like to see how the decision space, performance space, and other factors interact concurrently.

4.3 LRGV Case Study Implementation

4.3.1 Motivation

Although climate change and urbanization pose serious threats to water management, the rising cost of building new infrastructure motivates non-structural approaches such as water marketing [118] for ensuring a sufficient quantity of supply. In a water market, agents transfer quantities of water, either across different regions [141] or different user sectors within the same water system [142]. The presence of a water market does not imply that the city can easily ascertain the best combination of traditional and market-based instruments for their supply. Adding a new portfolio instrument (spot leasing) to a strategy that uses traditional supply and an options contract can reduce costs and surplus water [56]. The significant performance improvement achieved by adding new planning instruments underscores an important uncertainty in defining the planning process itself. In a water marketing system, alternative system assumptions and formulations can dramatically influence the discovery and exploitation of critical planning tradeoffs [32].

These issues motivate water marketing as a challenging planning context that can serve as an excellent test case for demonstrating the MORDM framework. Our case study develops portfolio-based strategies for a hypothetical single city in the Lower Rio Grande Valley (LRGV) of Texas [32, 56, 122, 123]. The city uses water market transfers from agriculture to municipal use to augment their traditional reservoir-based supply. The LRGV case study assumes that municipal use has the same priority as water for irrigation, with the goal of determining reliable portfolio strategies that can increase the city's reliability while ensuring sufficient water for other regional uses [5].

Our use of the LRGV test case in this study builds on the *de Novo* planning results and assumptions from chapter 3 (see also [32]). The *de Novo* planning framework seeks to formalize the discovery and evaluation of alternative formulations for challenging planning problems. *De Novo* planning does this by starting with an *a priori* selection of decision variables, objectives, and constraints, and subjecting this formulation to global sensitivity analysis. We then search a suite of alternative problem formulations, using evolutionary multi-objective optimization to develop alternative tradeoffs for the LRGV. The study ultimately sought

to identify which of these alternative formulations were "non-dominated", versus the more traditional focus on non-dominated solutions in a single formulation. In our prior work, our choice of a preferred model case and candidate solutions, though, was strongly dependent on the expected value calculations used to assess performance tradeoffs. These expectations assume that our LRGV Monte Carlo simulation using historical data fully captures the breadth of uncertain futures that occur for the system. Unfortunately, these assumptions could very plausibly be violated in future planning periods. For example, [120] highlights concerns over reduced inflow from the Mexican tributaries feeding the LRGV's reservoir system. Climate change is projected to cause increasing temperatures, which could increase reservoir evaporation and modify streamflows [181]. Fluctuations in agricultural commodity prices also influence irrigators' willingness to participate in a water market [182], possibly increasing market prices. Each of these challenges illustrates that the "expected" performance assessment for the LRGV could be negatively impacted by deeply uncertain risks for unexpected system changes and shifts in estimated likelihoods.

Our goal is to demonstrate how the MORDM framework addresses the following questions: Was our choice of a preferred model case and its component solutions critically biased by our use of historical data to assess the expected behavior for the LRGV system? If so, what are the controlling assumptions or conditions we should consider to improve water planning for the LRGV test case? The following sections discuss our computational experiment for each of the steps shown in Figure 4.1: problem formulation, alternatives generation, uncertainty analysis, and scenario discovery.

4.3.2 Problem Formulation

Our problem formulation is constructed using the four components of the XLRM framework: uncertainties, levers, relationships, and measures. This section focuses on the decision levers, performance measures, and quantitative relationship for the LRGV, with the uncertainties discussed in Section 4.3.4.

Decision levers are used to construct the city's portfolio of three water supply instruments: permanent rights, spot leases, and options. Permanent rights are a

non-market based instrument in which the city is allocated a percentage of reservoir inflow in each month, using a ratio of the volume of the city's rights (the decision lever, N_R) to the total volume of regional rights. Spot leases can be acquired at any month in the year with a variable price. An adaptive options contract reduces lease price volatility by guaranteeing a fixed price for water acquisitions made later in the year. The options contracts are controlled by up to three variables: a high and low volume ($N_{O_{\text{low}}}$, $N_{O_{\text{high}}}$), with a threshold (ξ) that decides which contract to activate depending on the available water supply [123]. Acquisitions of water from both leases and options are controlled by anticipatory thresholds that relate expected supply to demand. The thresholds control "when" (α) and "how much" (β) water the city must acquire when using the market. Since 85% of the water in the region is used for agricultural use, we assume that there will always be enough water to meet the transfer requests specified by the alpha and beta variables [122].

Recall that this chapter will test the *de Novo* results presented in chapter 3. In this context, we adapt our treatment of decision levers from that study (see section 3.4.2). To choose an appropriate configuration of decision levers, Sobol' variance decomposition [100] was used to analyze the effect of each decision lever on the LRGV's performance measures in the initial many objective formulation [56]. The results suggested that only the volume of permanent rights and the alpha strategy variables significantly influenced performance, and the performance was less sensitive to the adaptive options contract variables and beta. The fact that a small subset of variables had this sensitivity performance motivated [32] to explore the question: "what is the minimum level of formulation complexity that is justified and effective for the LRGV test case?" The four candidate formulations of decision levers were summarized in Table 3.4. Case I uses the volume of permanent rights, a single-volume non-adaptive options contract, and one variable to determine both when to go to the market and how much water to acquire. Case II varies that threshold decision by the time of the year, and case III separates the when and how much decision by explicitly searching for separate alpha and beta values. Case IV adds in the adaptive options contract to meet the full complexity of the *a priori* problem formulation from the initial study [56].

The "relationship" for our problem formulation uses a 10-year expected performance Monte Carlo simulation and an extreme drought, both with a monthly

timestep. For further details about the simulation model and its implementation, the reader is encouraged to consult section 2.4.1 and a series of prior studies [32,56,122,123]. The simulation model samples historical lease pricing, demand, and reservoir inflows to test how the supply portfolios would perform under a single best estimate of the LRGV's uncertainties. We also use an extreme drought scenario, which combines the driest year in the historical record with the maximum demands from the discrete demand distributions used in the simulation. The city begins with a volume of water controlled by initial rights, and must satisfy its demands using its portfolio planning strategy. The drought scenario provides a challenging test for the portfolios, since the drought conditions maximize the difference between the expected supply and demand used in alpha/beta calculations and the city's actual demand [56].

Each configuration of decision levers in Table 3.4 is tested using a many objective formulation of measures (objectives and constraints) as shown in the following equations.

$$\begin{aligned} \mathbf{F}(\mathbf{l}_k) &= (f_{\text{10 yr. cost}}, f_{\text{10 yr. surplus}}, f_{\text{10 yr. crit. rel}}, \\ &\quad f_{\text{10 yr dropped}}, f_{\text{10 yr num. leases}}, f_{\text{dr. trans. cost}}) \qquad (4.1) \\ &\quad \forall \mathbf{l}_k \in \Omega \end{aligned}$$

$$\text{Subject to : } c_{\text{10 yr. rel}} : \quad f_{\text{10 yr. rel}} \geq 0.98 \qquad (4.2)$$

$$c_{\text{10 yr. costvar}} : \quad f_{\text{10 yr. costvar}} \leq 1.1 \qquad (4.3)$$

$$c_{\text{10 yr. crit. rel}} : \quad f_{\text{10 yr. crit. rel}} \geq 0.99 \qquad (4.4)$$

$$c_{\text{dr. rel.}} : \quad f_{\text{dr. rel.}} = 1.00 \qquad (4.5)$$

In the equations, "10 yr." refers to the 10-year Monte Carlo simulation and "dr." refers to the drought. The vector $\mathbf{l}_k$ denotes the levers (decision variables), with the subscript k indicating that the set of objectives and constraints is used with each model case from Table 3.4. Recall that the term "measure" indicates a quantification of system performance, which is used as an objective or constraint for optimization. This section briefly reviews each of the measures from equations 4.1 through 4.5, but for the full definitions please consult section 3.2.2. These

measures are also used later in this study to define the robustness of selected solutions. They are broken into three groups, depending on the type of attribute they are meant to quantify.

The first group of measures quantifies the cost of the supply portfolio. $f_{\text{10 yr. cost}}$ is the sum of the average annual costs for rights, options, and leases calculated using the 10-year simulation. The high-tail of the cost distribution is measured by cost variability, $c_{\text{10 yr. costvar}}$, defined as the ratio of the mean of the costs falling above the 95th percentile divided by the average annual cost. In the drought, $f_{\text{dr. trans. cost}}$ quantifies the cost of market transactions.

The second group of measures uses various metrics of reliability [40] to quantify the performance of each portfolio. Reliability, $f_{\text{10 yr. rel}}$ quantifies the likelihood that the city will meet its required demand in the 10-year Monte Carlo simulation. Drought Reliability, $f_{\text{dr. rel}}$, is the same calculation, performed within the drought scenario. The Critical Reliability measure $f_{\text{10 yr. crit. rel}}$ treats "success" of each month differently than the basic Reliability; the city must meet at least 60% of sampled demand. Critical Reliability therefore measures the likelihood of very large failure events that would be difficult to mitigate using demand management or other techniques.

The third group considers measures associated with market use and the efficiency of each portfolio. Acquisitions of exercised options and leases are fully controlled by the supply/demand threshold. Therefore, some portfolios could specify a large number of leases or options and incur high transactions costs. Number of Leases ($f_{\text{10 yr num. leases}}$), minimizes the expected number of leases obtained in each portfolio. Transfers are modeled to expire after 12 months of nonuse, so the Dropped Transfers measure quantifies the volume of water dropped from nonuse ($f_{\text{10 yr dropped}}$). Finally, the Surplus Water measure ($f_{\text{10 yr surplus}}$) determines the average amount of water carried over from year to year by the portfolio. This measure is minimized as a proxy for other regional users; portfolios that carry too much surplus use water that could be put to other purposes.

In summary, the decision levers are the volumetric and strategy decisions shown in Table 3.4. The relationship is a Monte Carlo simulation and extreme drought scenario. The cost, reliability, and market use metrics defined in this section are the performance measures. Finally, the uncertainties in this problem formulation

include both classical and deep uncertainties. The classical uncertainties are assumed distributions of lease prices, inflows, losses, reservoir variation, and demand, and are incorporated into the model using Monte Carlo simulations. We then examine the extent to which solutions identified by the MOEA step are robust across a variety of deep uncertainties, which we discuss in more detail in Section 4.3.4.

4.3.3 Multi-objective Evolutionary Algorithm

The ε-NSGAII [85, 88, 89] was chosen to generate alternatives for this study. This algorithm extends the original NSGAII [84] by adding epsilon dominance archiving [87] and adaptive population sizing [90] to change the size of the population commensurate with problem difficulty. At present, ε-NSGAII is a top performing search tool, as tested by recent diagnostic comparisons of modern MOEAs [54, 86]. Other top-performing MOEAs could be used in subsequent MORDM applications, but the MOEA used should be selected carefully to ensure that it can reliably solve challenging problems (see Section 4.2.2). The MOEA was used to generate alternatives for each of the four problem cases in Table 3.4. For more details, please consult chapter 3.

4.3.4 Uncertainty Sampling

The previous step identified solutions on the Pareto approximate surface given the base case assumptions. This uncertainty sampling step aims to test how significantly the performance of each of these solutions varies if the base case assumptions turn out to be wrong. This step focuses on two types of deeply uncertain parameters: 1) those that define the base case probability distributions used in the Monte Carlo simulation model and 2) other parameters that the Monte Carlo simulation treats as fixed values. Table 4.1 lists parameters of the first type and Table 4.2 lists those of the second type.

We use a LHS sample over both types of uncertainties to generate 10,000 alternative SOW's against which to test the solutions (the base case represents one SOW). For the real-valued, non-probabilistic parameters (e.g. the initial reservoir level) each SOW has a value between the lower and upper bounds listed in Table 4.2. For the parameters defining the probability distributions (e.g. the distribution

of lease prices), we use "scaling factors" to renormalize the tails of the distribution as described in Section 4.3.4.1 below. Each SOW has a value for each scaling factor between the lower and upper bounds listed in Table 4.1.

We then run the Monte Carlo simulation model for each of the 10,000 SOW's and record its performance according to each of the measures. It is useful to compare the different purposes of the LHS sample used to generate the 10,000 SOW's and the Monte Carlo samples used by the simulation model to generate results for each individual SOW. The Monte Carlo samples weight points according to their estimated likelihood, so the model can sum over the sample to calculate in each SOW the mean values and variances for the model outputs. The LHS sample is a quasi-random design that weights points equally. It is used here to explore the performance of strategies over a wide range of plausible cases. There is no claim that all SOW in the sample are equally likely. Rather the sample aims to provide data that allows decision makers to understand which solutions are more or less sensitive to deviations from their base case assumptions and, in the scenario discovery step, what particular combinations of uncertainties would cause particular solutions to perform poorly. It should be noted that the specific LHS sample used here is meant only as one pragmatic example of how to explore deeply uncertain factors or probabilistic assumptions that influence the Pareto satisficing behavior of tradeoff solutions. This step offers a rich opportunity for future research to explore alternative schemes for exploring deep uncertainties.

4.3.4.1 Scaling Factors

The scaling methodology for our study is adapted from [31], in which probability distributions are renormalized to explore the consequences of potential misestimation of the likelihood of extreme events in the assumed baseline distribution. Here we renormalize the weight in the highest or lowest 25% of the distribution and use an integer scaling factor between 1 and 10 to control the reweighting. We re-run the simulation, where a non-uniform sampling procedure is used such that the highest or lowest 25% of the data becomes 1 to 10 times likelier, depending on the scaling factor. Each scaling factor is treated as an integer in the LHS. Note that subsequent MORDM analyses can use alternative scaling or distributional sampling methodologies. The key issue is to quantitatively explore the impacts of

Table 4.1. Scaling Factors

Input Variable	Lower Bound	Upper Bound
Low Inflows	1	10
High Losses	1	10
High Demands	1	10
High Lease Prices	1	10
Losses in Reservoir Storage	1	10

alternative likelihood assumptions on measures of system performance.

Table 4.1 presents the data scaling factors, and Figure 4.2 illustrates how these factors modify the cumulative distribution function (CDF) of the input data. To demonstrate how each SOW has different scaling factors across the data types, the figure shows example results attained using the 2, 4, 6, and 10 scaling factors, with the thick blue line indicating the original baseline data. Additionally, the figure shows how the data changes across two representative months: January (Figures 4.2a-4.2c and 4.2g-4.2i) and August (Figures 4.2d-4.2f and 4.2j-4.2l).

Lease pricing distributions are given in Figures 4.2a-4.2b and 4.2d-4.2e. Leases are purchased by sending a document to the Watermaster's office. The office arranges a one-time transfer of water from the irrigator to the municipality, at a variable price, from the main reservoir system. Lease pricing is modeled as a random variable sampled from an empirical monthly distribution. A prior analysis [122] showed that there are two distributions of lease prices, depending on whether or not the reservoir volume is below 1.76 $\times 10^{12}$ cubic meters. Changing agricultural commodity prices could impact willingness to participate in the market and therefore modify prices [182]. Additionally, the simulation assumes that any requested transfer can be fulfilled entirely, so in reality there could be volume limitations that would make prices increase. Finally, empirical price distributions developed for the LRGV test case are based on a limited number of data points, so the distribution could have a slightly different shape than what was observed. Lease prices are scaled to emphasize the highest 25% of the data distributions;

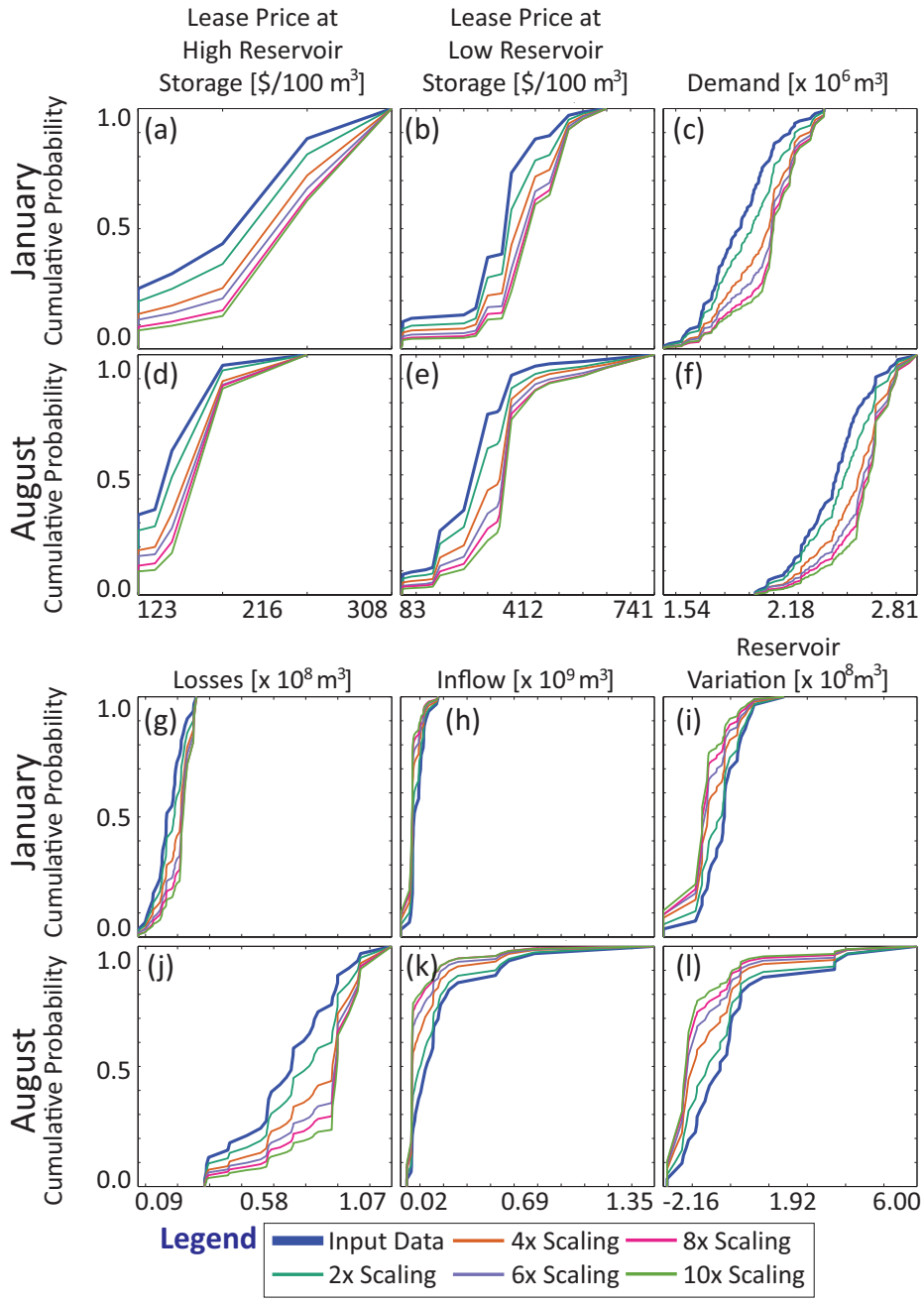

Figure 4.2. Comparison of cumulative distribution functions under different scaling factors. Figures 4.2a-4.2c and 4.2g-4.2i show the scaled data for January, whereas Figures 4.2d-4.2f and 4.2j-4.2l show the data for August. The thick line shows the baseline data, with the colored lines illustrating scaling factors between 2 and 10.

hereafter the scaled lease prices are termed "High Lease Prices"[7].

Demand distributions are presented in Figures 4.2c and 4.2f. Uncertain demand in the simulation is sampled based on Gaussian distributions with parameters estimated using historical data. The entire demand distribution also grows exponentially subject to a demand growth rate, as discussed below. Characklis et al. [5] suggested that regional demands could increase by a factor of three from 1990-2050. The assumption of normality in modeling demand could also be incorrect, motivating exploration of distributions that have larger "tails" with a larger proportion of likelihood in higher values. Demand distributions are scaled to emphasize the highest 25% of demand, termed "High Demands" hereafter.

Losses and inflows are presented in Figures 4.2g-4.2h and 4.2j-4.2k. Water allocated to permanent rights is calculated from reservoir inflows *pro rata* based on the volume of permanent rights compared to the total amount of regional rights. For example, if the modeled city has 50 units of rights, and all regional rights holders hold 100 units of rights, our modeled city would obtain half of all reservoir inflows in every month. The water available for allocation is the difference between a sampled inflow volume and a volume of losses, both estimated from empirical historical distributions. Inflows and losses are deeply uncertain since in future planning periods, climate change could shift the distribution of values toward higher losses and lower inflows [181]. Inflows are scaled to emphasize the lowest 25% and termed "Low Inflows", whereas losses are scaled to emphasize the highest 25% and termed "High Losses".

The reservoir variation is presented in Figures 4.2i and 4.2l. Reservoir variation refers to the aggregate change in reservoir storage for the water source; the variable can be positive for gains and negative for losses. The reservoir volume is monitored by the International Boundary and Water Commission (IBWC) that is charged with maintaining a 1944 treaty between the United States and Mexico. Recent concerns about lower inflows coming from Mexican tributaries [120] could cause the reservoir supply to decrease relative to historical conditions. Increased evaporation from climate change could also play a role. Therefore, the reservoir variation scaling focuses on the lowest 25% of the data. This is termed "Losses in Reservoir Storage"

[7]Two distributions of prices are still maintained, depending on reservoir storage (see Figure 4.2). Therefore, under the sampled SOW, the simulation determines which lease price distribution to use in a given month, and applies the same scaling factor for both distributions.

Table 4.2. Sampled Model Parameters

Parameter	Lower Bound	Upper Bound
Initial Rights	0.0	0.4
Demand Growth Rate(%)	1.1	2.3
Initial Reservoir Level [10^6 m^3]	987	2,714

since the scaled reservoir variation tends to cause more losses than in the baseline historical dataset.

4.3.4.2 Scalar Model Parameters

A second group of deeply uncertain variables represents point values of model parameters. The first sampled parameter is the initial water rights. This is an initial condition that represents the amount of water available to the city's supply in the first month of the simulation. The initial condition relates to the water carried over from year to year by the portfolio, and has been shown to be important in determining portfolios' performance [32]. The lower bound of 0.0 represents no water in the beginning of the simulation, such as a situation in which a supply failure occurs in the month before the simulation begins. The upper bound, 0.4, approximates a situation in which the city starts their supply with a volume equal to 40% of their permanent rights volume. Under extreme supply conditions of higher demands coupled with lower inflows, the city may not be able to maintain their preferred amount of supply from year to year. Thus this variable is treated as deeply uncertain, and sampled within the range 0.0 to 0.4.

The second deeply uncertain model parameter is the demand growth percentage. Based on an analysis of USGS water supply data and projections from the Texas Water Development Board presented in [32], we posited that the LRGV's demands are likely to grow with a rate between 1.1% and 2.3%. Since future demand growth will depend on a variety of factors that are difficult to predict, including future population growth, housing stock and weather patterns, this variable is sampled between 1.1% and 2.3% in the ensemble.

The initial reservoir volume is the final sampled model parameter. The main effect of the reservoir level in the simulation is on lease pricing, as mentioned earlier.

Table 4.3. Threshold Sets

Name	Measures	Thresholds
Market Use	Number of Leases	> 90th percentile
Reliability	Reliability, Critical Reliability Drought Reliability	< 10th percentile in any
Cost	Cost, Cost Variability	> 90th percentile in any

There is also a "dead storage zone" modeled in the mass balance simulation, below which no water is allocated. The range used in sampling, 987 million cubic meters to 2,714 million cubic meters, is adapted from prior work [122].

4.3.4.3 Quantifying Robustness

MORDM requires both a method to sample uncertainty as well as a method to quantify robustness for each tradeoff solution. For each tradeoff solution, there exists a distribution of performance for each output measure across all SOWs in the LHS ensemble. Our treatment of robustness focuses on the most extreme SOWs in the ensemble. We assume that if a solution performs well in these worst-case SOWs, it will also have robust performance for many deeply uncertain trajectories of the future. Specifically, we calculate a deviation metric, percent deviation (d_i):

$$d_i = \begin{cases} \frac{f_{i\,90} - f_{i\,base}}{f_{i\,base}} & \text{if } i \text{ is minimized} \\ \frac{f_{i\,base} - f_{i\,10}}{f_{i\,base}} & \text{if } i \text{ is maximized} \end{cases} \tag{4.6}$$

where i is the measure of interest, the 90 or 10 subscript indicates the 90^{th} or 10^{th} percentile in the uncertainty ensemble, and "base" indicates the measure value from the baseline simulation. In other words, the percent deviation calculation shows the magnitude that the most extreme 10% of the ensemble members deviates from the expected measure value in the baseline state of the world. We identify promising candidate solutions as those that have low percentage deviation across many output measures of interest.

In this study, our initial exploration of how each solution performed across different SOW suggested that we should focus on a subset of the 9 performance measures that were initially identified. We found that many of the solutions

actually improved their performance across 3 performance measures - surplus, dropped transfers, and drought transfers costs - in many SOW. Since the goal of the MORDM exercise is to find SOW in which the candidate solutions are vulnerable to plausible futures, and to suggest ways to mitigate those vulnerabilities, we focused on the remaining 6 measures (cost, cost variability, reliability, critical reliability, drought reliability, and number of leases).

4.3.5 Scenario Discovery

Our exploration of the "percent deviation" outcomes suggested a potential candidate solution that performed relatively well across many of the measures of interest. We then applied the SD process to this candidate solution in order to identify future SOW (values of uncertainties) under which it would perform poorly. To identify poor performance, we defined a set of thresholds for each group of performance measures: cost, reliability, and market use. Table 4.3 defines the threshold sets used to delineate SOWs that cause poor solution performance. In general, the candidate solution was considered to perform poorly in SOW in which any of the measures in the threshold set fell into the most extreme 10% of values.

4.4 Results

4.4.1 Generating Alternatives

Figure 4.3 presents multi-objective tradeoffs generated using a MOEA for each of the four model cases in Table 3.4. These results, adapted from [32], compare the model cases' performance with respect to multiple output measures. In this study, our purpose is two-fold: (1) evaluate which formulation yields solutions that are robust across a broad suite of SOW and (2) assess if our prior choice of preferred model case III was appropriate given the LRGV's deep uncertainties. Figure 4.3 is a glyph plot in which each portfolio solution is represented by a cone. The cone's coordinates represent the cost ($f_{10\text{ yr. cost.}}$), number of leases ($f_{10\text{ yr. num. leases}}$), and surplus water ($f_{10\text{ yr. surplus}}$) measures. Additionally, the orientation of the cones shows the dropped transfers ($f_{10\text{ yr. dropped}}$) and the cones' size shows the critical reliability ($f_{10\text{ yr. crit. rel}}$). On each axis, arrows indicate the direction of increasing

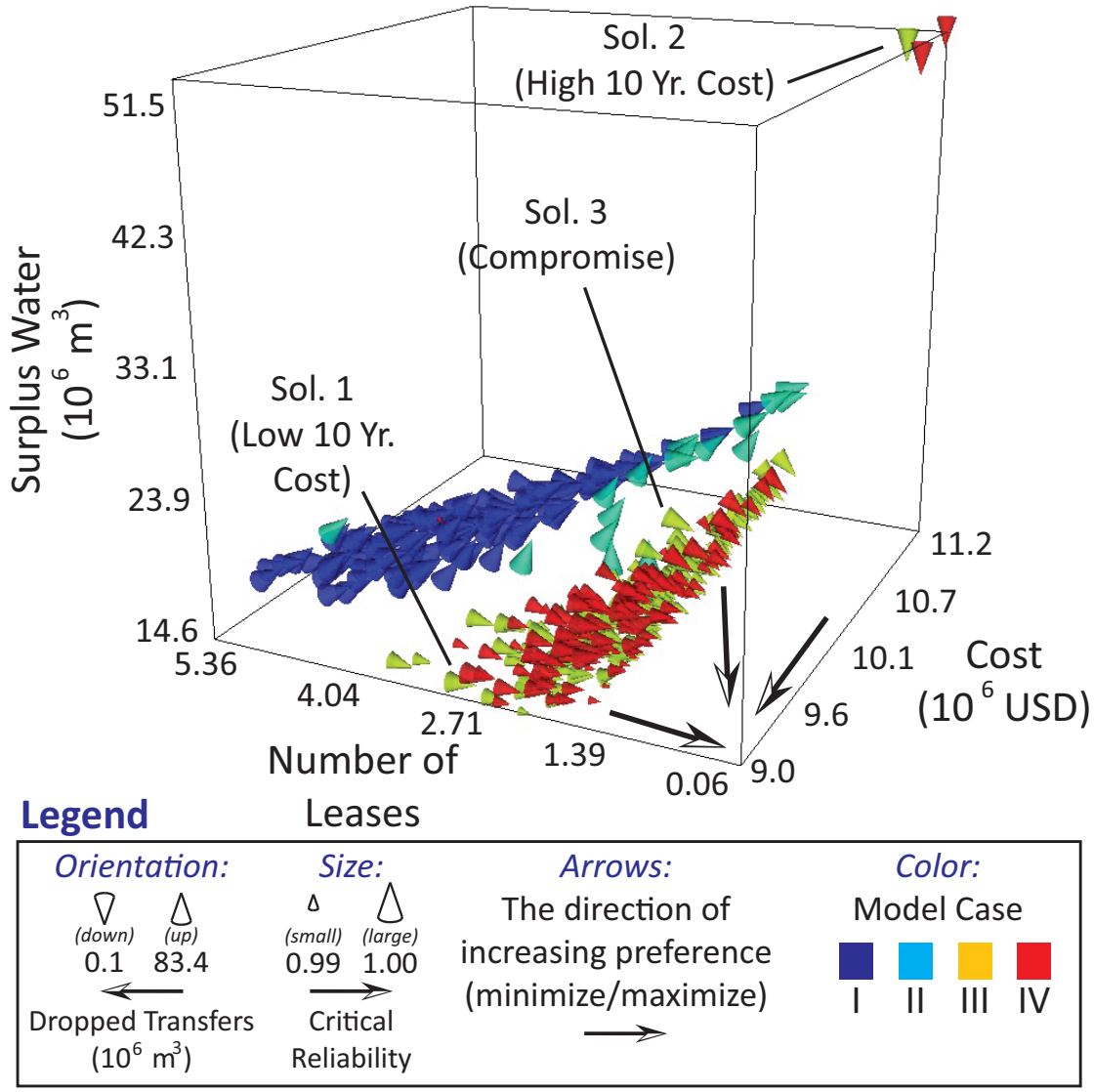

Figure 4.3. Non-dominated tradeoffs generated in prior work [32]. Each cone is an individual water portfolio solution, with spatial axes plotting solutions' cost, number of leases, and surplus water. The annotated solutions from model case III were chosen because of their performance in the plotted performance measures.

preference (i.e., whether a measure is minimized or maximized). Overall, Figure 4.3 shows that model cases III and IV have better expected performance with respect to the three measures plotted on the spatial axes, providing good performance with fewer leases and lower surplus water at every level of cost. Cases III and IV could be argued to be the "non-dominated" problem formulations based on their expected performance attained using the Monte Carlo simulation.

This result motivated choice of model case III in the prior study [32] and subsequent exploration of three selected solutions shown in Figure 4.3. Case III uses the distinct alpha variable to determine "when" to go to the market and beta to determine "how much" to buy. Additionally, the case III problem formulation separates these variables between January-April and May-December, but unlike case IV it does not use the adaptive options contract. In general, use of distinct alpha and beta values allows the city to "tune" market acquisitions to the input data and to buy only as many transfers as needed to meet reliability under the modeled conditions [122]. If the city typically has enough surplus water to avoid market transactions from January through April, for example, water managers can choose a portfolio that has lower values of alpha and beta in those months. These trends led to excellent performance with respect to their expected value measures (especially lowering the dropped transfers and surplus water due to efficient market use). However, our choice of the selected solutions is based on the single baseline SOW used to assess the tradeoffs. In the remainder of our results, the MORDM framework is employed to rigorously explore our choice of a formulation and solutions of interest for the LRGV test case. MORDM explores how changing assumptions about this historical baseline SOW can affect the performance of these solutions and possibly motivate selection of one of the alternative problem formulation's strategies for managing the city's water supply.

4.4.2 Percent Deviation of Performance Measures

MORDM uncertainty analysis globally samples many different combinations of exogenous assumptions (or alternate future conditions). Each set of exogenous parameters yields a new Monte Carlo simulation for the LRGV system that deviates from its original Monte Carlo simulation based on the historical SOW. Specifically, each tradeoff solution is tested with an ensemble of 10,000 SOWs as described in Tables 4.1 and 4.2. The percent deviation metric is used to interpret each solution's performance in the uncertainty ensemble, which calculates the difference between the most extreme 10% of the values in the uncertainty ensemble and the value in the baseline SOW. Figure 4.4 is an initial visualization of the percent deviation results for each solution across several measures. The figure is in the

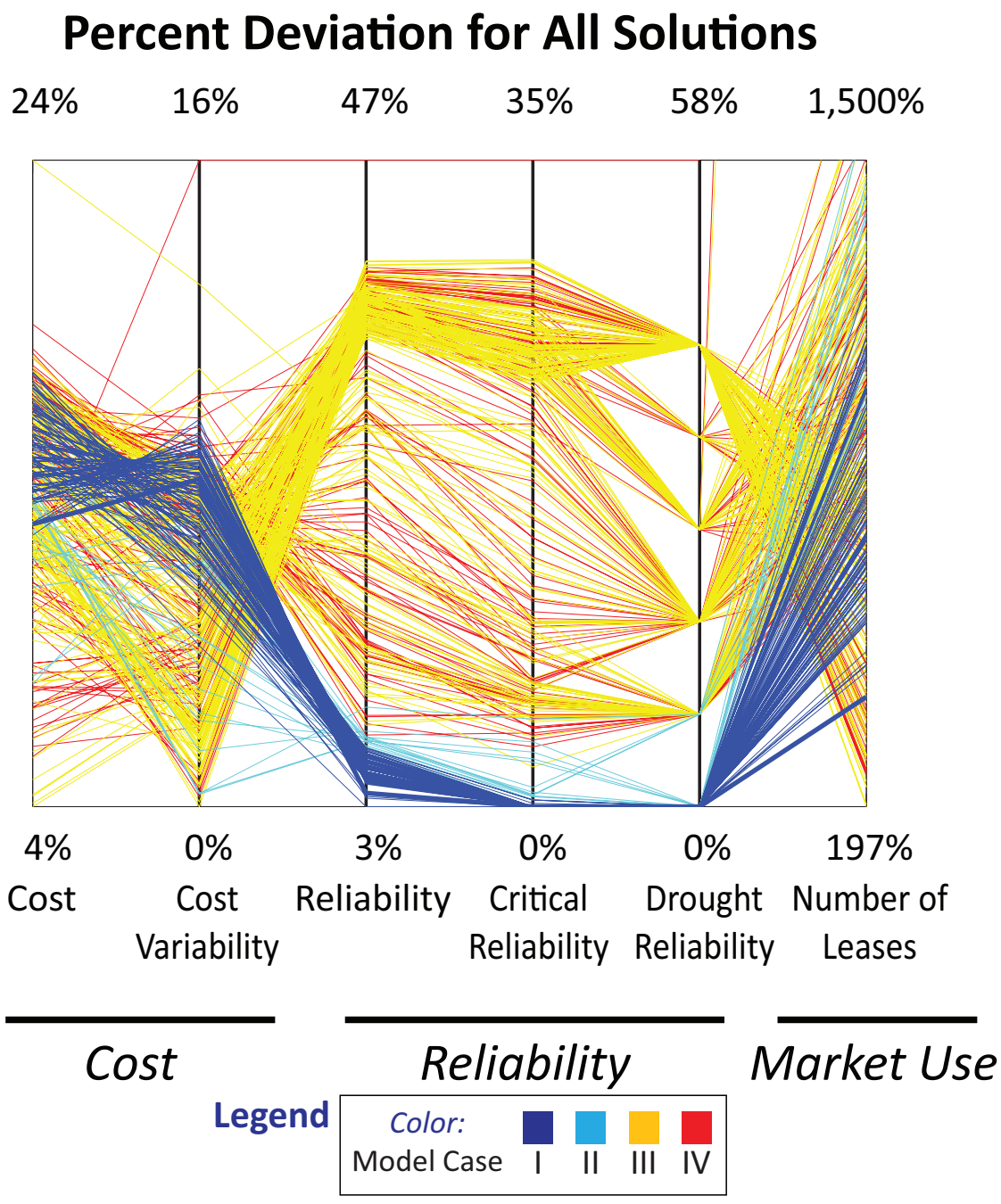

Figure 4.4. Parallel coordinate plot showing the percent deviation for all solutions, defined as the difference between the most extreme 10% measure value from the uncertainty ensemble compared to the baseline state of the world. Color shows the model case. Measures are grouped into three categories: cost, reliability, and market use.

form of a parallel coordinate plot, with color distinguishing the 4 model cases in the same manner as was used in figure 4.3. Each line in the figure shows a single solution, and the line's vertical position on each axis shows relative values of deviation for each measure, with values closer to the bottom axis representing more robust performance. Figure 4.4 allows us to observe the magnitude of percent deviation for each solution across the three groups of performance measures, and to determine trends across the model cases. Ideal performance in Figure 4.4 would be a horizontal line that intersects all of the vertical axes at zero deviation.

Cases I and II exhibit moderate to high deviation with respect to the cost and number of leases measures but improved performance with respect to reliability when compared to cases III and IV. Cases I and II use a single variable to exercise their options and leases and do not separate the "when" and "how much" decisions (see Table 3.4). Portfolios in these cases would need to maintain a high value for their alpha strategy variable in order to ensure a sufficient supply. The high threshold may lead to reduced performance with respect to the number of leases and surplus water measures in the baseline SOW (see figure 4.3). However, these large acquisitions of water from the market helped portfolios in cases I and II maintain high reliability in the uncertainty ensemble. A decision maker would likely be willing to accept higher costs, though, if it means that supply reliability is maintained under extreme conditions. In summary, the robust reliability performance of cases I and II exhibited by Figure 4.4 is a surprising result that conflicts with our prior choice of case III [32]. This shows that choosing solutions with respect to robustness can lead to markedly different decisions compared to observing performance measure results in a single SOW based on historical likelihoods.

4.4.3 Negotiation of a Robust Solution

Visual analytics within MORDM provides a rich opportunity for decision makers to interact with decision-relevant data for each of the suggested tradeoff solutions. This section demonstrates two specific techniques for interacting directly with the data visualizations. First, the decision makers can change the ordering of variables on visualization axes to better illustrate conflicts[8]. These plots show how it would

[8]When all axes are arranged such that the preferred direction is toward one half of the figure (i.e., at the bottom), crossing lines indicate that one cannot achieve good performance in a given

be very difficult to ascertain *a priori* what conflicts exist between measures [45] and which solutions would be robust to a wide array of plausible futures [9]. Moreover, decision makers can interact directly with solutions shown in the plot, using "brushing" to impose limits on the plotted data to reflect decision maker preferences [180]. Brushing yields a reduced group of diverse alternatives, allowing the decision maker to choose between a small number of maximally different alternatives [15]. Figure 4.5 expands our treatment of the percent deviation results to demonstrate these techniques. The measure axes have been rearranged to better illustrate the conflicts between percent deviation in reliability measures and the other performance measures. The figure also reflects brushing to focus on solutions that have lower than 5% deviation in any of the three reliability measures. Our focus in this demonstration is to choose solutions that have excellent performance in meeting reliability; decision makers who have other preferences can express them by using different arrangements of the axes and different brushing criteria. As a result of the brushing, solutions that do not meet the reliability criteria are shown in gray, with the remaining desired solutions shown in a color gradient. Choosing an appropriate color mapping is an interactive process, and the colors should be chosen to clearly illustrate trends in the data. In figure 4.5 the color scale demonstrates the percent deviation in cost of the remaining solutions (from 12% to 15%).

The solutions highlighted by brushing in Figure 4.5 mostly come from cases I or II, and have a range of performance with respect to deviations in cost, number of leases, and cost variability. The blue colored solutions, which have low deviations in cost, also have low deviations in terms of number of leases but exhibit higher deviations in cost variability. These solutions typically perform well across all of the cost and reliability measures, so we selected solution 4 since it has the lowest deviation with respect to number of leases in this set. Table 4.4 compares the decision levers of solution 4, termed the "robust" solution, with the previously selected solutions 1-3 from case III. Solution 4's strategy variables are each 1.69, which means that its expected supply must be 1.69 times its expected demand in all months[9]. Though this may lead to the portfolio acquiring a higher number of

measure without suffering degradation in another.

[9]Case I specified one alpha/beta threshold whereas case III specified four distinct values (see

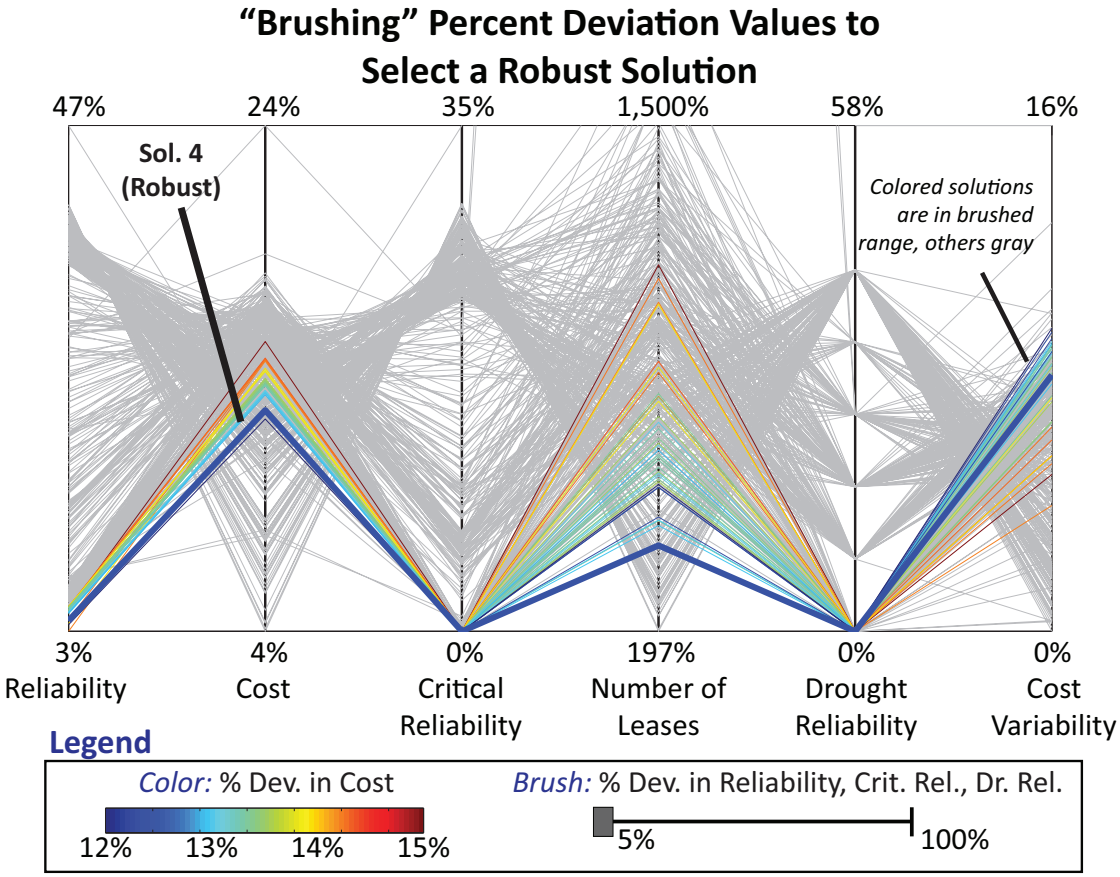

Figure 4.5. Parallel coordinate plot of percent deviation for all solutions. Here, the axes were rearranged to show maximal conflict between measures. Brushing was used to plot in gray all solutions that had greater than 5% deviation in the three reliability measures. Color was then rescaled to show the percent deviation in cost of the remaining solutions. Solution 4 was selected for further analysis.

leases and wasting more dropped transfers, this higher amount of market use leads to more robust performance in extreme SOWs. The solution's volumes of rights and options are also comparable to those of low cost solution 1. Furthermore, solution 4 comes from a simpler decision lever formulation (case I) that may be easier to implement in practice, with fewer decision levers and more straightforward rules for acquiring market transfers.

Figure 4.6 examines solutions' robustness performance in the context of the original expected-value performance measures in the optimization. The figure re-

table 3.4). Table 4.4 presents all solutions using four distinct variables, although in the MOEA search there was only one threshold value searched in case I.

Table 4.4. Selected Solutions' Properties

Solution	1	2	3	4
Name	Low 10 Yr. Cost	High 10 Yr. Cost	Comp-romise	Robust
Model Case	3	3	3	1
N_R (10^6 m^3)	37	61	46	38
N_O (10^6 m^3)	17	0	21	20
$\alpha_{\text{May}-\text{Dec}}$	1.29	1.29	1.24	1.69
$\beta_{\text{May}-\text{Dec}}$	1.53	1.34	1.60	1.69
$\alpha_{\text{Jan}-\text{Apr}}$	1.20	0.06	1.39	1.69
$\beta_{\text{Jan}-\text{Apr}}$	1.28	0.09	1.39	1.69

tains the same spatial axes, cone size, and orientation from Figure 4.3 and uses color gradients to superimpose the percent deviation results for cost (Figure 4.6a) and critical reliability (Figure 4.6b). This is a unique way to show the robustness of each region of the original tradeoff set. In prior work, solutions 1-3 were selected because of their good performance in cost, number of leases, and surplus water under the baseline SOW. Figure 4.6a illustrates that some of the solutions in this region also have low percent deviation in cost. The same region, though, exhibits poor performance with respect to percent deviation in critical reliability (see Figure 4.6b). This could lead to severe failures if assumptions about the LRGV's supply and demand conditions are wrong. In our initial tradeoff exploration [32], we did not select solution 4 because it has a high number of leases compared to the selected solutions in case III. Solutions in the region containing solution 4, however, are much more robust with respect to critical reliability, avoiding critical failures in SOW with low inflows, high demands, and other plausible future conditions.

4.4.4 Scenario Discovery

The prior sections focused solely on solutions' worst case performance in the uncertainty analysis compared to their Pareto approximate performance from optimization using the baseline SOW. Solution 4 was selected because of its acceptable performance with respect to a percent deviation metric across several measures. However, the analysis did not indicate what values of the uncertain exogenous factors caused Solution 4 to have poor performance in alternative SOW. This section

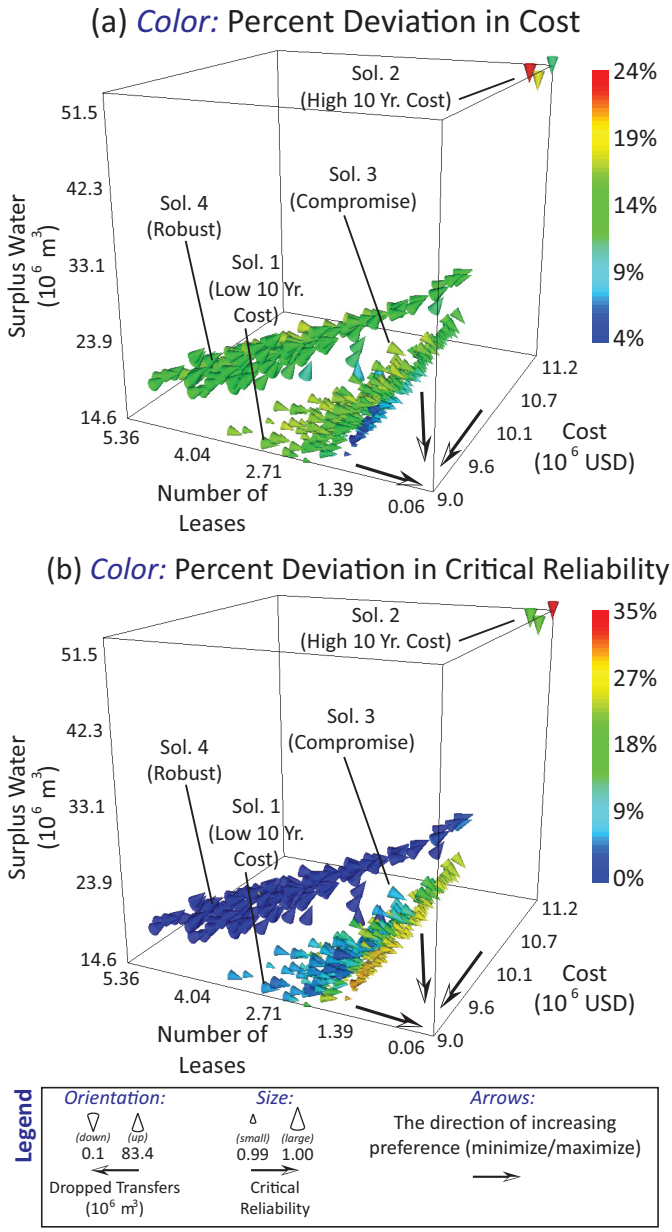

Figure 4.6. The selected solution in the context of the prior objectives. The cone coordinates, orientation, and size are retained from Figure 4.3, but the color shows the percent deviation in cost (Figure 4.6a) and critical reliability (Figure 4.6b). Note that the selected solution 4 is in a different region of the space than what was previously selected.

presents the results of computer-aided scenario discovery (SD) on Solution 4, with the goal of providing straightforward scenario descriptions of what factors cause poor performance for the solutions. Recall that SD uses PRIM to construct "scenario boxes" for three sets of performance thresholds defined in Table 4.3. The use of SD within MORDM follows the methodology of [78], using metrics of coverage and density to select appropriate scenarios[10]. Rows in Figure 4.7 show the factors sampled in the MODRM uncertainty ensemble (scaling factors in the first five rows and model parameters in the second three rows). The columns show the three threshold sets, or groups of thresholds on performance measures that were used to delineate poor performance from acceptable performance. The gray bars show the values that triggered vulnerabilities in the discovered scenarios. Each threshold set developed a unique set of vulnerable points. For example, points that caused high cost may not cause poor reliability, and so forth. The sets of box constraints are therefore independent across the three different threshold sets.

The cost threshold set was violated when low inflows were scaled to be 3.5 times likelier, high losses scaled 1.9 times likelier, high demands 3.4 times likelier, the demand growth rate higher than 1.4%, and the initial reservoir volume below 1.8 billion cubic meters. Each of the dimensions included in this scenario would cause the city to buy more water on the market and thus incur higher costs. Dimensions included in the scenario lower the amount of water supply (low inflows, high losses, and low reservoir volume) and increase the amount of demand (high demand growth rate). Moving to the reliability threshold set, we find that inflows, losses, and demands also appear in the scenario box. Poor performance in reliability also occurs when the initial rights variable is lower than 0.24. The initial rights variable is included in the simulation as a way to model the amount of water carried over from year to year by a portfolio (i.e., the amount of water available to the city before the simulation time period begins). Low initial rights is representative of the effects of a long term (potentially multi-year) drought [56]. The appearance of low initial rights in the reliability threshold set's results indicates that the portfolio's

[10]In using PRIM, we chose scenarios that had acceptable coverage and density and constrained the smallest number of dimensions possible. For example, the reliability threshold set only uses 4 out of a possible 9 dimensions. While higher-dimensional scenarios could also plausibly explain the dataset, using a small number of dimensions allows the scenario to be easy to interpret and also maximizes the density (i.e., most of the points in the scenario are indeed vulnerable).

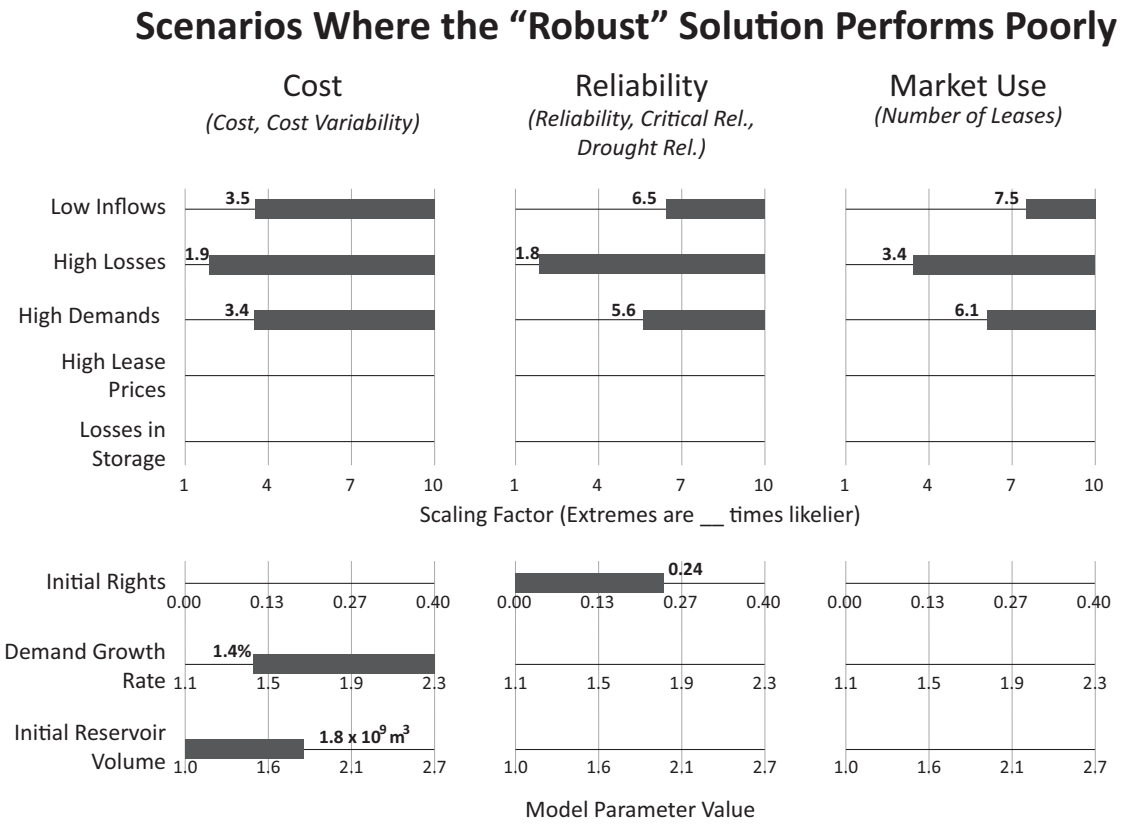

Figure 4.7. Results of scenario discovery. The bars indicate SOW where the "Robust" solution (see solution 4 on Figures 4.5, 4.6 and Table 4.4) performs poorly on three groups of performance measures.

market use strategy may sometimes fail if there is not sufficient water in the city's supply account, especially in these drought conditions. The third threshold set is market use. This set focuses only on SOW when the number of leases are in the highest 10%. The gray bars here show that inflows, losses, and demands are the most important factors for causing vulnerabilities.

It is not our goal to claim that each of the SOW in the discovered scenarios will actually occur in future planning. Ultimately, the decision makers would use these scenarios to inform discussions of what factors are most important for future planning. As an example, losses were critically important in all three threshold sets. Scaling factors that caused vulnerabilities for the losses were relatively low, triggered at 1.9 for cost, 1.8 for reliability, and 3.4 for market use. Low to moderate values for scaling factors could very plausibly occur in future planning periods, since they do not represent drastic changes compared to the distributions attained

using LRGV test case's historical data (see Figure 4.2). Therefore, the system should likely be monitored to ensure that high losses are not being observed in the system, since low losses are important for ensuring robust portfolio performance. Triggered scaling values for other dimensions indicate a higher level of robustness; for example, vulnerabilities in cost, reliability, and market use were only triggered when low inflows were scaled to be 3.5, 6.5, and 7.5 times likelier, respectively.

4.5 Discussion

Generating high-quality planning alternatives for complex environmental systems poses several significant challenges. First, the systems are often characterized by multiple, conflicting performance measures. Traditional design approaches often try to aggregate the multiple measures into a single metric of performance (e.g. [44]). This aggregation, though, favors certain aspects of performance over others in unpredictable ways [16, 17, 21, 24, 45, 154]. In contrast, MORDM's problem formulation considers multiple conflicting performance measures explicitly and simultaneously. The solution to a many objective problem formulation is a set of tradeoff solutions, each of which is non-dominated with respect to multiple performance measures. Considering the tradeoff as a whole allows decision makers to learn about trends and properties of their modeled system, such as how sensitive it is to change, how it responds to extreme events, and what its performance is with respect to multiple measures [157]. Tradeoffs for the LRGV, for example, showed that a water market can lower surplus water and cost while providing high reliability [56]. A second challenge comes from planning problems that impose severe constraints on performance that limit our ability to find feasible solutions (i.e., solutions that meet or exceed all planning constraints). MORDM uses an advanced solution generation technique (MOEAs) to find non-dominated planning alternatives for the formulation. The efficacy of MOEAs on problems with severe constraints was recently demonstrated using a severely constrained engineering systems design problem [183]. The authors used random Monte Carlo sampling to generate possible alternatives and found that only 4 out of a possible 50 million randomly generated solutions met the feasibility constraints. A MOEA, however, was able to find feasible solutions in a few thousand design evaluations.

MORDM's framework subjects the set of solutions identified by the MOEA to a rigorous, quantitative evaluation under a wide range of plausible future conditions, and uses visual interactive techniques to assist policymakers in selecting robust solutions. Interactive visual analytics is an important part of MORDM because it enables exploration across many performance tradeoffs, robustness measures, and critical exogenous factors simultaneously. We used the percent deviation metric to examine the performance of solutions across various plausible futures, and we started with showing the metric across six different performance measures at once (Figure 4.4). Figure 4.5 represents a "brushing" exercise that eliminated solutions with high deviation in the three reliability measures and utilized a color gradient to show deviation in cost. This reduced the number of solutions from which to choose and enabled the decision maker to choose from maximally different alternatives [15] that were also robust to deep uncertainties. MORDM's use of percent deviation for solution selection is a unique contribution of this work, compared to prior MOEA studies that focused solely on expected value performance measures in solution selection. The approach is a promising way to include decision maker participation in the planning process and to meet decision makers' desire to maintain a high level of performance even under deeply uncertain future conditions such as climate change [140].

Including climate change in system planning underscores a key motivating question for managing complex environmental systems: how do we generate plausible projections of future system conditions and use them to make decisions? Although planners could previously assume that the statistical properties of their systems would not change under long planning horizons, anthropogenic changes limit our ability to properly characterize expected conditions in the future [1, 30, 71, 76]. To plan for deeply uncertain future risks, decision makers often use scenario analysis, which seeks to provide coherent storylines of plausible future events [72]. The Special Report on Emissions Scenarios (SRES) is a well-known scenario analysis for climate planning, with a small number of axes designed to cover assumptions about the type of economic development and governance [74, 75]. An integrated assessment model quantifies these properties and uses point values of variables such as global carbon emissions as inputs to other models. While useful for creating a formalized discussion for climate change, the SRES example highlights key

issues with the classical scenario approach [30, 73]. The choice of socioeconomic data for the SRES scenarios is made independently of possible system vulnerabilities such as climate warming. Since traditionally there is no feedback between this choice of scenario data and ultimate system performance measures, there is no way to characterize how changes in those assumptions can affect likelihood of system vulnerability. In fact, decision makers will often hold optimistic views of future forecasts (and scenarios) while limiting their choice of mitigating actions due to their aversion to risk [59]. MORDM addresses these issues by simulating a wide array of plausible futures, requiring only an assumption of plausible ranges of exogenous factors. Subsequently, visualizations of solution performance can identify the important values of uncertainties without *a priori* estimates of values that cause performance failures. Problem formulations in the framework promote innovative decision levers [59] and facilitate selection of solutions that have acceptable performance across a wide array of simulated futures.

Our MORDM demonstration culminated with interactive scenario discovery (SD) for the LRGV test case. The goal of SD is to find simple explanations of which factors cause vulnerabilities for selected robust solutions [107]. We tested a large number of different scaling factors to perturb the forcing data for the LRGV, and SD indicated that performance vulnerabilities were not strongly controlled by some deeply uncertain variables (such as lease pricing). However, the analysis suggested that even for small perturbations in losses, there could be large deviations in cost, reliability, and market use for our selected solution. Prior studies with the LRGV simulation did not strongly focus on the losses variable [122, 123], and this demonstrates a decision bias termed cognitive hysteresis [58]. Cognitive hysteresis refers to a situation in which the initial formulation of a problem strongly motivates the selected solution for the problem. SD can help decision makers formulate new hypotheses about their system and identify the most important factors for future planning. In the LRGV, this means closely monitoring the amount of water coming into the reservoir and properly adjusting design decisions and water balances for losses.

Planning for complex environmental systems should acknowledge that the planning problem itself is constantly evolving. This phenomenon is partly due to new system conditions (such as the impacts of climate change or rapid population

growth) that emerge over time. However, the planning formulation also changes due to learning that occurs as decision makers actually solve different iterations of their problem. In the LRGV, solving for tradeoffs showed us the dramatic effect of the market, and SD results demonstrated the importance of certain deeply exogenous factors. MORDM can therefore be considered an iterative process, with each step feeding back to the problem formulation to generate new hypotheses [57] about a system's decision levers, performance measures, uncertainties, and the governing relationships between actions and outcomes. MORDM's optimization and systems analysis focus is on enhancing learning and collaborative problem construction when supporting decision makers [29, 67, 70, 170]. In this manner our approach echoes the concerns of [65], in that the formulation of a planning problem is uniquely coupled to the actual solution to the problem. Figure 4.7 gave an example of three such discovered scenarios for the LRGV test case that could cause concern for water managers. The scenario showed that low inflows and high losses could cause managers' costs to increase, their reliability to suffer, and their market use to be higher than anticipated in a baseline SOW. After examining these results, decision makers may want to pose a new problem that focuses on low inflow simulations and tries to optimize their water acquisitions under these conditions [184].

4.6 Conclusion

In their discussion of "wicked" public domain planning problems, Rittel and Webber [65] argued that these problems have no definitive solution, since the problem formulation is constantly modified as planners learn more about the system and conditions evolve. The authors also argue that, for such "wicked" problems, the lack of any public consensus undercuts the usefulness of a single objective philosophy (e.g., increasing national welfare), in favor of a multi-objective perspective that can engage diverse stakeholders with differing worldviews. These challenges are compounded by a recent focus on human-induced changes, such as climate and land use change [1, 71, 159]. The MORDM framework presented in this paper seeks to address these challenges with multi-objective planning that considers deep uncertainties while evaluating alternative problem formulations.

MORDM employs many objective search using multi-objective evolutionary algorithms to address conflicting performance measures by developing Pareto approximate tradeoff sets. MOEA search allows analysts to generate tradeoff sets even under severe performance constraints and uncertainties that arise as systems undergo change. We also address an important issue in analyzing tradeoffs: How do we inform the final negotiated selection of a solution from the tradeoff set? MORDM helps decision makers select solutions that perform well under a wide array of deeply uncertain future trajectories (i.e., robust alternatives). The framework samples solutions under a wide array of plausible futures and calculates robustness metrics across all performance objectives using RDM. Furthermore, MORDM uses statistical data mining algorithms to clarify which exogenous factors control performance failures for the system. The framework represents the first time MOEA optimization has been used in combination with RDM techniques.

The MORDM framework is demonstrated using a risk-based water portfolio planning problem in the Lower Rio Grande Valley of Texas, USA. The tradeoff solutions exhibit a wide range of performance under the MORDM uncertainty ensemble. Our results indicated that the simplest of four problem formulations had the most robust performance. The chosen formulation was able to achieve high performance with respect to critical reliability (i.e., avoiding catastrophic failures). A subsequent demonstration using scenario discovery characterized which exogenous factors strongly controlled LRGV performance failures. Losses in reservoir inflows influenced failures across a broad suite of measures, even when they were scaled only slightly differently than the baseline historical data. This finding frames how the MORDM framework can be used to inform adaptive management of complex environmental systems undergoing change. For the LRGV, this could include monitoring evaporation rates and triggering new planning in the event of extreme droughts. Such iterative, interactive methods aid in helping find more sustainable solutions to "wicked" environmental planning and management problems and can aid in building consensus across a broad range of decision maker preferences. As a further demonstration, chapter 5 will present an additional implementation of MORDM to explore water resources management in a different context.

Chapter 5

Planning in the Susquehanna: Formulation Biases and Consequences of Deep Uncertainty

This chapter demonstrates how Many Objective Robust Decision Making (MORDM), introduced in this dissertation, can be used to test the robustness of management alternatives for water supply infrastructure in the eastern United States. The Lower Susquehanna case study explored in this chapter exploits stage-storage relationships, demands, and inflows from an existing study by the Susquehanna River Basin Commission (SRBC). The SRBC study reflects their ongoing efforts to actively coordinate the competing demands between the municipal water supply and thermoelectric power generation stakeholders within the Lower Susquehanna. Two many objective problem formulations are developed and evaluated using the Interactive River Aquifer Simulation (IRAS)-2010 [185, 186], with a daily time step. The first formulation is reflective of the classic history-based deterministic water resources systems planning approach that still dominates practice [34–39] conditioned on the historical inflow and evaporation record from 1932-2000. The second formulation moves beyond the historical record by considering streamflow and evaporation uncertainties modeled using a K-Nearest Neighbor stochastic simulation strategy. MORDM is then used to contrast the robustness of the two formulations' Pareto approximate solutions' robustness to deep uncertainties, associated with water demand targets, prolonged drought, and increasing

interannual variability of streamflow and evaporation.

The optimization-simulation and MORDM components of this study posed severe computational demands in excess 1,000,000 hours of computing. Consequently, each optimization formulation was solved using the massively parallel Borg Multiobjective Evolutionary Algorithm (MOEA), which can efficiently exploit hundreds of thousands of computing cores simultaneously. Moreover, this study contributes a parallel extension of the MORDM framework that can exploit massively parallel computing. A journal article based on this chapter, co-authored by Evgenii Matrosov of University College London and Patrick Reed, will be submitted in Spring 2013.

5.1 Introduction

Chapters 3 and 4 proposed and demonstrated two new planning innovations: *de Novo* Planning as a means of simultaneous evaluation of competing formulations, and Many Objective Robust Decision Making (MORDM) as a means of aiding the selection of formulations and/or specific alternatives that are robust to deeply uncertain conditions. These early chapters demonstrated the *de Novo* Planning and MORDM frameworks using the LRGV test case. Although the LRGV test case has significant value, the problem's planning concerns only a single city as the sole stakeholder, and the potential for broad regional water transfers is very specific to the legal context of the western U.S. In contrast, the water planning challenges in the eastern U.S. require a transition to a more infrastructure-dominated multi-sector perspective.

Reservoir operations research has a long history of contributions to the literature [26, 37, 187, 188]. There are three major problem classes that have been addressed [26]: (1) determining how large a new reservoir should be, (2) long-term planning for evolving system demands, and (3) real-time operations at timescales of minutes to hours. In each category of planning, classical water systems planning approaches have most often used a single benefit function to optimize which water allocations are made at each timestep [39]. As demonstrated in the prior chapters of this thesis, the many objective approach employed in this study helps to avoid decision biases associated with aggregating several benefit functions to-

gether [45]. An aggregated function obscures the needs of different constituents, such as the conflict between human needs and ecosystem services [42]. In contrast, the MOEA-based optimization approach advanced in this work enables the direct analysis of system tradeoffs. MOEAs can be directly coupled to complex infrastructure simulations (figure 2.4), allowing for uncertainty analysis and explicit consideration of multiple sectors. Such an approach was advocated as early as the 1960's by the Harvard Water Program [24], for the specific purpose of reservoir systems planning.

While MOEAs can enhance the complexity of systems and objectives that we can consider, another point of concern in these problems stems from the data used to drive the analysis (i.e., the input hydrologic data). Milly et al. discuss that using a single historical timeseries of data to determine system performance is problematic under climate change, land use modifications, and population growth [1]. Climate change has the potential to modify the magnitude of flows across the U.S. [131, 189–191]; even though a system may have performed well in the historical record, it may not fare well under a modified climate. Lins and Cohn argue that the most important effect of climate change may be the "increased uncertainty" introduced into the systems analysis [190]. Stochastic generation techniques evaluate water resource systems for a wider range of hydrologic conditions [192–195]. Hypotheses for hydrologic change can be proposed, and the stochastic generation techniques can then statistically generate multiple flow sequences that still reflect regional hydrology. Methods for generating stochastic hydrologic data have been shown useful in a variety of applications: finding proper reservoir storage [196], exploring the impacts of climate change [197, 198], determining performance of water transfer schemes [199], and estimating the risk of reservoir depletion [200]. Across each of these domains, considering a broader suite of hydrologic uncertainties poses a severe computational challenge, since this requires larger ensembles of system simulations (i.e., Monte Carlo simulations based on synthetic hydrology). When coupled with MOEAs, these Monte Carlo simulations are particularly demanding because they increase the time required to perform search. Furthermore adding uncertainty could lead to degradations in the search process [54] if an inappropriate optimization technique is selected.

A further challenge stems from the concept of deep uncertainty, in which de-

cision makers do not know or cannot agree upon the full suite of risks that are posed in their system (see section 2.2.1). Beyond introducing the ability to generate synthetic data, managers may also be concerned with how the *properties* of their system could change in an uncertain future. This chapter presents an illustrative MORDM analysis that explores several significant uncertainties for water planning. We hypothesize that a simple measure of interannual persistence can show a system's resilience to changes in long-term memory, that is, how easily can the water resources system respond to large perturbations in the record. Increased persistence could cause significant long-term droughts [184]. In the presence of such droughts, are deterministic planning techniques sufficient to generate alternatives? Is a stationary statistical model sufficient to generate better alternatives, or do those solutions exhibit severe performance vulnerabilities? In addition to system persistence, we also test how changes in the variance of annual records affects system performance. Both sampled dimensions recognize that there is often a significant uncertainty in estimating statistical model parameters, which can propagate to modifications in output system performance [201]. A further source of deep uncertainty in our analysis comes from modifications in water demands. Reservoir planning is typically done making fixed assumptions of water demand targets. If alternative designs are generated using fixed demands, what is the impact of changes in demand on the alternatives that were optimized using fixed demand targets?

This chapter explores the robustness of management alternatives for the Lower Susquehanna (see section 2.4.2) that result from two alternative formulations: (1) a many objective deterministic optimization–simulation based on historical SRB hydrology and (2) a probabilistic many objective optimization under uncertainty that more comprehensively explores the impacts of streamflow and evaporation uncertainties. The competing formulations are subjected to MORDM analysis to contrast their potential vulnerabilities to deeply uncertain exogeneous assumptions (e.g., demand targets, prolonged droughts, and increased variability in streamflow and evaporation). The probabilistic formulation and MORDM analysis requires millions of simulations of the Lower Susquehanna system across their ensembles. This poses a major computational challenge and requires a transition beyond the SRBC's current OASIS simulation of the system because the model's proprietary

LP solver is not portable to large scale HPC environments (see section 2.4.2). In order to create rapid evaluations of the system in an HPC environment, the test case is built using a rule-based simulation model called the Interactive River Aquifer Simulation (IRAS)-2010 [186]. Its fast evaluation time allows hundreds of thousands of model evaluations to be done in parallel on an HPC system in a reasonable amount of time.

The Lower Susquehanna test case focuses on the interaction between conflicting uses, where reducing demand for one sector's use will facilitate storage that can be used to fulfill other demands. To this end we generate alternative management plans for reducing several conflicting demands, as well as environmental flow-based regulations downstream of the dam. Beyond the case study-specific insights, a key goal in this study is to expose the potential negative consequences or formulation biases that result from the use of deterministic planning-based historical hydrology. Although our deterministic many objective formulation contributes a more comprehensive characterization of the Lower Susquehanna test case's multi-sector demand tradeoffs relative to classic single objective planning [34–39], deeply uncertain changes in the SRB have the potential to exacerbate drought risks. Moreover, this study explores if the sole consideration of hydrological uncertainties (i.e., stochastic generation of streamflow and evaporation) is sufficient to discover robust alternatives for the system while simultaneously identifying the key deep uncertainties that control its vulnerabilities.

5.2 Methods

5.2.1 System Representation

5.2.1.1 IRAS-2010 Water Resource Simulator

IRAS-2010 [186] is an updated version of the original IRAS developed by Loucks et al. [185][1]. Similar to the SRBC's OASIS model of the system, the IRAS-2010

[1] As documented by Matrosov et al. [186], IRAS-2010 represented several major improvements to the original IRAS. Improvements relevant to our use of IRAS-2010 in the current chapter include better calculation of water deficits, the ability to include demand restrictions, performance measure outputs, text-based input and output files, a loop for multiple simulation runs used in the stochastic analysis, and a faster run time. We completed a C++ translation of the original

model treats a water resource system as a collection of nodes and links. Nodes represent points where water can be stored (i.e. lakes, reservoirs, groundwater aquifers, wetlands), where water enters the system (gauge nodes), or where water is demanded or used (demands, consumptions, hydropower, or pumping) [186]. A link conveys water between nodes. Water enters the system through gauge nodes, and it is then diverted and stored based on user-defined rules and IRAS-2010 demand priority calculations. Since IRAS does not use an optimization "engine" to allocate its water, diversions and storage releases are made based on the rules specified by the user and not a system-wide benefit function.

There are two types of water release from storage nodes. Supply-driven releases are based on the storage at each node and are limited by a minimum and maximum release at each level of storage. To coordinate across multiple storage nodes, balance tables can be provided, which specify a target volume of each reservoir in the balance group based on the volume in the rule-site (independent) reservoir or aggregated group storage. Release from the rule-site reservoir is regulated by a release rule function. A second type of release is termed demand-driven release, in which a demand node requests an amount from storage, and the storage node releases the specified amount. IRAS-2010's demand-driven release system is the most similar to OASIS, since OASIS uses a set of constraints to ensure that exact volume are delivered to demand nodes if possible.

IRAS can be used to simulate the effects of changing conditions in a water resource system. For example, Brandao and Rodrigues [202] used IRAS to simulate how changes in reservoir storage affect downstream flow conditions. Matrosov et al. [203] studied the effect of infrastructure expansion in the Thames Basin, including structural improvements (new reservoirs) and demand management (reductions in pipe leakage). This chapter focuses on demand management in the absence of new structural improvements. In the simulation model, restrictions in a demand node can be linked to conditions in a supply node using two variables, the Reservoir Threshold and the Restriction Value. Consider a Reservoir Threshold of 85% and a Restriction Value of 10%. Under these parameters, when the storage volume goes below 85% of the maximum storage capacity, the demand

IRAS-2010 code and added some additional features to help meet the needs of this chapter, including the ability for the model to accept timeseries of evaporation.

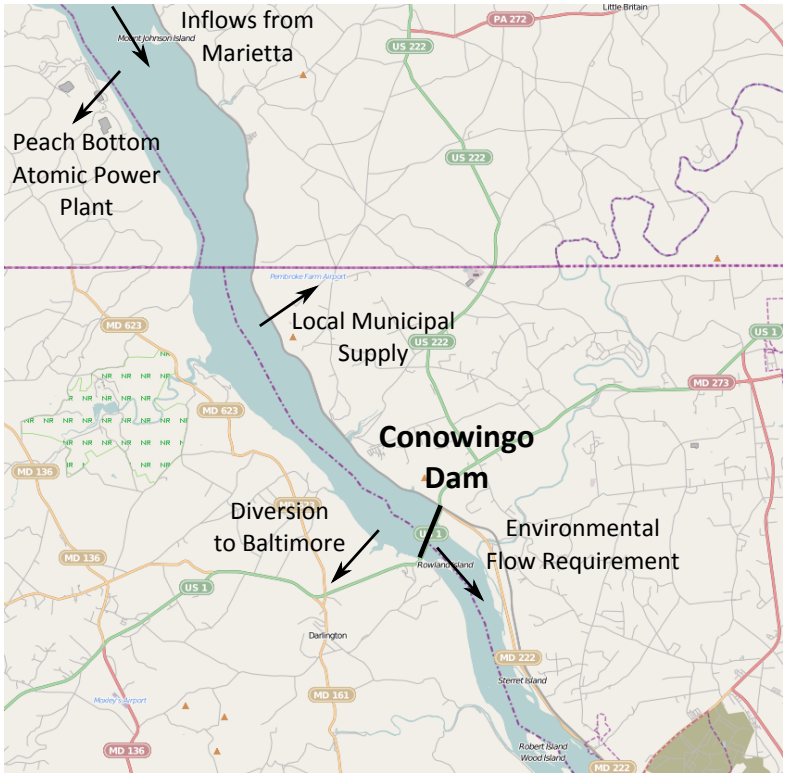

Figure 5.1. Schematic of withdrawals from the Conowingo dam for the Susquehanna River Basin test case. Map adapted from figure I-2 in Conowingo Pond Management Report [43], using a base map from the OpenStreetMap project, http://www.openstreetmap.org/.

must be reduced by 10%. IRAS-2010 uses a variable timestep as small as one day. User-defined data can be entered in different time increments; for example, some data can change seasonally and other data can change for each day of the year. A set of performance measure calculations keeps track of the frequency, duration, and magnitude of meeting user-defined storage targets. Additionally, performance measures track the desired demand in each time step and how much of that demand is met. The reader is encouraged to consult references for the full set of equations and details about IRAS and IRAS-2010 [185, 186].

5.2.1.2 Lower Susquehanna Network

Figure 5.1 is a schematic of the Lower Susquehanna test case network, showing the location of the Conowingo dam and associated water demands. Inflow to the system begins at Marietta, using data from USGS gauge 0157600 at Marietta,

Table 5.1. Conowingo Storage Performance Targets

Target	Reservoir Stage (ft)	Percentage of Capacity
Normal Pond	108.5	97.2
Summer Weekend Recreation Level	106.5	91.7
Muddy Run Shuts Down	104.0	85.0
Critical Level for Peach Bottom	103.5	83.6
Minimum Pond	100.5	75.6

PA [204][2] Marietta's flows serve as inflow to Conowingo, the main supply for the system. As shown in the figure, Conowingo then releases to local municipal supply, the Peach Bottom Power Plant for cooling, a diversion to the Baltimore system for water supply, and downstream releases to meet minimum federally-mandated low flow requirements (i.e. Environmental Flows).

The simulated Conowingo dam has a realistic stage-storage relationship based on Federal Energy Regulatory Commission (FERC) disclosures [205] and the stage-storage relationship in the SRBC's OASIS model of the Lower Susquehanna. Relevant storage targets in the Lower Susquehanna test case are shown in table 5.1. Recall that OASIS uses a series of LP constraints and goals to allocate water to Conowingo's demands. In IRAS, we retain the target levels for several representative demands, but we also add constraints to the stage-storage relationship to reflect maximum releases and system geometry as documented for the FERC disclosure. For example, according to the table, the reservoir is constrained to only release water above the minimum pond level of 100.5 feet. Additionally, the Peach Bottom Power Plant can only withdraw water when storage is above 103.5 feet.

Table 5.2 outlines the demand targets used in the case study. A local municipal demand[3] and the Peach Bottom cooling water demand both use monthly patterns

[2]In OASIS, a much larger portion of the basin is included, so the boundary condition of the model occurs much further upstream. Therefore, relevant inflows come from modeled allocations of water at relevant nodes directly upstream of the dam. For demonstration purposes, we use the observed Marietta flows as input to the model. By using the observed flows, we assume that actual historical management above Conowingo is already reflected in the record, allowing us to model how Conowingo would need to react under those circumstances.

[3]Demand data here was adapted from a demand node from the SRBC's OASIS simulation,

Table 5.2. Demand Targets

Target	Range (mgd)	Notes
Local Municipal	0.79-2.89	Monthly Pattern
Peach Bottom	13.15-26.69	Monthly Pattern
Baltimore	300	Constant
Environmental Flow	2,262.10-6,463.15	Seasonal Pattern (See Table 5.3)

Table 5.3. Environmental Flow Requirements

Beginning Date	Ending Date	Flow (cfs)
January 1	March 31	3,500
April 1	April 30	10,000
May 1	May 31	7,500
June 1	December 31	3,500

adapted from SRBC's OASIS data. Baltimore and the Environmental Flow requirements are modified slightly from the data used by the SRBC. For illustrative purposes, our Environmental Flow condition requires the release of a regulated amount shown in table 5.3, regardless of the observed flow at the Marietta gauge. In the real system, the flow target is linked to flow at Marietta, such that the amount of release is reduced if the observed flow is lower than the target[4]. Evaporation data at Conowingo comes from SRBC's OASIS model and the Conowingo Pond Management Report [43], with a daily evaporation rate given for each day in the historical time series.

The planning horizon spans from 1932 to 2000, with a daily time step. The date range was chosen to correspond with data availability for both the USGS gauge data and the SRBC's OASIS-based evaporation rates. Note that our IRAS system representation was designed to capture key Lower Susquehanna demands

which represents a small municipal demand in the proximity of the reservoir. The case study assumes this small demand cannot be modified using drought restrictions. Failures for this demand, though, can still occur if there is not sufficient water in the system to meet the desired demand target.

[4]Effectively, our low flow regulations are higher than in the real system, representing a worst-case condition. We are determining if these higher environmental flow demands will still allow the reservoir to meet the other conflicting water uses.

and releases from the SRBC's OASIS simulation. The SRBC's OASIS simulation was developed for the full SRB and models many more demands and upstream releases than this chapter's Lower Susquehanna test case. The primary focus of the Lower Susquehanna test case is to capture critical stakeholders' demands and risks given the deep uncertainties posed by the regional demands, increasing hydrologic variability, and potential for extended droughts.

5.2.2 Performance Measure Definitions

After running the simulation for a given inflow and evaporation time series, quantitative performance measures are used to capture system performance, or how well the storage nodes were able to meet supply targets and volumetric demands. As argued by Hashimoto et al [40], quantitative performance measures are a critical part of multiobjective water systems analysis, designed to show how systems perform in an "uncertain future." The set of quantitative measures used in this chapter is adapted from a recent many objective analysis of London's water supply [203]. This section will present the equations of each type of measure. Since each measure can be quantified in multiple locations in the supply network, this section will present generic descriptions of the measure without regard to where the value is quantified. Section 5.2.4 will then present how these measures are combined into many objective problem formulations for the Lower Susquehanna. In the following equations, $\mathbf{x}$ represents the vector of decision variables, and n_{ts} is the number of timesteps.

The first three measures used in the formulation are calculated at storage nodes. The measures quantify the system's ability to meet predefined storage targets (set percentages of the total maximum storage at a node). Figure 5.2 shows an illustrative example of a storage timeseries. The vertical axis plots volume, and the horizontal axes plots 7 timesteps. The bold horizontal lines on the storage figure indicate the two storage target levels, indicated by the subscript k.

Water managers are often interested in how often their reservoir or storage site can meet given storage targets. For example, fishing and boating often requires that reservoir storage levels are kept moderately high, which could conflict with objectives of flood control or water supply. Specifically, let S_t to represent the

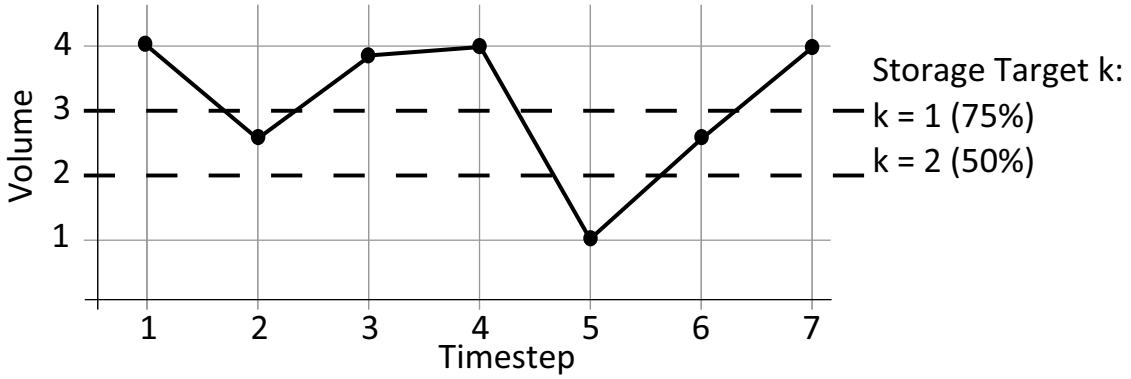

Figure 5.2. A generic timeseries used to illustrate storage performance measures. The horizontal axis plots time while the vertical axis plots volumetric storage, and a hypothetical storage timeseries is shown. Two storage targets are indicated with bold dashed lines.

node's volumetric storage and g_k to represent the k^{th} volumetric storage target. Then, we define B_t to equal 1 if there is a failure in timestep t (i.e., the storage falls below the target):

$$\begin{aligned} \text{if} \quad S_t < g_k \quad \text{then} \quad & B_t = 1 \\ \text{else} \quad & B_t = 0 \end{aligned} \tag{5.1}$$

Then let nf_k equal the number of times that storage falls below target k.

$$nf_k = \sum_{t=1}^{n_{ts}} B_t \tag{5.2}$$

Using the prior equations, reliability[5] for target k indicates how often the storage was above the target, shown in equation 5.3.

$$r(\mathbf{x})_k = 1 - \frac{nf_k}{n_{ts}} \tag{5.3}$$

In figure 5.2, the storage is above target 1 for 4 timesteps, yielding a reliability of 4/7 or 57%. Similarly, the storage is above target 2 for 6 timesteps, with a

[5]Recall that chapters 3 and 4 used a measure of reliability to quantify performance of the LRGV's water supply portfolios. The reliability used in this chapter is similar, but it quantifies the ability of a reservoir to effectively retain storage above a certain volume, instead of capturing how a system meets flow demands. Additionally we are treating the entire series as one record instead of breaking up the time into distinct planning years.

reliability of 6/7 or 85%. Figure 5.2 uses target 2 as an additional storage target. Although failures at target 2 will occur less often than target 1, the failures will be more severe and more difficult from which to recover. The user can specify as many storage targets as desired, with the ability to calculate reliability at each of the storage targets.

Beyond considering user-defined storage targets, managers also want to avoid catastrophic, severe failures that occur infrequently. The Minimum Storage Level objective was designed to quantify the severity of such failures, similar to the vulnerability metric posed by Hashimoto et al. [40] and used in Chapter 3. The objective is defined in equation 5.4,

$$\text{Minimum Storage Level}(\mathbf{x}) = \min_{t \in n_{ts}} \frac{S_t}{S_{\text{cap}}} \tag{5.4}$$

where S_{cap} denotes the storage capacity. Note that the Minimum Storage Level is expressed as a percentage of capacity, in order to allow measure values to be compared in systems with multiple storage nodes. In figure 5.2, the minimum storage level occurs at timestep 5, yielding a measure value of (1/4) or 25%.

Demand measures are calculated based on the target and delivery at each time step. Figure 5.3 presents an illustrative example of meeting demand targets. In the figure, time is plotted on the horizontal axis and volumes are plotted on the vertical axis. In each timestep, the system tries to meet demand targets indicated by the dots. The system actually delivers an amount shown with a gray bar; if the bar falls below the dot there is a failure in the time step[6].

Volumetric reliability [203], measures how much of the desired demand was met for the entire time series,

$$vr(\mathbf{x}) = \sum_{t=1}^{n_{ts}} \frac{Y_t}{D_t} \tag{5.5}$$

where Y_t is the delivery and D_t is the demand. In figure 5.3, during the entire timeseries there was a total of 23 units demanded, but only 17 were delivered.

[6]Demand delivery results are presented in a manner similar to prior bar graphs in this dissertation (i.e., figure 3.7), but with several differences. In contrast to the prior figure, the height of the bar shows the amount delivered to the demand, which can never be higher than the demand target. Additionally, the source of the water is not shown, since the simulation does not distinguish between whether the delivery comes from a release from reservoir storage, inflow from a gauge node, or inflow from points upstream of the demand in the network.

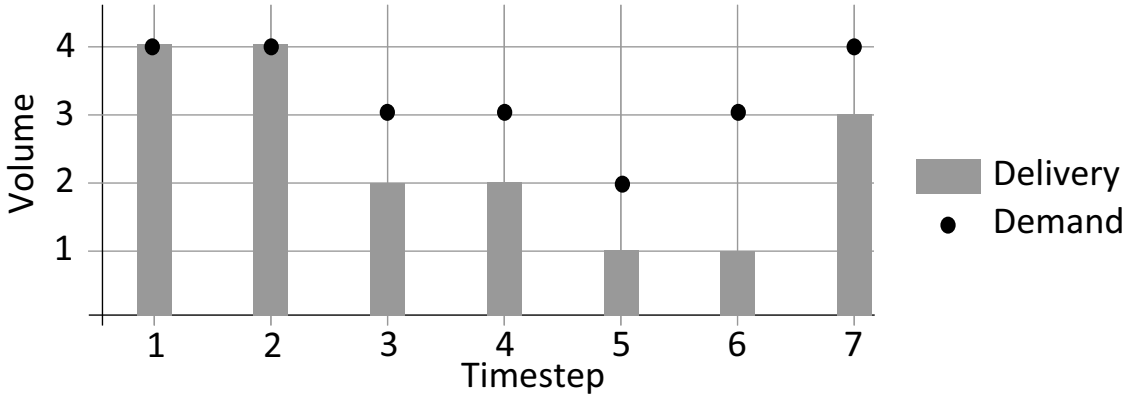

Figure 5.3. A generic timeseries used to illustrate demand performance measures. The horizontal axis plots time while the vertical axis plots volumes. The dots show the demand target at each time step while the bars indicate the delivery for the demand. The dot having a higher vertical position (such as at time steps 3-7) indicates the system was not able to meet the demand.

This is equivalent to a Volumetric Reliability of (17/23) or 74%. The benefit of this measure is that it distinguishes the volume of shortage deficit at each time step, with higher values of the measure indicating a larger percentage of demand met. However, at certain demand locations, managers may want to avoid severe failures where the shortage is a high percentage of the water demand. The Shortage Index (SI) measure [206, 207] penalizes large shortages by taking the sum of the squared shortage amounts at each time step, as shown in equation 5.6.

$$SI(\mathbf{x}) = \sum_{t=1}^{n_{ts}} (\frac{D_t - Y_t}{D_t})^2 \tag{5.6}$$

Higher values for SI indicate poorer performance. To calculate SI for series in figure 5.3, we add the squared shortages at each step (i.e., $(1/3)^2$, $(1/2)^2$), yielding a value of 1.48. SI is sensitive to large shortages, so we hypothesize that adding SI to a formulation that already includes Volumetric Reliability will ensure that if shortages occur, they are likely to be less severe (to ensure good performance in the SI measure).

In summary, this chapter uses several types of generic performance measures for both supply and demand nodes to quantify system performance. Note that each point in the system can exhibit each of these measures, and it is up to the user to choose appropriate measures for their problem formulation. In this manner,

we can construct multiple problem formulations that use different measures as we learn more about the implication of each objective on the final set of negotiated solutions [29][7].

5.2.3 Stochastic Hydrology

Measures presented in the prior section quantify several aspects of engineering performance for the Lower Susquehanna test case. However, insights using the deterministic measures will be biased by the specifics of the historical record. For example, the Minimum Storage Level objective is contingent on the single day with the most severe failures in the historical record. Using a broader ensemble of simulated streamflow sequences allows a water resources manager to evaluate the performance of their system under a better sampling of plausible conditions. Methods for generating synthetic streamflow sequences began with the work of Thomas and Fiering in 1962 [208] and Box and Jenkins in 1970 [209]. Rajagopalan et al. [195] reviews modern approaches, separating them into parametric approaches, that assume a standard functional form for the data, and nonparametric approaches that use the empirical distribution of the data for resampling. They advocate a K-Nearest Neighbor (K-NN) approach [210] that finds "similar" data points at each time step and uses the neighbors to create new synthetic records based on the historical data. The approach has been demonstrated across several applications [198, 210–213].

The sampling approach used in this chapter follows the technique of Nowak et al. [213]. We have an observed daily flow record of $n_{\text{historical yr}}$ historical years. Initially the annual flow for each historical year is calculated, and subsequently this record of annual flows to train an order 1 autoregressive (AR(1)) model. These models take the form given in equation 5.7,

$$X_t = c + \phi X_{t-1} + N(0, \sigma) \tag{5.7}$$

[7]The IRAS-MOEA linkage used in this chapter and presented in Matrosov et al. [203] makes use of a markup language that allows the user to interchange different objectives and decisions without recompiling the simulation or optimization. This feature facilitates the exploration of multple problem formulations, and it complements IRAS's text files that allow changes in infrastructure to be quickly made by the user or in combination with a GUI.

where X_t is the variable at time t, c is the mean of the original series, and σ is the standard deviation of the original series with its mean removed. The AR(1) model can then be used to simulates a series of $n_{\text{sim yr}}$ simulated years[8].

Suppose that in a given year, the AR(1) model simulates an annual value of Z. We require a method of disaggregating the simulated yearly flows into a plausible daily series that retain the statistical properties of the historical record. First a proportion matrix $\mathbf{p}$ is constructed that contains the proportion of annual flow occurring in each day of the year. Define K "nearest neighbors" with respect to the flow Z. In other words, the nearest neighbors are the K historical years that have the lowest absolute difference between historical and simulated flow. Following prior work [210, 213], we choose the number of nearest neighbors to be $K = \sqrt{n_{\text{historical yr}}}$. Next, a weight function [210] is used to sample from $i = 1$ to K, to determine which neighbor's set of flow proportions is used within the simulated year. The weight function $W(i)$ for each index i is defined in equation 5.8,

$$W(i) = \frac{\frac{1}{i}}{\sum_{i=1}^{K} \frac{1}{i}} \tag{5.8}$$

used as a discrete probability distribution for random sampling to determine which of the i years is used for the daily proportions of annual flow Z.

The following illustrative example, adapted from Nowak et al. [213], is used to illustrate the method. Consider a hypothetical year that is four days long, with 5 years of historical data as shown:

$$\mathbf{p} = \begin{pmatrix} 1967 & 0.1 & 0.3 & 0.4 & 0.2 \\ 1968 & 0.15 & 0.25 & 0.35 & 0.25 \\ 1969 & 0.1 & 0.2 & 0.5 & 0.2 \\ 1970 & 0.05 & 0.15 & 0.65 & 0.15 \\ 1971 & 0.2 & 0.2 & 0.4 & 0.2 \end{pmatrix} \qquad \mathbf{z} = \begin{pmatrix} 1967 & 35 \\ 1968 & 40 \\ 1969 & 33 \\ 1970 & 52 \\ 1971 & 43 \end{pmatrix} \tag{5.9}$$

Assume the AR(1) simulates the annual value Z for a given year to be 45. The

[8]For example, consider a historical record of 100 years. Managers may be interested in planning over a smaller time horizon such as 25 years, so they can set $n_{\text{historical yr}} = 100$ and $n_{\text{sim yr}} = 25$, using the AR(1) model to simulate multiple, 25 year sequences. This reduces the computational expense of a single run, allowing more synthetic sequences to be generated in a given amount of computer time.

Table 5.4. Decision Variables

Site	Type	Range
Baltimore	Reservoir Threshold	0.0-1.0
Baltimore	Restriction Level	0.0-1.0
Environmental Flow	Reservoir Threshold	0.0-1.0
Environmental Flow	Restriction Level	0.0-1.0

weighted resampling algorithm samples from the K nearest neighbors according to the weights in equation 5.10 and selects year 1968. According to the $\mathbf{p}$ matrix, in 1968, 0.15, or 15% of the flow occurred on the first day, 25% occurred on the second day, and so forth. We apply the proportions from the second row of p to Z to create a series of disaggregated flows:

$$Z \times \mathbf{p_2} = 45 \times [0.15\ 0.25\ 0.35\ 0.25] = [6.75\ 11.25\ 15.75\ 11.25] \tag{5.10}$$

In summary, the stochastic generation technique in this chapter is a combination of two methods: a simulation of annual flow or evaporation, and a K-nearest neighbor disaggregation technique for simulating daily flows. The method maintains plausible properties of the historical record but can also generate new data values not seen previously. For more information on this procedure the reader is encouraged to consult Nowak et al. [213].

5.2.4 Problem Formulations

5.2.4.1 Design Levers

Due to the fiscal challenges of building new infrastructure, recent reservoir planning is often focused on better management of existing systems [39]. Decision levers in both the deterministic and stochastic formulations are based on demand reduction targets that are linked to reservoir storage levels [203, 214–217]. The decision variables are listed in table 5.4.

Recall that the Reservoir Threshold controls what storage level is used to activate the restriction at the demand site. The Restriction Level then determines the percentage reduction in the demand. Separate decision levers are provided for

the diversion to Baltimore and the Environmental Flow requirement. In the actual SRB management, Baltimore diversions occur intermittently, and the Baltimore demand may be a lower priority when other demands stress the system. Here the Baltimore diversion is a constant demand, so the restriction levers explore how often the managers would have to reduce its diversions to free up water for other uses.

Regarding the environmental flow requirement, there are several mitigation strategies available to supplement the releases from Conowingo during low flow conditions [43]. Although those systems are not modeled in this case study, a demand restriction of Environmental Flow could be considered equivalent to calling for those supplemental flows. By minimizing our use of the Environmental Flow restrictions, we are determining how much the requirement is in conflict with the other uses.

5.2.4.2 Deterministic Formulation

Equation 5.11 provides the full many objective formulation for the deterministic case, using the definitions provided in section 5.2.2.

$$\begin{aligned}\mathbf{F}(\mathbf{l}) \;=\; & (f_{\text{Conowingo Rel Rec}}, f_{\text{Conowingo Rel Muddy}}, f_{\text{Conowingo Minimum Storage Level}}, \\ & f_{\text{Baltimore Vol Rel}}, f_{\text{Env Flow SI}}, \\ & f_{\text{Env Flow Vol Rel}}, f_{\text{Peach Bottom Vol Rel}}) \end{aligned} \tag{5.11}$$

The first three objectives focus on storage targets at Conowingo. $f_{\text{Conowingo Rel Rec}}$ quantifies the likelihood of meeting the summer weekend recreation level, or 91.7% of Conowingo's storage capacity. Recreation and fishing requires a relatively high storage level, so small to moderate drawdowns in reservoir storage could negatively impact this objective. The second reliability objective, $f_{\text{Conowingo Rel Muddy}}$, focuses on more severe failures at Conowingo. Its storage target is set at the intake level for the Muddy Run power plant, equivalent to 85% of Conowingo's storage capacity. Although Muddy Run is not explicitly modeled in this case study, if the reservoir drops below 85% it will be difficult to meet this unmodeled demand. Following figure 5.2, we have included two storage targets to capture storage failures of varying magnitude. A third storage objective is the Min-

imum Storage Level. This objective is calculated during the day with the lowest storage in the timeseries, expressed as a percentage of capacity. The two reliability objectives and Minimum Storage level are all maximized in the optimization.

Four objectives are included that focus on water demands. Volumetric reliability is maximized for Baltimore ($f_{\text{Baltimore Vol Rel}}$), the Environmental Flow requirement ($f_{\text{Env Flow Vol Rel}}$), and Peach Bottom ($f_{\text{Peach Bottom Vol Rel}}$). Peach Bottom is included in the objectives but it has no decision lever to help control its demands in low flow periods. The many objective analysis is designed therefore to show whether or not Peach Bottom's needs can be met by changing the other conflicting demands. The motivating assumption is that the nuclear power system is a dominant risk, and its operations are expensive to change. In addition to the volumetric reliability objectives, a shortage index is calculated for the Environmental Flow requirement ($f_{\text{Env Flow SI}}$) following prior work [203]. The shortage index objective, a sum of the squared shortages between delivered water and the required target, is minimized. Including both Volumetric Reliability and Shortage Index for the Environmental Flow objective captures two properties of its performance – the ability to meet environmental flow demands on average in each timestep, and avoidance of more severe failures that could occur less frequently. The objectives were chosen to determine if there is a clear conflict between retaining storage and meeting water demands. In order to fully meet the demands, the reservoir may need to draw down below its storage thresholds.

5.2.4.3 Stochastic Formulation

The deterministic formulation provides several measures of storage-based and demand-based performance for the Lower Susquehanna. Since objectives such as the Minimum Storage Level are conditioned to the single worst day on the record, results from the deterministic optimization could be "optimistically biased" [218]; in other words managers could be led to believe that results could be no worse than what was found in the historic record. As recognized by Matalas [192], "the historic sequence is very unlikely to be repeated in the future, [therefore] the single response obtained from this sequence is unlikely to be representative of the system's future response." The stochastic simulation technique used in this chapter, adapted from Nowak et al. [213] has the key advantage of being able to generate

new flow values that were not seen in the record, including the ability to generate plausible droughts that are worse than the historical record. Furthermore, our use of multiple records allows a wider array of possibilities to calculate performance, increasing the chance that we generate interesting sequences of data with long persistence, extreme magnitudes, and other critical properties that stress the system.

Each stochastic simulation constitutes 100 equally plausible traces of a 35 year long daily timeseries generated using the annual model and proportion vectors (section 5.2.3). To illustrate how stochastic objectives are calculated, consider the Minimum Storage Level objective for synthetic record i:

$$\text{Minimum Storage Level}(\mathbf{x})_i = \min_{t \in n_{\text{stoch ts}}} \frac{S_t}{S_{\text{cap}}} \tag{5.12}$$

where $n_{\text{stoch ts}}$ denotes the number of time steps in one stochastic record (365 days $\times$ 35 years). We then construct a stochastic version of the objective using a worst case (min/max or max/min) approach. The objective is to maximize the minimum value for Minimum Storage Level over the 100 replicates:

$$f_{\text{Stoch Minimum Storage Level}}(\mathbf{x}) = \min_{i \in [1,100]} \text{Minimum Storage Level}_i \tag{5.13}$$

Each objective in equation 5.11 is expanded in a similar manner, with the maximization objectives set to maximize the minimum value over the 100 replicates, and the minimization objectives set to minimize the maximum value. The advantage to this approach is that each replicate stochastic simulation is functionally equivalent to the deterministic simulation, only with different forcing data and run for a different time horizon (35 years for the stochastic case and 68 years for the deterministic case). Therefore the numeric values for the objectives can be easily compared.

5.2.5 Many Objective Search

The prior chapters demonstrated how many objective search can help stakeholders find high-quality planning alternatives to problems with conflicting objectives. In general, a successful MOEA will *converge* close to the true Pareto optimal set for

a problem while maintaining *diversity* across the full suite of objectives for the problem. In a recent review of state of the art MOEAs, Reed et al. motivated two poignant issues for MOEAs. [54]. The first is that new auto-adaptive algorithm technologies [93, 94] can help solve challenging problems by adapting their search properties without requiring input from the user. Second, emerging use of parallel computing can allow users to solve problems that would otherwise be intractable due to long function evaluation times.

This chapter builds on the studies presented in Chapters 3 and 4, introducing a new problem that runs a computationally expensive infrastructure simulation for each function evaluation. In order to solve this problem effectively, we employ a parallel version of the Borg MOEA [86,93]. The algorithm uses a steady state [95], epsilon-dominance archiving (see section 2.3.1.1 and [87]), adaptive population sizing [85], time continuation [91] and adaptive operators [94]. Several recent applications have diagnosed its superior performance relative to other state-of-the-art MOEAs [54, 86, 96].

Recall that MOEAs use a set of operators that control their selection of good candidate solutions and variation of those solutions to evolve approximations to the Pareto optimal tradeoff set. Selecting the correct operators in a MOEA can be challenging, since different operators could thrive depending on the search landscape for a given problem. The Borg MOEA overcomes this issue by employing multiple variation operators, using them in proportion to how effectively they can produce solutions in the ε-dominance archive. The operators are simulated binary crossover (SBX [219]), polynomial mutation (PM [219]), differential evolution (DE [220]), uniform mutation (UM), parent-centric crossover (PCX [221]), uniform normally distributed crossover (UNDX [222]), and simplex crossover (SPX [223]). Additionally, the Borg MOEA uses a concept termed ε-progress to increase the population size if the search has not improved by significant distances measured by the user's epsilon settings. For more information on the Borg MOEA please consult Hadka et al. [86, 93].

5.2.6 MORDM

The prior sections motivated our transition from two many objective formulations: the first using deterministic objectives from the historical timeseries and the second transitioning to an ensemble of stochastically generated streamflow and evaporation timeseries to calculate performance. A second motivating question for this study is: does the use of stochastic input data allow us to generate alternatives that are robust to deeply uncertain exogenous conditions? To explore this question, we use an MORDM analysis that subjects each candidate solution to a large ensemble of deeply uncertain conditions. We visualize the percent deviation results (i.e., the severity of changes between the worst-case performance in the ensemble and baseline conditions), and then use Scenario Discovery to determine values of the uncertainties that cause performance vulnerabilities, according to the methods introduced in chapter 4.

Our analysis of deep uncertainties for the Lower Susquehanna has two major components. The first component uses four factors to scale the system's annual data and explore several hypotheses about possible system changes. The first hypothesis relates to the persistence effects of annual data. Recent studies have indicated broad shifts in streamflow patterns in projections of climate change [131, 191], and we test the effect of these changes using a measure of the "persistence" of annual flows. By modifying the first parameter in our annual AR model, we are increasing and decreasing the effect of the previous year's flow or evaporation on the current year's condition. These shifts could have broad implications. For example, a higher level of persistence increase the probability that multiple years with uncharacteristically low flow may occur in sequence. The second hypothesis is that climate change could increase the variance in interannual flows. Shortle et al. project that streamflow variability for Pennsylvania could increase in a changing climate [125], motivating an investigation of increasing variance in our stochastic generation. Sampling these variables uniformly over a given range allows us to see a range of possible future climate trajectories, a "robustness" approach advocated by Wilby and Dessai [224]. The two modifications are employed by adding scaling factors θ to equation 5.7:

$$X_t = c + \theta_{\text{AR}}\phi X_{t-1} + N(0, \theta_{\text{Noise}}\sigma) \tag{5.14}$$

Table 5.5. Dimensions for Uncertainty Analysis

Variable	Lower Bound	Upper Bound
Inflow θ_{AR}	0.8	1.2
Inflow θ_{Noise}	0.8	1.2
Evaporation θ_{AR}	0.8	1.2
Evaporation θ_{Noise}	0.8	1.2
Peach Bottom Demand Scaling	0.8	1.2
Baltimore Demand Scaling	0.8	1.2
Environmental Flow Requirement Scaling	0.8	1.2

The second component of our deep uncertainty analysis considers the demand targets for Peach Bottom, Baltimore, and the Environmental Flow requirement. Although a set of fixed demand targets is used in both the deterministic and stochastic optimization, these demand targets might turn out to be wrong as usage patterns shift in the system. Ng [3] shows that demand uncertainty can bias the calculation of reliability performance objectives. Frederick [4] discusses the effect of climate change on modifying water demands. The list of parameters is provided in table 5.5. For both the statistical model parameters and demand scaling, each variable was sampled in a range of 0.8-1.2, representing a 20% decrease and increase respectively.

5.3 Computational Experiment

5.3.1 Stochastic Hydrology

The stochastic generation technique presented in section 5.2.3 is applied to two input variables: inflow at Marietta and evaporation at the Conowingo reservoir. The procedure of Nowak et al. [213] was originally coded in R [225] and provided by Balaji Rajagopalan at the University of Colorado.

The first step in the sampling routine is to generate models for annual data for both inflow and evaporation. The AR generation package in R takes as input the original series of data with its mean removed and chooses the best model order by minimizing the Akaike Information Criterion [226]. For inflow, the result was a zero-order "white noise" process (in other words the autoregressive parameter

in equation 5.7 $\phi = 0$), using $\sigma = 26,920$ million cubic meters/day (Mcm/d) and $c = 108,615$ Mcm/d. For the evaporation rate, the best model was an AR(1) model with $\phi = 0.1805$ and $\sigma = 0.214$ meters/day (m/d). The arima.sim package in R was then used to simulate 10,000 equally plausible timeseries using each model. This large database of plausible annual time series was sampled within the optimization evaluations, with 100 different realizations sampled during each evaluation[9].

Section 5.2.6 introduced the θ scaling factors for the autoregressive parameter and standard deviation for the simulated noise in annual data. To ensure that the θ factors could be used equally for both inflow and evaporation, a new order-1 model was trained for inflow, with $\phi = 0.0264$. In the uncertainty analysis calculation, for each parameter set, AR(1) models were generated online with the new values for the two AR(1) parameters, using the mean of the original series as an initial condition. Ranges for θ can be found in table 5.5.

5.3.2 Implementation of the Borg MOEA

A key challenge for using MOEAs is how to find the best parameterization for ensuring effective search. The Borg MOEA[10] has been demonstrated to be relatively insensitive to the choice of parameters, showing a high probability of attaining successful search no matter what the parameter setting, as long as the algorithm is run for a sufficient duration [54, 86]. The algorithm's insensitivity to parameter choice therefore justifies using the default parameterization, as suggested by Hadka et al [93] and shown in table 5.6. This parameterization was used for each of the two experiments: the deterministic formulation (section 5.2.4.2) and the stochastic formulation (section 5.2.4.3).

Initial visualizations of search progress were used to determine an appropriate run duration, as reported in table 5.6. Each algorithm run therefore was set to terminate after 100,000 function evaluations, and following previous work, each algorithm run was performed for 50 random seed replicates (i.e. each algorithm run

[9]The proportion calculations and weighted sampling were coded in C++ and performed online during each simulation. The routines utilized the Boost set of C++ libraries (http://www.boost.org)

[10]For more information, please see http://www.borgmoea.org/

Table 5.6. Borg Parameterization

Parameter	Value
Initial Population Size	100
Max Evaluations	100,000
Injection Rate	0.25
SBX Rate	1.0
SBX Distribution Index	15.0
PM Rate	$1/n_{\text{decvar}}$
PM Distribution Index	20.0
DE Crossover Rate	0.1
DE Step Size	0.5
UM Rate	$1/n_{\text{decvar}}$
PCX Number of Parents	3
PCX Number of Offspring	2
PCX Eta	0.1
PCX Zeta	0.1
UNDX Number of Parents	3
UNDX Number of Offspring	2
UNDX Eta	0.5
UNDX Zeta	0.35
SPX Number of Parents	3
SPX Number of Offspring	2
SPX Epsilon	0.5

requires $50 \times 100,000 = 5$ million function evaluations). This experimental design presents a significant computational challenge. One deterministic simulation takes about 6 seconds, so 5 million function evaluations would require almost 1 year of continuous computing. The stochastic simulation itself contains 100 replicates of the system hydrology, so it requires approximately 5 minutes per function evaluation. The same experiment using stochastic simulations therefore requires almost 48 years of continuous computing. In order to make the problems computationally tractable, this chapter uses the Ranger cluster, part of the Texas Advanced Computing Center (TACC) at the University of Texas Austin[11]. Ranger contains 62,976 cores with 123 TB of memory and 1.7 PB of disk space, with a theoretical peak performance of 579 TFLOPS. Each job used 1024 cores simultaneously with

[11]Consult http://www.tacc.utexas.edu/ for more information on HPC systems available at TACC. Some simulations were also performed on the latest system, Stampede, after Ranger was decommissioned.

Table 5.7. Epsilon Settings

Objective	Search Epsilon	Sort Epsilon
Conowingo Reliability - Recreation	0.01	0.005
Conowingo Reliability - Muddy Run Threshold	0.01	0.005
Conowingo Minimum Storage Level	0.01	0.005
Baltimore Vol. Reliability	0.001	0.005
Environmental Flow Shortage Index	0.01	0.005
Environmental Flow Vol. Reliability	0.001	0.005
Peach Bottom Vol. Reliability	0.001	0.005

varying run duration[12], for 563,200 requested hours of computation in total.

Following section 2.3.1.1, epsilon dominance is used to set the significant precision of each objective. During search, the epsilon settings are estimated such that a sufficient number of unique values are found with respect to each objective. After the 50 random seed replicate runs are completed, a separate non-dominated sorting routine is completed using the recently introduced MOEAframework package[13] using a different epsilon setting to reduce the set size [88]. The epsilon settings were the same for the stochastic and deterministic formulations and are presented in table 5.7.

5.3.3 Uncertainty Sampling and Robustness Analysis

This chapter uses the MORDM procedure developed in chapter 4 to explore the implications of deep uncertainty for the SRB test case. This section gives a brief overview of the MORDM implementation for the SRB, and for the full details the reader is encouraged to refer to Chapter 4 and Kasprzyk et al. [33]. The goal of uncertainty analysis within MORDM is to diagnose how the performance of planning alternatives is transformed if fundamental assumptions behind performance calculations are wrong. Similar to chapter 4, we construct a Latin Hypercube Sample (LHS) across different types of uncertain parameters – parameters that control the generation of stochastic hydrology, and point-valued scaling factors on assumed data. The LHS ensemble samples 10,240 values for the deeply uncertain

[12]Fifty 1-hour long jobs for the deterministic case, and 50 10 hour jobs for the stochastic case.
[13]Available at http://www.moeaframework.org.

parameters[14], and evaluates the stochastic simulation for each parameter set, for all solutions. The computational experiment for uncertainty analysis therefore requires $n_{\text{solutions}} \times 10,240$ evaluations. In chapter 4, it was tractable to simulate the entire uncertainty ensemble for each solution in serial, submitting multiple serial jobs in an HPC environment to complete the experiment. However, the SRB stochastic simulation's evaluation time of 5 minutes is several orders of magnitude larger than the LRGV simulation used in that chapter, motivating use of the TACC Ranger HPC resources and a massively parallel MORDM implementation. Simulations were performed using the C++ MPI libraries; the program runs on 1024 cores with each core doing a fraction of the uncertainty ensemble for each solution. We requested 65 jobs using 1024 cores, with duration of 10 hours each, for a total of 665,600 requested hours of computation. After the simulation ensemble is complete, calculations of percent deviation, performance thresholds, and PRIM sensitivity analysis are carried out using the procedures outlined in chapter 4.

5.4 Results

5.4.1 Comparing Formulations

The Borg MOEA was used to discover Pareto-approximate alternatives for two formulations; the first used a deterministic simulation based on the historical SRB hydrology (section 5.2.4.2) and the second used a stochastic simulation with uncertainty in inflows and evaporation (section 5.2.4.3). A main goal of this chapter is to critically compare the two alternative formulations, determining the relative merit of adding the stochastic generation of streamflow and evaporation, and later the effects of deep exogenous uncertainties. Figure 5.4 is a glyph plot that superimposes results from both formulations. The figure shows the two Pareto-approximate solution sets, with red spheres denoting solutions from the stochastic set and blue spheres denoting solutions from the deterministic set[15]. The spatial

[14]The sample size was chosen to be slightly larger than chapter 4 in order to give an even workload to each of the 1024 processors in the parallel implementation.

[15]These results came from a sort of 50 trials each of the deterministic and stochastic Borg runs. The sets were sorted together and new epsilon settings were used to reduce set size for plotting and further analysis. The deterministic set contains 189 solutions, while the stochastic set contains 368 solutions. Note that by changing the epsilon settings, a larger or smaller number

coordinates show three of the demand-based objectives in the formulation, with each sphere representing an individual design in the problems' Pareto-approximate sets. The horizontal axes plot the volumetric reliability for Baltimore water supply and Peach Bottom cooling water, while the vertical axis plots the shortage index measure for the Environmental Flow requirement. Size of the spheres plots one of the supply-based objectives, the Minimum Storage Level at Conowingo, with small spheres having lower values and large spheres having higher values[16]. Following the convention of chapters 3 and 4, arrows in the figure show the direction of increasing preference.

Each set in figure 5.4 exhibits similar qualitative properties. The sets exhibit a surface that "curves" away from the ideal region indicated by the preference arrows. This characteristic shape (see also figure 2.1) indicates a tradeoff between the water demands (i.e., it is not possible to meet all demands perfectly). Additionally, highlights (i) and (ii) show how the demand objectives conflict with achieving the desired storage targets. Large solutions at highlight (i) have preferred performance for the Minimum Storage Level objective, and good performance for Peach Bottom. However, these solutions have high values for the Environmental Flow Shortage Index, indicating that they had potentially large shortages for meeting the regulated low flow requirement during the record. The same pattern is exhibited by large and small solutions in the blue deterministic set.

Figure 5.4 is also an effective way to compare the two formulations. The figure indicates that the deterministic set was able to achieve significantly lower values for the Environmental Flow Shortage Index than the stochastic set. This result implies that there was a much lower probability of severe shortfalls in meeting the regulated low flow requirement, in solutions generated using the historical record. Each set exhibits a range of values for Volumetric Reliability for Baltimore and Peach Bottom, with the deterministic set also exhibiting slightly better results. These differences are meaningful from the standpoint of reservoir management, since the Volumetric Reliability effectively represents the portion of demands that are able to be delivered on average in the planning period (see equation 5.5). In

of solutions can be visualized according to user preference [88].

[16]Recall that high levels for Minimum Storage Level are preferred. An objective value of 0.92 means that the reservoir never went below 92% capacity, whereas a value of 0.72 means that reservoir drawdown got as low as 72% during the run.

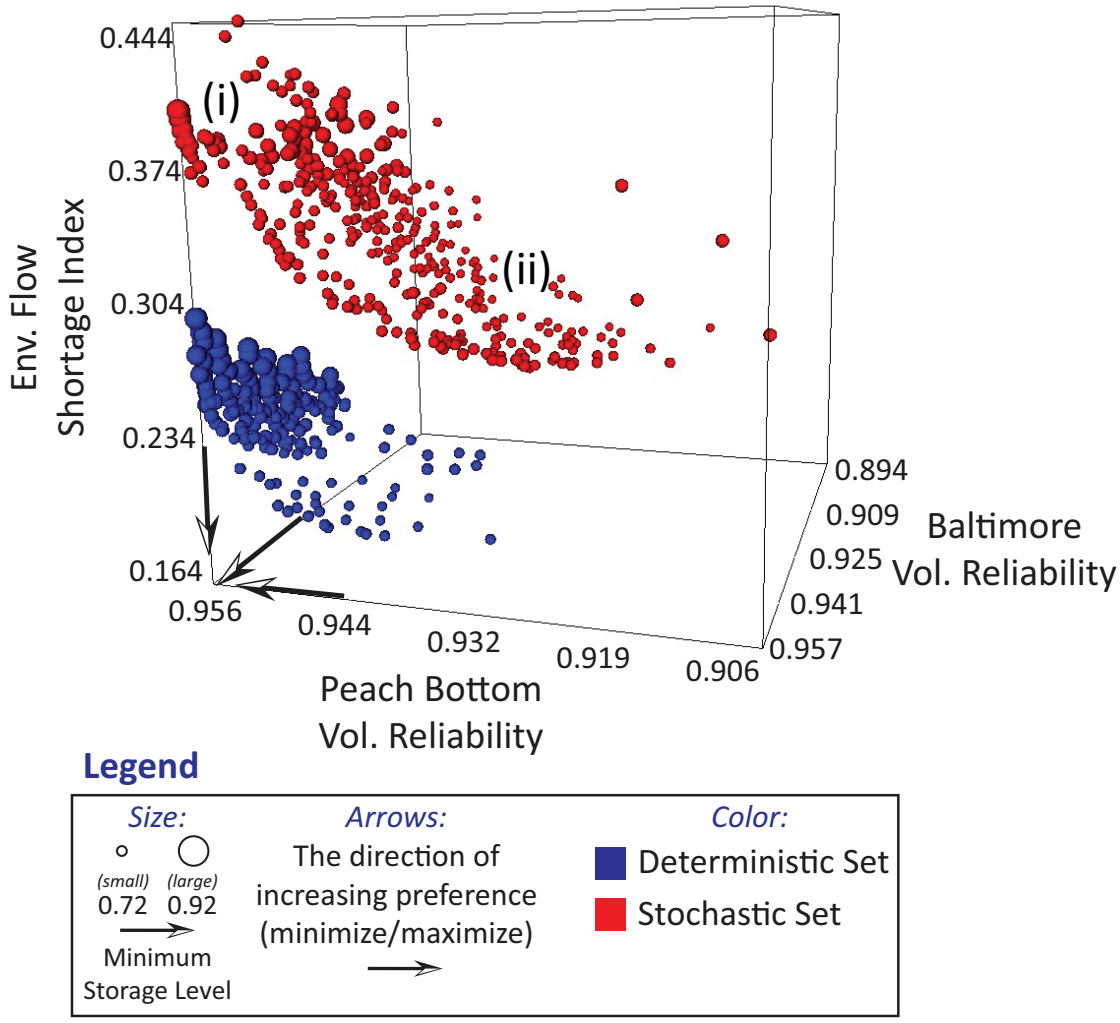

Figure 5.4. Comparison of formulations for Lower Susquehanna test case. Spatial axes show three of the demand-based objectives (Baltimore Volumetric Reliability, Peach Bottom Volumetric Reliability, and the Environmental Flow Shortage Index), and size shows the Minimum Storage Level for each solution. Blue solutions are generated using the deterministic simulation (i.e. with historical timeseries as input) whereas red solutions use the stochastic simulation.

summary, initial analysis of figure 5.4 would seem to suggest that the deterministic solutions were better able to meet environmental flow demands while maintaining high volumetric reliability for municipal water supply and thermoelectric power.

To further investigate this result, figure 5.5 shows two many objective views of the two sets using parallel coordinates. Recall that parallel coordinates are an effective tool for many objective analysis since they allow the analyst to visualize

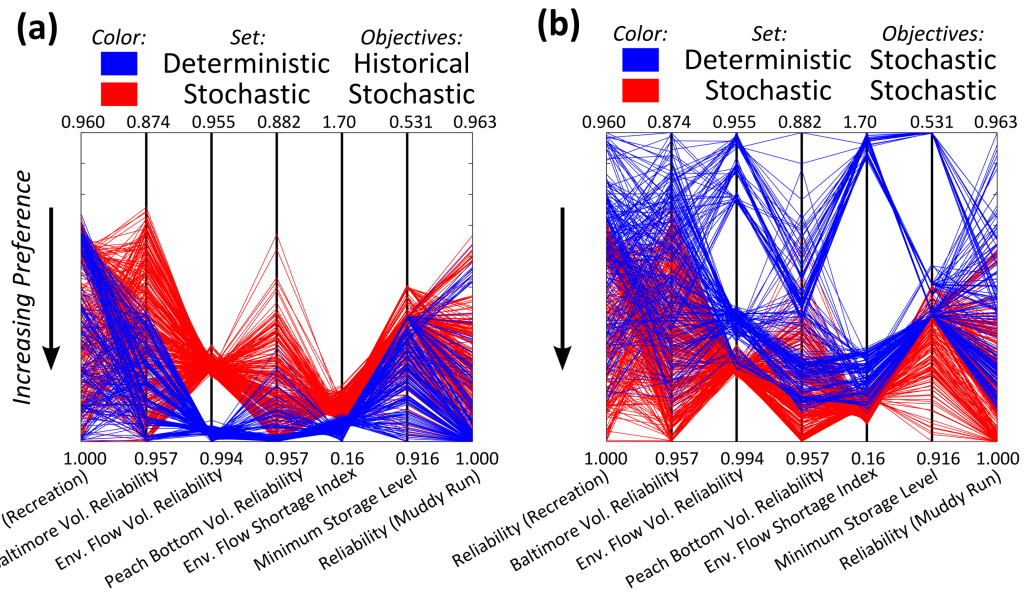

Figure 5.5. Another visualization of the Pareto approximate sets shown in figure5.4, with color retained from the previous figure. Figure 5.5a shows the sets using their respective objectives (the deterministic set shows the deterministic objectives, and so forth). Figure 5.5b subjects the deterministic set to a stochastic analysis. In each figure, lines represent each Pareto-approximate solution. Note the vertical separation in 5.5b, showing the formulation bias of the deterministic case.

the effects of all objectives simultaneously. Figure 5.5a uses the same objective calculations as figure 5.4: for the deterministic set, objective function calculations are based on the deterministic historical record, whereas the stochastic set used an ensemble of 100 realizations of a 35 year synthetic record of inflow and evaporation. Figure 5.5b uses the stochastic ensemble to evaluate decision variables from both sets. In both parallel coordinate plots, red and blue lines indicate individual solutions from the sets, with their vertical position on each axis indicating their relative objective function value. Each objective is plotted such that preferred solutions are towards the bottom of the figure. This plotting convention also allows us to show conflicts between neighboring objectives [51]. For example, solutions that tend to have preferred values for the Environmental Flow Shortage Index tend to have low values for the Minimum Storage Level objective (indicating that meeting low flow requirements often requires severe reservoir drawdowns during the planning horizon).

Figure 5.5a reinforces the result from figure 5.4, indicating better numerical

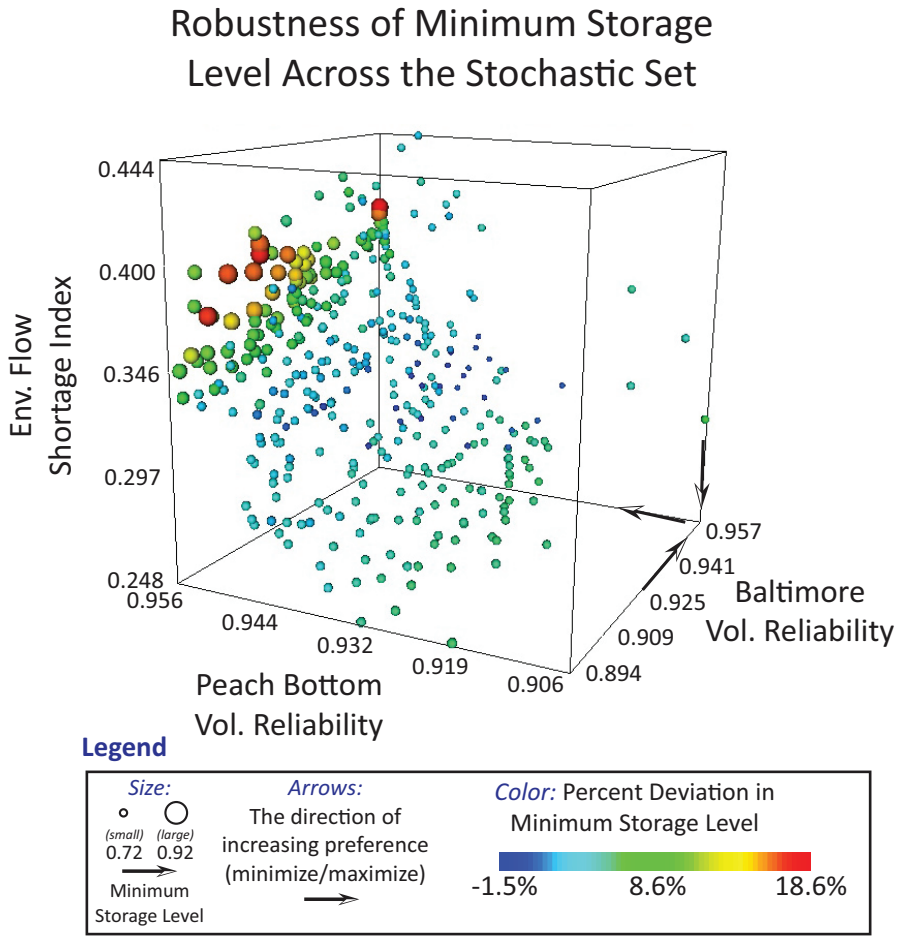

Figure 5.6. Robustness results for the Minimum Storage Level objective, for the stochastic set. Spatial axes and size are retained from figure 5.4 (four of the objectives from MOEA search). Color shows the percent deviation results of the Minimum Storage Level.

performance of the deterministic set, especially with respect to the Environmental Flow objectives. However, figure 5.5b indicates that when evaluating the deterministic set with respect to stochastic objectives, the deterministic set is biased towards severe failures on each of the demand objectives, with a wide divergence in performance with respect to the storage objectives. The Minimum Storage Level objective, for example, goes as low as 0.531 in the re-evaluated deterministic set. For this solution, this result indicates that in the worst generated inflow and evaporation timeseries, the reservoir storage level dropped to 53.1% of its capacity. In contrast, the range of performance for the stochastic set was between 0.916 and 0.72. The improved performance of the stochastic set is expected, since the solu-

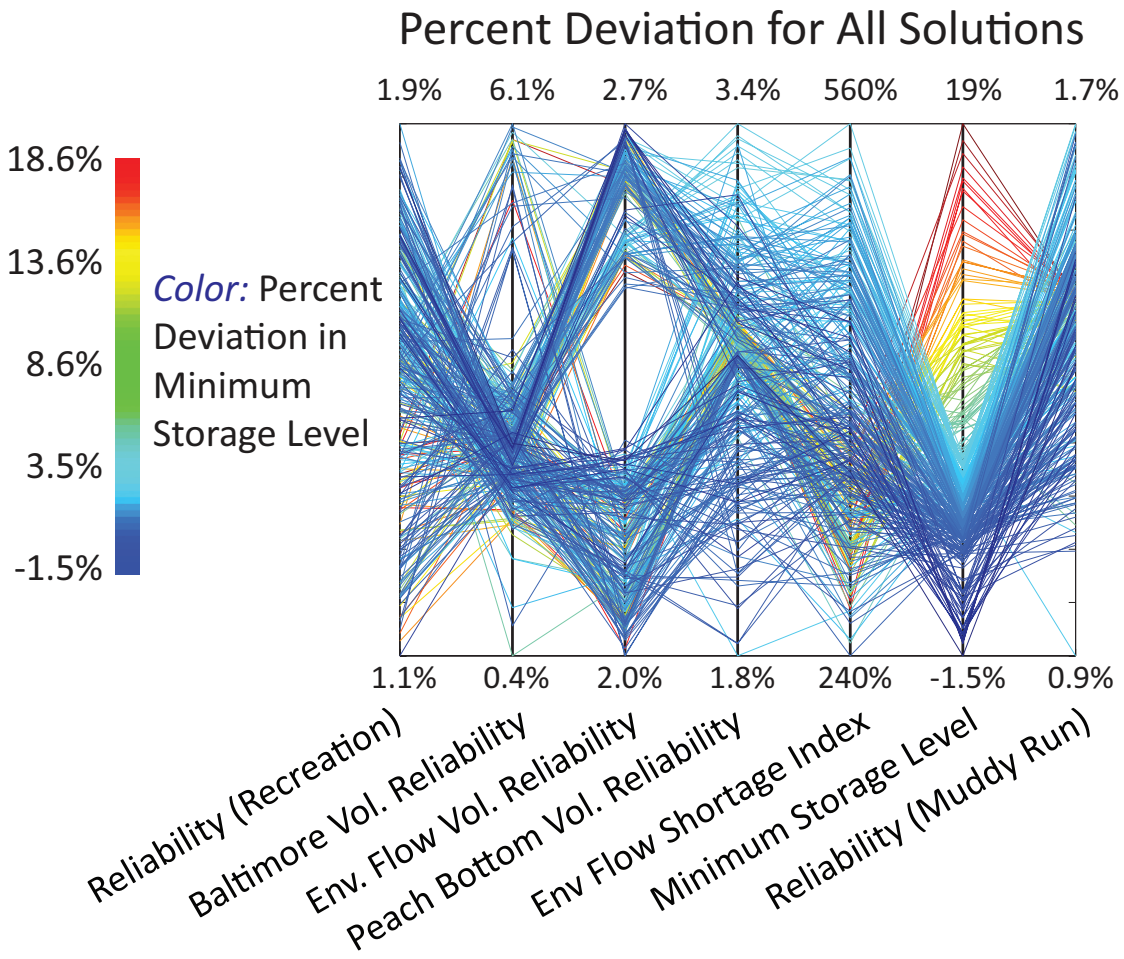

Figure 5.7. Parallel coordinate visualization of the percent deviation results, across all objectives in the stochastic set. Color is maintained from figure 5.6, showing percent deviation in the Minimum Storage Level.

tions were evaluated using the large synthetic ensemble in the MOEA. A unique contribution of figure 5.5, though, is the ability to estimate the magnitude of discrepancy between the historical formulation and its stochastic counterpart. The figure can also be used to inform selection of objectives in order to emphasize the differences. The storage reliability objectives for recreation, for example, do not change as severely between the deterministic and stochastic sets as the shortage index, because failures in the relatively high recreation storage level are less affected by extremes in the record than the other performance targets.

5.4.2 Selecting Robust Alternatives

The results in figure 5.5b explicitly show the consequence of depending on historical hydrology when seeking optimal planning tradeoffs. The results show that the deterministic, history-based planning is strongly inferior and lacks robustness when a broader range of hydrologic uncertainties are considered. Therefore, the remainder of the results focus exclusively on the stochastic set and its associated robustness to deep uncertainties. Each solution in the stochastic set was subjected to an ensemble of 10,240 randomly-generated values of the deep uncertainties in table 5.5. Following the analysis of chapter 4, the percent deviation metric (equation 4.6) is used to compare solutions' worst-case performance in the uncertainty ensemble compared to their performance in the baseline SOW. Figure 5.6 shows a representative result, using the Minimum Storage Level result to show the robustness of each solution's ability to maintain storage under deep uncertainties. Color shows percent deviation in Minimum Storage Level, between -1.5% and 18.6%[17]. Recall that solutions on the left side of the figure had good performance with respect to Minimum Storage Level in the optimization, which meant they were able to keep a high storage level in the face of conflicting demands. In the deep uncertainty ensemble, their performance degrades significantly; the percent deviation of 18% is as large as the entire range of objective performance in figure 5.4. Interestingly, smaller solutions toward the right side of the figure had poorer performance in Minimum Storage Level in the baseline condition, but their robustness was better, with less severe performance degradations.

In order to understand the solutions' multivariate performance in the MORDM ensemble, we use a parallel coordinate plot showing the percent deviation results across all objectives[18]. Figure 5.7 uses the same arrangement of axes as figure 5.5, with color plotted as the percent deviation in Minimum Storage Level to

[17]According to equation 4.6, positive values for percent deviation indicate poorer performance with respect to the baseline SOW. If an objective is to be maximized, deviation of 18.6% means the objective was 18.6% lower in the ensemble than the baseline. The small negative value here at the lower end of the range means that the performance improved slightly in the ensemble. This is possible because of the large multivariate changes in the uncertainty analysis. A solution that had poor performance in Minimum Storage level in the baseline could have actually been well served by the larger variance under deep uncertainty, not achieving such severe storage failures than in the normal condition. Later, figure 5.7 indicates that these solutions did indeed have positive percent deviation for the other objectives.

[18]For reference, this figure uses the plotting conventions of figure 4.4 in Chapter 4.

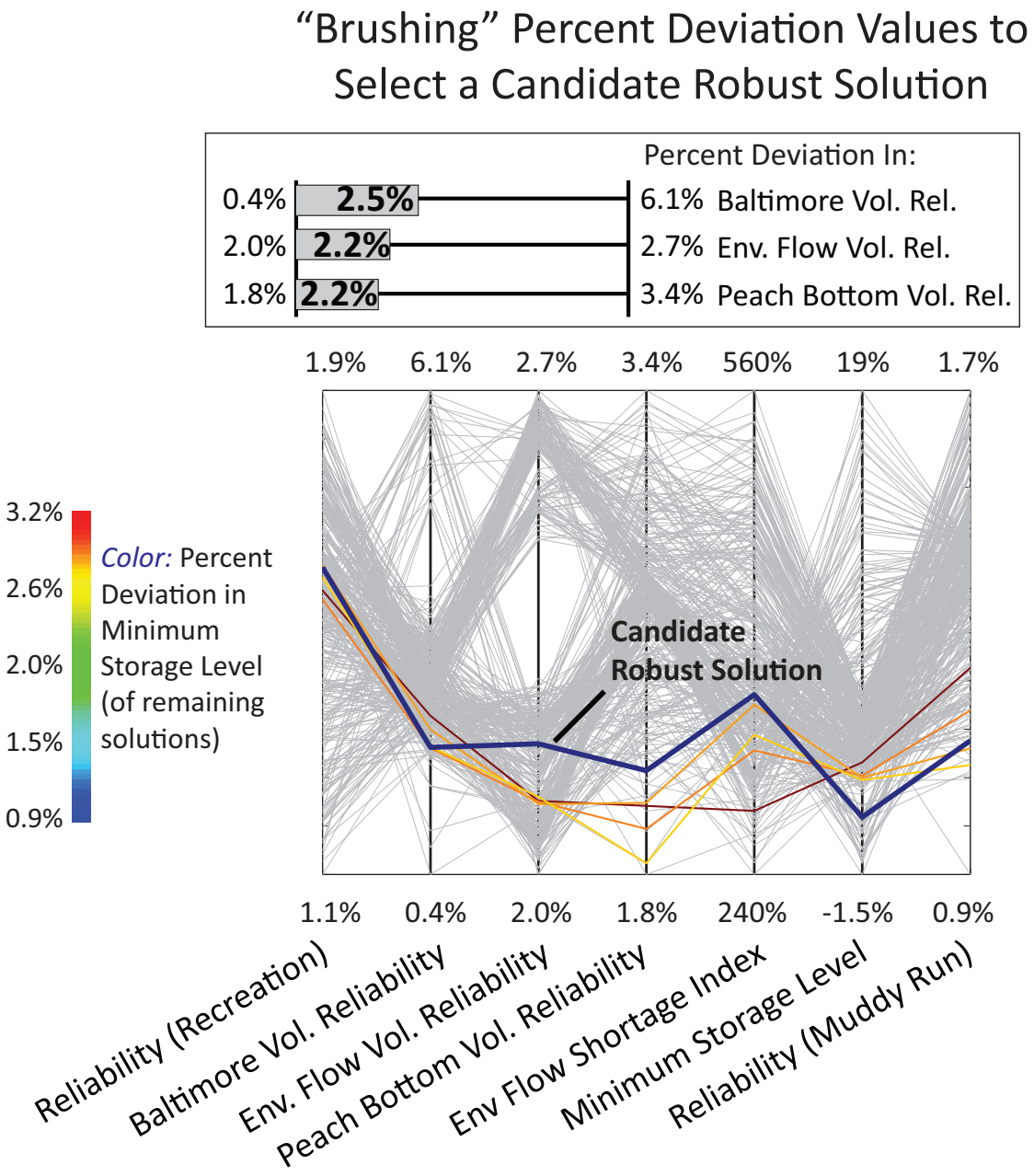

Figure 5.8. A demonstration of interactive brushing of figure 5.7, in which conditions can be imposed in order to reduce the set size. As shown with the slider bars, the colored solutions are forced to have deviations of less than 2.5%, 2.2% and 2.2% for Baltimore, Environmental Flow, and Peach Bottom Volumetric Reliability respectively. Note that values of the color scale are shifted to match the range of the remaining solutions. All other solutions are plotted in gray. The candidate robust solution is plotted with a thick line.

Table 5.8. Candidate Robust Solution Properties

Decision Variable	Value
Baltimore Reservoir Threshold	0.986
Baltimore Restriction Level	0.464
Env. Flow Reservoir Threshold	0.900
Env. Flow Restriction Level	0.414

Table 5.9. Candidate Robust Solution Performance

Baseline Objective	Value	Percent Deviation (%)
Reliability (Recreation)	0.975	1.6
Reliability (Muddy Run)	0.994	1.1
Minimum Storage Level	0.737	0.9
Baltimore Vol. Rel	0.926	1.9
Env. Flow Vol. Rel	0.985	2.2
Env. Flow SI	0.335	363

maintain continuity with figure 5.6. Some solutions with low deviation in Minimum Storage Level, plotted as blue in the figure, also have low deviations in some of the demand objectives. However, it is difficult to determine how to choose an appropriate candidate solution only using figure 5.7. Following the analysis in chapter 4, we use brushing to impose restrictions on the percent deviation results and negotiate a candidate solution to explore further. This analysis is presented in figure 5.8. Here, plotting conditions are retained from the previous figure, but three conditions are imposed on the colored solutions. Percent deviation in Baltimore's Volumetric Reliability is constrained to be less than 2.5%, and percent deviation in the Environmental Flow and Peach Bottom are both constrained to be less than 2.2%. These conditions drastically reduce the set size to consider, from 368 solutions to 5. In order to differentiate the performance of the remaining solutions, we also modify the color scale based on the data range of the remaining solutions.

The candidate robust solution is chosen as the one with the lowest Minimum Storage Level deviation from this subset. Focus on the Minimum Storage Level bounds the expectations of water managers on the severity of possible storage failures and avoids severe drawdown events that could be difficult to mitigate. Figure 5.9 provides an alternative way of viewing the selected solution. The figure retains

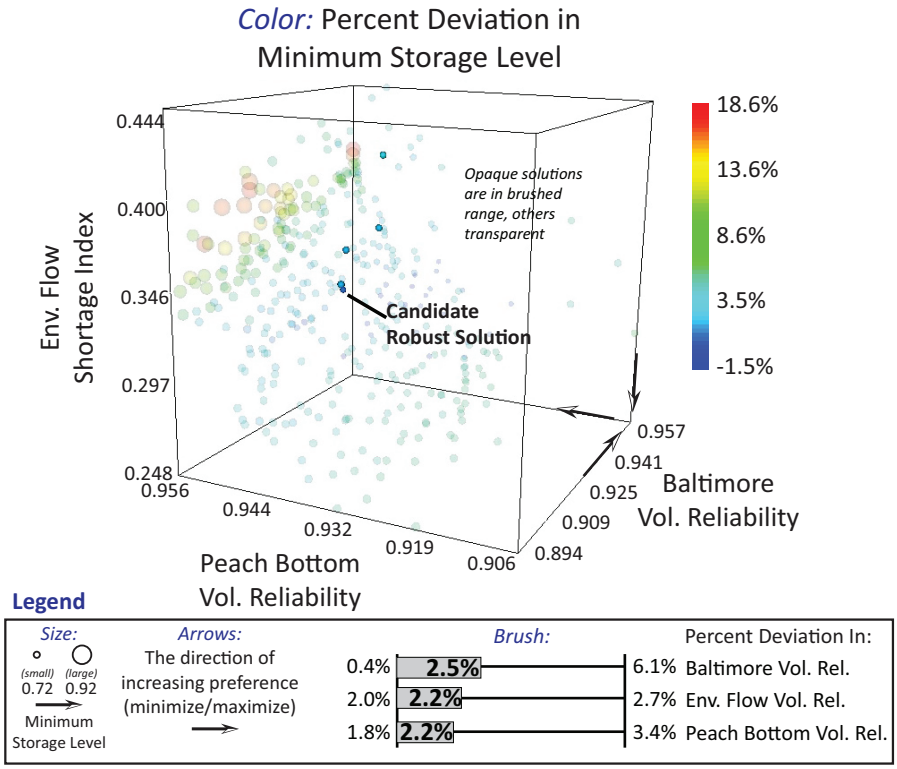

Figure 5.9. An illustration of brushing with a glyph figure. The same conditions as figure 5.8 are imposed on the set, with plotting conventions retained from figure 5.6. Note that the selected candidate solution is in the interior of the tradeoff space, which would be very difficult to isolate without interactive techniques.

Table 5.10. Threshold Set for Lower Susquehanna Storage Performance

Measure	Thresholds
Reliability (Recreation)	$<$ 10th percentile
Reliability (Muddy Run)	$<$ 10th percentile
Minimum Storage Level	$>$ 90th percentile

properties of figure 5.6, showing the percent deviation results in context of the baseline objectives. The selected solution represents a compromise in performance between maintaining high storage and meeting each demand. Solutions in this compromise region would be very difficult to distinguish using a glyph plot by itself, due to the number of points plotted. Solution properties are reported in table 5.8, and its performance both in the baseline SOW and the percentage deviation are presented in table 5.9.

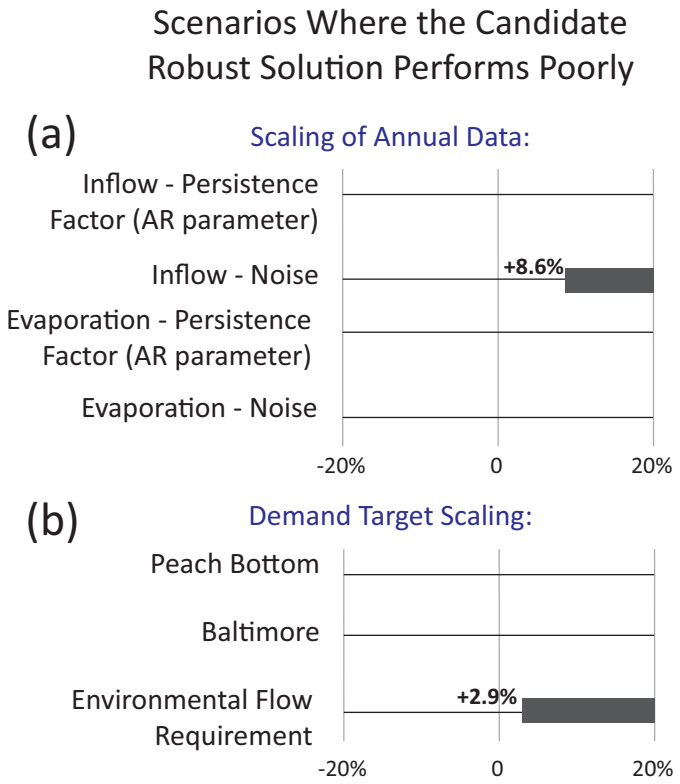

Figure 5.10. Results of scenario discovery for the candidate robust solution. Bars indicate SOW in which the identified solution performs poorly, using the threshold set in table 5.10. Each dimension in the ensemble represents a type of scaling factor that was scaled to range between a 20% increase and 20% decrease.

5.4.3 Identifying Critical Thresholds

The candidate robust solution has acceptable performance across a wide range of performance objectives. Using percent deviation to quantify robustness provides an expectation of how the solution's performance can degrade across an array of sampled exogenous uncertainties. The scenario discovery analysis in chapter 4, though, showed that there could still be important scenarios, or sets of values of the uncertainties, that nevertheless cause poor performance. Scenario discovery in this chapter will help us explore a similar result for the candidate solution. Moreover, the scenario discovery procedure will also help answer a key motivating question of this study: does solely considering hydrological uncertainties allow us to discover robust alternatives, or does the presence of deep uncertainties cause insurmountable performance failures for reservoir planning?

In initial tests, scenario discovery was initially performed for several potential threshold sets, following the procedure in chapter 4. To maximize our ability to use the diagnostic information in the PRIM-based sensitivity analysis, we chose the threshold set in table 5.10 for the final analysis. Table 5.10 uses the union of three performance measures to characterize vulnerable performance. For the candidate robust solution, this threshold set resulted in a subset of 2,046 out of 10,240 parameter sets classified as being vulnerable. PRIM was then used to characterize values of the uncertainties that caused this poor performance. The result is shown in figure 5.10. Recall that there were two types of sampled variables in the uncertainty ensemble: factors controlling the generation of streamflow and evaporation, and the demand targets. Two variables were important for causing poor performance: the "noise" or variance of annual streamflow and the scaling of the Environmental Flow requirement. For scaling of the noise in inflow, factors above 8.6% were considered significant. This is relevant because of the recently reported projections of increasing streamflow variance in Pennsylvania [125] in a report on climate change impacts. Interestingly, almost any scaling of the environmental flow requirement (increase of 2.9% and higher) caused poor performance for the identified solution.

One result of the full MORDM analysis presented in the last several sections is that we found a way to identify a good compromise between competing demands for our Lower Susquehanna system while also diagnosing the most important control on its performance. The strong importance of the environmental flow requirement, both in its response in robustness space and its importance in the Scenario Discovery analysis, indicate that it controls the ability of the system to maintain storage and consequently meet each of the other demands. According to the Conowingo management report, the minimum release flow requirements help dilute industrial and wastewater effluent downstream of the dam and also maintain the ecological integrity of wetlands surrounding the river [43]. This large demand, if not managed correctly, has the potential to penalize Conowingo's ability to maintain storage during low flow periods, as modeled in our IRAS-2010 simulation. Examining the candidate solution's properties in table 5.8, we see that the solution imposes fairly strict demand management policies on the system. The reservoir threshold for Baltimore restrictions is 98.6%, which means almost any drawdown

requires a reduction in deliveries to Baltimore. For the environmental flow restriction, the threshold is less severe at 90%. The restriction levels for Baltimore and the low flow requirement are 46.4% and 41.4% respectively, which means a substantial cut to the delivery for both demands. Even with these restrictions, the solution still has failures in the baseline stochastic ensemble. However, its performance does not substantially degrade under the deep uncertainty ensemble, unless the noise in streamflow is greatly increased or the environmental flow target is increased almost any amount. This shows the effectiveness of the stochastic formulation in the MOEA, which can generate a good compromise that is robust to deep uncertainties.

5.5 Conclusions

This chapter developed the Lower Susquehanna test case, providing an example of management for infrastructure systems in the eastern U.S. The test case combined a water resource system simulator, IRAS-2010, with the recently introduced Borg MOEA, using HPC systems to evaluate hundreds of thousands of planning alternatives for the Lower Susquehanna. The main purpose of this experiment was to determine the influence of formulation biases and deep uncertainty on infrastructure planning. Formulation biases means that results from one type of planning formulation may appear to be critically biased when evaluated in a different context. The comparative figures in section 5.4.1 showed that the deterministic case appeared to have better performance when evaluating its solutions using only the deterministic case, but its performance was severely degraded when the solutions were evaluated under the stochastic ensemble. This was especially evident with the shortage index measure, which tries to minimize the likelihood of severe demand shortfalls. Within the stochastic simulation, the performance of the best solution from the deterministic set (for shortage index) was worse than the worst solution from the stochastic set. The implication of this result is that the history-based planning approach that dominates current practice is strongly inferior to methods that can consider a broader range of hydrologic uncertainties.

This insight was further tested by using the MORDM uncertainty analysis to investigate the effect of deep uncertainties for Lower Susquehanna planning. There

were two types of deep uncertainty in this chapter – scaling of the annual data and scaling of demand targets. The ensemble was designed to capture a range of possibilities for future water availability, from the standpoint of supply (modified inflow and evaporation) and demand. MORDM in this chapter made extensive use of interactive visual analytics. First, visualizing the percent deviation shows the difference between worst-case ensemble performance and the original baseline condition, indicating a measure of robustness across the tradeoff set. Robustness with respect to a single measure may not necessarily mean that a solution is robust across all the multivariate aspects of performance. Interactive techniques are particularly helpful in these situations, allowing the analyst to change visualization axes and "brush" solutions with desired criteria. Our demonstration began with a plot of the robustness for Minimum Storage Level (figure 5.6), and used a series of interactive techniques to find a compromise solution that had low deviation values for the demand objectives. The interactive "brushing" procedure coupled with the robustness analysis is an important contribution to the methodology of using MOEAs, because it can identify solutions that would be difficult to identify using other techniques such as qualitative visual inspection of the tradeoffs. The selected solution represents a compromise between the multiple performance objectives, but it is important to use a technique to guide selection since solutions with similar objective performance may have vastly different robustness performance.

Our MORDM demonstration culminated with Scenario Discovery for an identified candidate solution. The most important parameters that controlled performance were increases in annual variance of system inflow and scaling of the environmental flow requirement. Visualizations such as figure 5.10 help stakeholders "discover" relevant scenarios for their system, or groups of uncertainties that potentially cause performance issues. But in a larger context, MORDM can be seen as a method of constructive decision aiding [227], documenting the most important part of a constructed system and motivating changes in system simulations, analysis techniques, and problem formulations. As mentioned in the introduction, our Lower Susquehanna test case was built to capture what we denoted as the most important factors for planning in the system. After the MORDM procedure, we found that the environmental low flow requirement could transform the system's ability to meet any other demand as modeled. Additionally, some of our objectives

were very sensitive to changes in formulation (from historical to stochastic). These insights would be very helpful in future work, modifying our understanding of the system and improving our ability to better manage eastern U.S. infrastructure in the face of environmental change.

Chapter 6

Concluding Remarks

6.1 Conclusions

This dissertation has explored the challenges of water supply planning given deep uncertainties, including environmental change and population growth. Two frameworks were introduced to enhance many objective planning. The first framework, *De Novo* Planning, addressed structural uncertainties in the problem formulation itself (decisions, objectives, and constraints). The second framework, Many Objective Robust Decision Making (MORDM), extended this approach to address situations in which decision makers do not know or cannot agree upon the full set of risks to their system. This section provides major conclusions of the work.

The dissertation used an advanced search tool called multiobjective evolutionary algorithms (MOEAs). The MOEAs can solve complex, uncertain planning problems with severe constraints. Tradeoff sets that come from the MOEAs show that severe conflicts exist between competing planning objectives in water systems. Our first conclusion is that *many objective analysis using MOEAs helps analysts find solutions that would be difficult or impossible to construct without a many objective view of the problem.*

For the Lower Rio Grande Valley (LRGV), minimizing the cost of a city's water supply portfolio is in direct conflict with the likelihood that the portfolio can maintain a sufficient supply of water. Adding a water market to the city's portfolio helped meet reliability targets at a lower cost. However, market solutions would have never been found if decision makers adhered to avoiding "wasted" market

transfers. The many objective LRGV problem helped planners find solutions that could meet their conflicting planning goals simultaneously. Considering surplus water as a proxy for other regional uses especially complicates the LRGV problem. The Lower Susquehanna system was introduced as a means to further explore regional planning. In the Lower Susquehanna, municipal drinking water came from the same source as water for thermoelectric cooling and environmental releases. Fully meeting municipal demands, therefore, is in conflict with water for the other sources. The many objective analysis presented for the Lower Susquehanna showed how demand management strategies could be employed to help the system better meet its conflicting demands.

The second conclusion addresses how to process the tradeoffs provided by the MOEA. *Interactive visual analytics facilitates stakeholder interaction and negotiation of promising planning solutions.* This dissertation introduced extensive use of parallel coordinate and three-dimensional glyph plots. The plots helped eludicate the importance of each objective on the performance of the system. For example, in the LRGV, there were two distinct groups of solutions that were found by plotting surplus water on a glyph plot. The first group of solutions used a large amount of reservoir rights, and the second group had drastically lower surplus water performance by using the market. The stakeholder can use interactive plots to explore these regions and learn more about solution properties in each region. Interactive plots were also used within the *de Novo* Planning calculations and in MORDM. For *de Novo* Planning, we compared objective function performance of multiple versions of a planning problem. Then, in MORDM, we used parallel coordinate plots to show solutions' percent deviation in multiple realizations of deep uncertainty.

Third, *multiple problem formulations can help our understanding of water resources systems.* Each example used multiple formulations to explore different system properties. Chapter 2 began with LRGV planning using only reservoir rights. A second formulation added market instruments for the LRGV, and the second formulation exhibited lower costs and surplus water relative to the initial formulation. In chapters 3 and 4, we looked at different configurations of decision variables. Simpler configurations may be easier to implement in practice; chapter 3 showed, though, that the simpler formulations could trade off performance

in some objectives (relative to adding more degrees of freedom in the problem). Chapter 4 further explored how the formulations perform in terms of robustness. Finally, chapter 5 used two formulations to compare historically-based planning with stochastic streamflow generation. In each example, the dissertation reinforced the idea that "problem formulation" is constantly evolving, as stakeholders learn more about the system and its properties.

In this thesis, MORDM was presented as a means to addressing deep uncertainty in modeling assumptions. Scenario Discovery (SD) within MORDM allows stakeholders to find the most important values for factors that control system performance. The fourth conclusion is that *SD techniques should be used to support ex post monitoring and adaptation for water systems.* In other words, after the planning exercise is complete, managers should be cognizant of what conditions could cause performance vulnerabilities for their systems. The Lower Susquehanna system provides an illustrative example. In the Lower Susquehanna, almost any change in environmental demands causes vulnerabilities in system storage. SD results can also motivate new planning exercises, such as new optimization or improved simulation modeling.

Each planning example in the dissertation utilized high performance computing (HPC), which will continue to be an important tool for systems analysis. The computational experiment in chapter 5 represented over 1,000,000 hours of computing, an experiment that can only be achieved using HPC systems. Advanced solution tools and model diagnostics will expand our ability to solve more complicated problems with greater fidelity as the availability of HPC systems increases.

6.2 Contributions

The first contribution of this dissertation seeks to improve many objective decision making, regardless of the problem domain. In most prior studies with MOEAs, a single problem formulation or conceptualization is used. The two frameworks introduced in this thesis, *de Novo* Planning and MORDM, provide novel ways to incorporate problem learning and deep uncertainty into systems planning. These innovations can be extended beyond the field of water supply planning to fields such as infrastructure planning, environmental management, and complex engineered

systems design.

For water supply planning in general, the dissertation has reflected the growing concern that prior supply and demand records cannot be considered fully reflective of future conditions. Using stochastic streamflow generation in combination with simulation models, the research explored the effect of deeply uncertain conditions such as the increasing incidence of severe drought.

Lastly, we sought to provide regional insights for two planning basins. In the LRGV, the work showed how a city could implement a portfolio-based approach to water supply by augmenting permanent rights with market transfers. In the Lower Susquehanna, our work demonstrated that MORDM can be used to investigate the most important aspects of a complex infrastructure network and improve the management of built infrastructure under deep uncertainties.

The contributions of this dissertation have been disseminated through several peer-reviewed journal articles. Chapters 3 and 4 were adapted from two papers published in *Environmental Modelling and Software* [32,33]. In addition, the work on planning in the Lower Susquehanna will be submitted to a journal in the summer of 2013. Additionally, the approaches have been shared with real-world decision makers[1].

Other publications, external to this dissertation, have also shown the feasibility of the methods described here. The first study that contributed to this effort was briefly discussed in Chapter 2 of this dissertation and published in *Water Resources Research* [56]. That paper contributed the first many objective analysis of the LRGV's water supply, combining volumetric objectives that sought to release water for non-municipal uses (such as ecological low-flow requirements) with economic and risk management objectives. Results of the study showed that using a market to augment the city's supply lowers costs while maintaining high reliability and lower water surplus. I was involved with two other related examples of many objective analysis for water resources systems. Many objective analysis was used to design water distribution systems [228]. Additionally, IRAS system simulations

[1]Results from the initial LRGV study were shared in a university news brief (http://live.psu.edu/story/43895). As part of a continuing conversation with the California Department of Water Resources, the MORDM framework was discussed in the 2013 Update to the California Water Plan (http://www.waterplan.water.ca.gov/cwpu2013/index.cfm, advisory committee review draft chapter 5.)

were used with a MOEA to explore London's water supply [203]. Although MOEAs have been shown to be promising tools for water resources systems analysis, there is a critical need to diagnose their successes and failures across multiple problem types. I was a co-author on a diagnostic assessment of ten MOEAs on a suite of water resources applications published in *Advances in Water Resources* [54].

6.3 Future Work

This section illustrates three possible areas of future work: improved decision support systems, extension to new problem domains, and stakeholder interaction.

6.3.1 Improved Decision Support Systems

There is a rich opportunity to provide new decision support innovations based on the work presented in this thesis. The first framework introduced in the thesis, *de Novo* Planning, explored deep uncertainties in our problem formulation (i.e., decisions, objectives, and constraints). Recently, Woodruff et al. [53] extended these ideas to compare aggregated single and two-objective formulations for a complex engineered system, extensively using visual analytics to show the decision biases in the different problems. Woodruff et al. showed that "aggregated" one and two objective versions of a problem make it difficult to generate high quality solutions. Future work within the *de Novo* Planning paradigm can continue to advance techniques to include human judgement in the computer-aided decision support process. One possible study could show the marginal improvement in including new planning instruments, objectives, and decisions. The work would address the appropriate number of objectives that should be used to fully capture decision makers' preferences. The project would also explore how including a large number of dimensions affects users' ability to sort and visualize tradeoff solutions. New technologies can likely aid this process, such as tablet computing, internet data sharing, and improved techniques for three dimensional visualization.

Future work can also extend MORDM. In MORDM, an analyst selects a robust solution that performs well across many realizations of deeply uncertain conditions. Other solutions, though, could perform better under specific conditions.

For example, states of the world (SOW) that exhibited high losses caused performance problems for the robust LRGV solution. An improvement to MORDM could use new methods to find so-called ameliorating solutions that do well under critical SOW. As mentioned in section 4.2.4, some robust decision making analyses [61, 79, 176, 177] have utilized this additional step. Another possible extension of MORDM could improve interactive visualizations for the framework. There is a large database of SOW simulations for each candidate solution in the tradeoff set. In this thesis, a percent deviation calculation distilled solutions' performance across the entire ensemble. Heat maps (i.e., figure 3.3) could be used to visualize the entire ensemble for each solution, to complement the quantitative percent deviation results. These figures would give additional information to the decision makers that would show the most important SOW and their system's performance vulnerabilities.

A third future contribution could improve uncertainty sampling within simulation models. There is a wide array of simulations used in modern engineering practice. For water resources planning, multiple subsystems (agricultural, ecological, economic, etc.) are often combined into "holistic" water systems models [158]. However, increasing the complexity of simulation models often increases the amount of computational time required to run them. In this dissertation, two quantitative simulations exhibited vastly different computational requirements[2] The computational experiments used in the dissertation could require hundreds of thousands of model evaluations, motivating the use of High Performance Computing (HPC) implementations.

Although the availability of faster HPC systems will likely continue to increase in the future[3], our ability to carry out MORDM-type analyses with more complex models may require innovations in how uncertainty sampling is performed for these systems. Bayer et al. [231] proposed a stack-ordering technique that fo-

[2]The Lower Rio Grande Valley (LRGV) simulation had function evaluations of less than 1 second, whereas the stochastic version of the Lower Susquehanna IRAS-2010 simulation took approximately 5 minutes per evaluation. This is likely a lower bound of computational complexity, because adding multiple stakeholders and a longer planning horizon will drastically increase the simulation times required.

[3]Moore's law states that the number of transistors on a circuit will double approximately every two years [229]. However, recently some have thought that the law cannot continue forever, due to concerns about power consumption and the limits of device physics [230].

cuses on critical model realizations within a Monte Carlo simulation, reducing the computational demand of the simulation by up to 90%. Kirsch et al. [123] proposed a control variate reduction technique that reduces noise in a simulation by mapping the variance in simulation inputs with the variance in the outputs. Such approaches can be incorporated into MORDM simulations in order to increase the number of model realizations that could occur in a reasonable amount of time. Eventually, future work could evaluate multiple deeply uncertain SOWs within the many objective optimization itself. This new type of robust optimization has the potential to identify better solutions compared to using one SOW realization during search.

6.3.2 New Application Areas

Coello Coello provides an excellent review of the many application areas that have been aided by MOEA optimization [46][4]. This section covers new application areas that will be explored in future work.

- *Water Resources Systems Infrastructure* Most infrastructure planning is done using expensive, proprietary licensed software. The software is often very powerful. However, such software is often not available for research, costly to obtain, and difficult to integrate with tools such as MOEAs or model diagnostics. The open source license of IRAS-2010 is a first in the water resource community. This dissertation has shown how IRAS-2010 can be combined with MORDM to improve decision making. The reasonable computational time of IRAS-2010 allows hundreds of thousands of evaluations to be performed on the HPC clusters or cloud computing. Water resources planners can easily create new IRAS-2010 input files and models to explore their own systems in this manner. For coupling to the MOEA and other tools, frameworks such as MOEAframework and the Borg MOEA are also available[5], opening up the methods presented here to a wide audience.

- *Multi-sector Planning for Wastewater Treatment and Water Supply* Effec-

[4]For example, Coello Coello [46] covers applications such as environmental engineering, robotics, structural engineering, medicine, computer science, design, scheduling, and finance.

[5]See http://www.borgMOEA.org and http://www.MOEAframework.org/

tive water resources planning requires the comprehension and optimization of complex systems. To date, the science has not adequately captured the complexity of the problem. For example, small, decentralized water treatment plants are being proposed as a solution for certain new housing developments. However, there are uncertainties about how these decentralized plants affect water quality [232]. Moreover, water quality concerns are not often included in basin-wide river planning [70]. Many objective analysis can be used with water treatment models, river basin models, and stakeholder input to improve water treatment designs.

- *Engineered Injection and Extraction (EIE) for Contaminated Groundwater* During in-situ groundwater remediation, a treatment solution is injected directly into contaminated groundwater, in order to react with the contamination and degrade it. EIE refers to a technique in which clean water is injected and extracted in surrounding wells in order to speed up the reaction [233]. Prior work has used a single system design, or sequence of injections and extractions. Many objective analysis can be used to optimally create sequences of injections and extractions that balance multiple objectives such as maximizing the amount of reaction while minimizing the required amount of treatment solution. This is a challenging problem because of the computational demands of the flow and transport simulations required to evaluate each EIE design, and the high amount of uncertainty in the groundwater flow fields and the ability to characterize reaction rates. The project is being carried out in forthcoming work [234].

6.3.3 Stakeholder Interaction

The dissertation suggested how *de Novo* Planning and MORDM can be used to make better decisions in water resources planning. The identified robust solution in chapter 4 used a simple decision rule to control the city's water acquisitions. We later identified some SOW in which the solution performed poorly. In our results, we presented visualizations of these SOW, arguing that decision makers could monitor and adapt to these conditions under the future planning horizon. Both frameworks in the thesis were designed to be interactive, where visualiza-

tions of tradeoff sets, model results, and uncertain scenarios would be shown to stakeholders. The hypothesis of the work is that *developing tradeoffs that balance multiple objectives and visualizing uncertainty can lead to better decision making.*

An important area of future research will test this hypothesis for real-world systems. Some important questions would be raised by the work. What is the best way to communicate tradeoffs (i.e., tables, two dimensional plots, three dimensional plots, narratives)? Does more information about uncertainty aid decision makers or confuse them? How can MORDM be extended to situations with multiple actors or agents negotiating with one another? The proposed project would be carried out using hand-on workshops and learning with stakeholders. The methodology can build on an extensive literature on collaborative research with stakeholders. Holling and Chambers developed approaches for resource management [235], Baker et al. performed a workshop developing future water planning scenarios [236], and Groves et al. tested different representations of climate change uncertainties for water resources planners [237]. The goal behind such a project would be to use the feedback from real-world stakeholders to motivate further extensions and improvements to the research presented here.

Bibliography

[1] Milly, P. C. D., J. Betancourt, M. Falkenmark, R. Hirsch, Z. W. Kundzewicz, D. P. Lettenmaier, and R. J. Stouffer (2008) "Stationarity is Dead: Whither Water Management?" *Science*, **319**, pp. 573–574.

[2] Kundzewicz, Z., L. Mata, N. Arnell, P. Doll, P. Kabat, B. Jimenez, K. A. Miller, T. Oki, Z. Sen, and I. Shiklomanov (2007) "Freshwater resources and their management," in *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (M. Parry, O. Canziani, J. Palutikof, P. J. van der Linden, and C. E. Hanson, eds.), Cambridge University Press, Cambridge, pp. 173–210.

[3] Ng, W. S. and G. Kuczera (1993) "Incorporating demand uncertainty in water supply headworks simulation," *Water Resources Research*, **29**(2), pp. 469–477.

[4] Frederick, K. D. (1997) "Adapting to Climate Impacts on the Supply and Demand for Water," *Climatic Change*, **37**, pp. 141–156.

[5] Characklis, G. W., R. C. Griffin, and P. B. Bedient (1999) "Improving the ability of a water market to efficiently manage drought," *Water Resources Research*, **35**(3), pp. 823–831.

[6] Rahm, B. G. and S. J. Riha (2012) "Toward strategic management of shale gas development: Regional, collective impacts on water resources," *Environmental Science and Policy*, **17**, pp. 12–23.

[7] Knight, F. H. (1921) *Risk, Uncertainty, and Profit*, Houghton Mifflin, Boston, MA.

[8] Langlois, R. N. and M. M. Cosgel (1993) "Frank Knight on Risk, Uncertainty, and the Firm: A New Interpretation," *Economic Inquiry*, **31**(3), pp. 456–465.

[9] Lempert, R. J. (2002) "A new decision sciences for complex systems," *Proceedings of the National Academy of Sciences*, **99**, pp. 7309–7313.

[10] Griffin, R. C. (1998) "The fundamental principles of cost-benefit analysis," *Water Resources Research*, **34**(8), pp. 2063–2071.

[11] Morgan, M. G., M. Kandlikar, J. Ribsey, and H. Dowlatabadi (1999) "Why conventional tools for policy analysis are often inadequate for problems of global change," *Climatic Change*, **41**, pp. 271–281.

[12] Griffin, R. C. (2008) "Benchmarking in water project analysis," *Water Resources Research*, **44**(W11418).

[13] Hoehn, J. P. and A. Randall (1989) "Too Many Proposals Pass the Benefit Cost Test," *The American Economic Review*, **79**(3), pp. 544–551.

[14] Hird, J. A. (1991) "The Political Economy of Pork: Project Selection at the U.S. Army Corps of Engineers," *The American Political Science Review*, **85**(2), pp. 429–456.

[15] Brill, E. D., Jr., J. M. Flach, L. D. Hopkins, and S. Ranjithan (1990) "MGA: A Decision Support System for Complex, Incompletely Defined Problems," *IEEE Transactions on Systems, Man, Cybernetics*, **20**(4), pp. 745–757.

[16] Banzhaf, H. S. (2009) "Objective or Multi-Objective? Two Historically Competing Visions for Benefit-Cost Analysis," *Land Economics*, **85**(1), pp. 3–23.

[17] Maass, A. (1966) "Benefit-Cost Analysis: Its Relevance to Public Investment Decisions," *The Quarterly Journal of Economics*, **80**(2), pp. 208–226.

[18] Haveman, R. H. (1967) "Benefit-Cost Analysis: Its Relevance to Public Investment Decisions: Comment," *The Quarterly Journal of Economics*, **81**(4), pp. 695–699.

[19] Maass, A. (1967) "Benefit-Cost Analysis: Its Relevance to Public Investment Decisions: Reply," *The Quarterly Journal of Economics*, **81**(4), pp. 700–702.

[20] Castle, E. N. and R. C. Youmans (1968) "Economics in Water Research and Policy," *American Journal of Agricultural Economics*, **50**(5), pp. 1655–1666.

[21] Maass, A. and D. C. Major (1970) "Economics in Water Research and Policy: Comment," *American Journal of Agricultural Economics*, **52**(1), pp. 144–145.

[22] Castle, E. N. and R. C. Youmans (1970) "Economics in Water Research and Policy: Reply," *American Journal of Agricultural Economics*, **52**(1), pp. 145–146.

[23] Nicklow, J., P. Reed, D. Savic, T. Dessalegne, L. Harrell, A. Chan-Hilton, M. Karamouz, B. Minsker, A. Ostfeld, A. Singh, and E. Zechman (2010) "State of the Art for Genetic Algorithms and Beyond in Water Resources Planning and Management," *Journal of Water Resources Planning and Management*, **136**(4), pp. 412–432.

[24] Maass, A., M. M. Hufschmidt, R. Dorfman, H. A. Thomas, Jr., S. A. Marglin, and G. M. Fair (1962) *Design of Water-Resource Systems: New Techniques for Relating Economic Objectives, Engineering Analysis, and Governmental Planning*, Harvard University Press, Cambridge.

[25] Loucks, D. P., J. R. Stedinger, and D. A. Haith (1981) *Water Resource Systems Planning and Analysis*, Prentice-Hall, Inc., Englewood Cliffs, N.J.

[26] Simonovic, S. P. (1992) "Reservoir Systems Analysis: Closing Gap Between Theory and Practice," *Journal of Water Resources Planning and Management*, **118**(3), pp. 262–280.

[27] Morgan, M. and M. Henrion (1990) *An Overview of Quantitative Policy Analysis*, Cambridge University Press, New York, pp. 16–46.

[28] National Research Council (2009) *Informing Decisions in a Changing Climate*, National Academy Press.

[29] Tsoukias, A. (2008) "From decision theory to decision aiding methodology," *European Journal of Operational Research*, **187**(1), pp. 138–161.

[30] Brown, C., W. Werick, W. Leger, and D. Fay (2011) "A Decision-Analytic Approach to Managing Climate Risks: Application to the Upper Great Lakes," *Journal of the American Water Resources Association*, **47**(3).

[31] Dixon, L., R. J. Lempert, T. LaTourrette, and R. T. Reville (2008) *The Federal Role in Terrorism Insurance: Evaluating Alternatives in an Uncertain World*, RAND, Santa Monica, CA.

[32] Kasprzyk, J. R., P. M. Reed, G. W. Characklis, and B. R. Kirsch (2012) "Many-objective de Novo water supply portfolio planning under deep uncertainty," *Environmental Modelling and Software*, **34**, pp. 87–104.

[33] Kasprzyk, J. R., S. Nataraj, P. M. Reed, and R. J. Lempert (2013) "Many Objective Robust Decision Making for Complex Environmental Systems Undergoing Change," *Environmental Modelling and Software*, **42**, pp. 55–71.

[34] Hall, W. A., W. S. Butcher, and A. Esogbue (1968) "Optimization of the Operation of a Multiple-Purpose Reservoir by Dynamic Programming," *Water Resources Research*, **4**(3), pp. 471–477.

[35] Jacoby, H. D. and D. P. Loucks (1972) "Combined Use of Optimization and Simulation Models in River Basin Planning," *Water Resources Research*, **8**(6), pp. 1401–1414.

[36] Stedinger, J. R., B. F. Sule, and D. P. Loucks (1984) "Stochastic Dynamic Porgramming Models for Reservoir Operation Optimization," *Water Resources Research*, **20**(11), pp. 1499–1505.

[37] Yeh, W. W.-G. (1985) "Reservoir Management and Operations Models: A State-of-the-Art Review," *Water Resources Research*, **21**(12), pp. 1797–1818.

[38] Lund, J. R. and M. Israel (1995) "Water transfers in water resource systems," *Journal of Water Resources Planning and Management*, **121**(2), pp. 193–205.

[39] Harou, J. J., M. Pulido-Velazquez, D. E. Rosenberg, J. Medellin-Azuara, J. R. Lund, and R. E. Howitt (2009) "Hydro-economic models: Concepts, design, applications, and future prospects," *Journal of Hydrology*, **375**, pp. 627–643.

[40] Hashimoto, T., J. R. Stedinger, and D. P. Loucks (1982) "Reliability, Resiliency and Vulnerability Criteria for Water Resource System Performance Evaluation," *Water Resources Research*, **18**(1), pp. 14–20.

[41] Rosenberg, D. E. (2009) "Integrated water resources management and modeling at multiple spatial scales in Jordan," *Water Policy*, **11**, pp. 615–628.

[42] Suen, J.-P. and J. W. Eheart (2006) "Reservoir management to balance ecosystem and human needs: Incorporating the paradigm of the ecological flow regime," *Water Resources Research*, **42**, W03417.

[43] Susquehanna River Basin Commission (2006) *Conowingo Pond Management Plan, Tech. Rep. 242*, Susquehanna River Basin Commission.

[44] Cross, N. (1989) *Engineering design methods*, Wiley, Chichester.

[45] Franssen, M. (2005) "Arrow's theorem, multi-criteria decision problems and multi-attribute preferences in engineering design," *Research in Engineering Design*, **16**, pp. 42–56.

[46] Coello Coello, C. A., G. B. Lamont, and D. A. Van Veldhuizen (eds.) (2007) *Evolutionary Algorithms for Solving Multi-Objective Problems*, Genetic and Evolutionary Computation, 2 ed., Springer, New York.

[47] Lempert, R. J., S. W. Popper, and S. C. Bankes (2003) *Shaping the next one hundred years: new methods for quantitative, long-term policy analysis*, RAND, Santa Monica, CA.

[48] Cohon, J. and D. Marks (1975) "A review and evaluation of multiobjective programming techniques," *Water Resources Research*, **11**(2), pp. 208–220.

[49] Deb, K. (2001) *Multi-objective Optimization Using Evolutionary Algorithms*, John Wiley, New York.

[50] Reed, P. M. 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), pp. 140–149.

[51] Fleming, P. J., R. C. Purshouse, and R. J. Lygoe (2005) "Many-Objective Optimization: An Engineering Design Perspective," in *EMO 2005: The Third International Conference On Evolutionary Multi-Criterion Optimization* (C. Coello Coello, A. Aguirre, and E. Zitzler, eds.), no. 3410 in Lecture Notes in Computer Science, Springer Verlag, pp. 14–32.

[52] Kollat, J. B., P. M. Reed, and R. M. Maxwell (2011) "Many-objective groundwater monitoring network design using bias-aware ensemble Kalman filtering, evolutionary optimization, and visual analytics," *Water Resources Research*, **47**.

[53] Woodruff, M. J., P. M. Reed, and T. Simpson (In-Press) "Many Objective Visual Analytics: Rethinking the Design of Complex Engineered Systems," *Structural and Multidisciplinary Optimization*.

[54] Reed, P., D. Hadka, J. Herman, J. Kasprzyk, and J. Kollat (2013) "Evolutionary Multiobjective Optimization in Water Resources: The Past, Present and Future," *Advances in Water Resources*, **51**, pp. 438–456.

[55] Zeleny, M. (2005) "The Evolution of Optimality: De Novo Programming," in *EMO 2005: The Third International Conference On Evolutionary Multi-Criterion Optimization* (C. Coello Coello, A. Aguirre, and E. Zitzler, eds.), no. 3410 in Lecture Notes in Computer Science, Springer Verlag, pp. 1–13.

[56] Kasprzyk, J. R., P. M. Reed, B. R. Kirsch, and G. W. Characklis (2009) "Managing population and drought risks using many-objective water portfolio planning under uncertainty," *Water Resources Research*, **45**.

[57] Hogarth, R. M. (1981) "Beyond Discrete Biases: Functional and Dysfunctional Aspects of Judgemental Heuristics," *Psychological Bulletin*, **90**(2), pp. 197–217.

[58] Gettys, C. and S. Fisher (1979) "Hypothesis plausibility and hypothesis generation," *Organizational Behavior and Human Performance*, **24**(1), pp. 93–110.

[59] Kahneman, D. and D. Lovallo (1993) "Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk-Taking," *Management Science*, **39**, pp. 17–31.

[60] Kollat, J. B. and P. M. Reed (2007) "A Framework for Visually Interactive Decision-making and Design using Evolutionary Multiobjective Optimization (VIDEO)," *Environmental Modelling & Software*, **22**(12), pp. 1691–1704.

[61] Lempert, R. J., D. G. Groves, S. W. Popper, and S. C. Bankes (2006) "A General, Analytic Method for Generating Robust Strategies and Narrative Scenarios," *Management Science*, **52**(4), pp. 514–528.

[62] Ellsberg, D. (1961) "Risk, Ambiguity, and the Savage Axioms," *The Quarterly Journal of Economics*, **75**(4), pp. 643–669.

[63] LeRoy, S. F. and L. D. Singell, Jr. (1987) "Knight on Risk and Uncertainty," *The Journal of Political Economy*, **95**(2), pp. 394–406.

[64] Runde, J. (1998) "Clarifying Frank Knight's discussion of the meaning of risk and uncertainty," *Cambridge Journal of Economics*, **22**, pp. 539–546.

[65] Rittel, H. and M. Webber (1973) "Dilemmas in a General Theory of Planning," *Policy Sciences*, **4**, pp. 155–169.

[66] Hitch, C. J. (1960) *On the Choice of Objectives in Systems Studies*, *Tech. Rep. P-1955*, The RAND Corporation.

[67] Liebman, J. C. (1976) "Some Simple-Minded Observations on the Role of Optimization in Public Systems Decision-Making," *Interfaces*, **6**(4), pp. 102–108.

[68] Zeleny, M. (1989) "Cognitive Equilibrium: A New Paradigm of Decision Making?" *Human Systems Management*, **8**, pp. 185–188.

[69] ——— (1981) "On the Squandering of Resources and Profits via Linear Programming," *Interfaces*, **11**(5), pp. 101–107.

[70] Reed, P. M. and J. R. Kasprzyk (2009) "Water Resources Management: The Myth, the Wicked, and the Future," *Journal of Water Resources Planning and Management*, **135**(6), pp. 411–413.

[71] Polasky, S., S. R. Carpenter, C. Folke, and B. Keeler (2011) "Decision-making under great uncertainty: environmental management in an era of global change," *Trends in Ecology and Evolution*, **26**(8), pp. 398–404.

[72] Mahmoud, M., Y. Liu, H. Hartmann, S. Stewart, T. Wagener, D. Semmens, R. Stewart, H. Gupta, D. Dominguez, F. Dominguez, D. Hulse, R. Letcher, B. Rashleigh, C. Smith, R. Street, J. Ticehurst, M. Twery, H. van Delden, R. Waldick, D. White, and L. Winter (2009) "A formal framework for scenario development in support of environmental decision-making," *Environmental Modelling and Software*, **24**, pp. 798–808.

[73] Groves, D. G. and R. J. Lempert (2007) "A new analytic method for finding policy-relevant scenarios," *Global Environmental Change*, **17**, pp. 73–85.

[74] IPCC (2000) *Emissions Scenarios: A Special Report of Working Group III of the IPCC*, Cambridge University Press, Cambridge, UK.

[75] Arnell, N. W., M. Livermore, S. Kovats, P. Levy, R. Nicholls, M. Parry, and S. Gaffin (2004) "Climate and socio-economic scenarios for global-scale climate change impacts assessments: characterising the SRES storylines," *Global Environmental Change*, **14**, pp. 3–20.

[76] Brown, C. (2010) "The End of Reliability," *Journal of Water Resources Planning and Management*, **136**(2).

[77] Hine, D. and J. W. Hall (2010) "Information gap analysis of flood model uncertainties and regional frequency analysis," *Water Resources Research*, **46**, W01514.

[78] Bryant, B. P. and R. J. Lempert (2010) "Thinking inside the box: A participatory, computer-assisted approach to scenario discovery," *Technological Forecasting and Social Change*, **77**, pp. 34–49.

[79] Hall, J., R. J. Lempert, K. Keller, A. Hackbarth, C. Mijere, and D. J. McInerney (2012) "Robust Climate Policies Under Uncertainty:

A Comparison of Robust Decision Making and Info-Gap Methods," *Risk Analysis*.

[80] Glantz, M. H. (1982) "Consequences and Responsibilities in Drought Forecasting: The Case of Yakima, 1977," *Water Resources Research*, **18**(1), pp. 3–13.

[81] Reuss, M. (2003) "Is it Time to Resurrect the Harvard Water Program?" *Journal of Water Resources Planning and Management*, **129**(5), pp. 357–361.

[82] Cohon, J. and D. Marks (1973) "Multiobjective Screening Models and Water Resource Investment," *Water Resources Research*, **9**(4), pp. 826–836.

[83] Cohon, J. L., R. L. Church, and D. P. Sheer (1979) "Generating Multiobjective Trade-Offs: An Algorithm for Bicriterion Problems," *Water Resources Research*, **15**(5), pp. 1001–1010.

[84] Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan (2002) "A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, **6**(2), pp. 182–197.

[85] Kollat, J. B. and P. M. Reed (2006) "Comparing state-of-the-art evolutionary multi-objective algorithms for long-term groundwater monitoring design," *Advances in Water Resources*, **29**(6), pp. 792–807.

[86] Hadka, D. and P. Reed (In Press) "Diagnostic Assessment of Search Controls and Failure Modes in Many-Objective Evolutionary Optimization," *Evolutionary Computation*.

[87] Laumanns, M. (2002) "Combining Convergence and Diversity in Evolutionary Multiobjective Optimization," *Evolutionary Computation*, **10**(3), pp. 263–282.

[88] Kollat, J. B. and P. M. Reed (2007) "A Computational Scaling Analysis of Multiobjective Evolutionary Algorithms in Long-Term Groundwater Monitoring Applications," *Advances in Water Resources*, **30**(3), pp. 335–353.

[89] Reed, P. M., J. B. Kollat, and V. K. Devireddy (2007) "Using interactive archives in evolutionary multiobjective optimization: A case study for long-term groundwater monitoring design," *Environmental Modelling and Software*, **22**, pp. 683–692.

[90] Harik, G., E. Cantu-Paz, D. Goldberg, and B. Miller (1997) "The gambler's ruin problem, genetic algorithms, and the sizing of populations," in *IEEE International Conference on Evolutionary Computation*, pp. 7–12.

[91] Goldberg, D. E. (2002) *The Design of Innovation: Lessons from and for Competant Genetic Algorithms*, Kluwer Academic Publishers, Boston.

[92] Tang, Y., P. Reed, and T. Wagener (2006) "How effective and efficient are multiobjective evolutionary algorithms at hydrologic model calibration?" *Hydrology and Earth System Science*, **10**, pp. 289–307.

[93] Hadka, D. and P. Reed (In Press) "Borg: An Auto-Adaptive Many-Objective Evolutionary Computing Framework," *Evolutionary Computation*.

[94] Vrugt, J. A. and B. A. Robinson (2007) "Improved evolutionary optimization from genetically adaptive multimethod search," *Proceedings of the National Academy of Science*, **104**(3), pp. 708–711.

[95] Deb, K., M. Mohan, and S. Mishra (2005) "Evaluating the Epislon-Domination based Multiobjective Evolutionary Algorithm for a Quick Computation of Pareto-optimal Solutions," *Evolutionary Computation Journal*, **13**(4), pp. 501–525.

[96] Woodruff, M., D. Hadka, P. M. Reed, and T. W. Simpson (2012) "Auto-Adaptive Search Capabilities of the New Borg MOEA: A Detailed Comparison on Alternative Product Family Design Problem Formulations," in *14th AIAA/ISSMO Multidisicplinary Analysis and Optimization Conference*.

[97] Saltelli, A., S. Tarantola, and F. Campolongo (2000) "Sensitivity Analysis as an Ingredient of Modeling," *Statistical Science*, **15**(4), pp. 377–395.

[98] Tang, Y., P. Reed, T. Wagener, and K. van Werkhoven (2007) "Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation," *Hydrology and Earth System Science*, **11**, pp. 793–817.

[99] Saltelli, A. and P. Annoni (2010) "How to avoid a perfunctory sensitivity analysis," *Environmental Modelling and Software*, **25**, pp. 1508–1517.

[100] Sobol', I. M. (1993) "Sensitivity estimates for nonlinear mathematical models and their Monte Carlo estimation," *Math. Model. Comput. Exp.*, **1**, pp. 407–417.

[101] Friedman, J. H. and N. Fisher (1999) "Bump hunting in high-dimensional data," *Stat. Comput.*, **9**, pp. 123–143.

[102] Hasofer, A. (2009) "Modern sensitivity analysis of the CESARE-Risk computer fire model," *Fire Safety Journal*, **44**, pp. 330–338.

[103] Saltelli, A., M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Galtelli, M. Saisana, and S. Tarantola (2008) *Global Sensitivity Analysis: The Primer*, Wiley.

[104] Sobol', I. M. (1967) "On the distribution of points in a cube and the approximate evaluation of integrals," *USSR Comput. Math. Math. Phys.*, **7**, pp. 86–112.

[105] Saltelli, A. (2002) "Making best use of model evaluations to compute sensitivity indices," *Computer Physics Communications*, **145**, pp. 280–297.

[106] Archer, G., A. Saltelli, and I. Sobol (1997) "Sensitivity measures, ANOVA-like techniques and the use of bootstrap," *J. Statist. Comput. Simul.*, **58**, pp. 99–120.

[107] Lempert, R. J. (2012) "Scenarios that illuminate vulnerabilities and robust responses," *Climatic Change*.

[108] Lempert, R. J., B. P. Bryant, and S. C. Bankes (2008) *Comparing Algorithms for Scenario Discovery*, *Tech. Rep. WR-557-NSF*, RAND.

[109] Bryant, B. P. (2009) *sdtoolkit: Scenario Discovery Tools to Support Robust Decision Making*, r package version 2.31.
URL `http://CRAN.R-project.org/package=sdtoolkit`

[110] Keim, D., F. Mansmann, J. Schneidewind, and H. Ziegler (2006) "Challenges in Visual Data Analysis," in *Proceedings of Information Visualization*, IEEE Computer Society, London, UK, pp. 9–16.

[111] Thomas, J. and K. A. Cook (2006) "A visual analytics agenda," *IEEE Computer Graphics and Applications*, **26**(1), pp. 10–13.

[112] Thomas, J. and J. Kielman (2009) "Challenges for visual analytics," *Information Visualization*, **8**(4), pp. 309–314.

[113] Andrienko, G., N. Andrienko, U. Demsar, D. Dransch, J. Dykes, S. Fabrikant, M. Jern, M. Kraak, H. Schumann, and C. Tominski (2010) "Space, time, and visual analytics," *International Journal of Geographical Information Science*, **24**(10), pp. 1577–1600.

[114] Woods, D. (1986) "Paradigms for intelligent decision support," in *Intelligent Decision Support in Process Environments* (E. Hollnagel, G. Mancini, and D. Woods, eds.), Springer-Verlag, New York, pp. 153–173.

[115] Woods, D. D. and E. Hollnagel (2006) *Joint Cognitive Systems: Patterns in Cognitive Systems Engineering*, Taylor and Francis, Boca Raton.

[116] Inselberg, A. (1985) "The Plane with Parallel Coordinates," *The Visual Computer*, **1**, pp. 69–91.

[117] Loughlin, D. H., S. R. Ranjithan, E. D. Brill, and J. W. Baugh (2001) "Genetic Algorithm Approaches for Addressing Unmodeled Objectives in Optimization," *Engineering Optimization*, **33**(5), pp. 549–569.

[118] Anderson, T. L. and P. J. Hill (eds.) (1997) *Water Marketing: The Next Generation*, The Political Economy Forum, Rowman and Littlefield, Lanham.

[119] Schoolmaster, F. A. (1991) "Water Marketing and Water Rights Transfers in the Lower Rio Grande Valley, Texas," *Professional Geographer*, **43**(3), pp. 292–304.

[120] Levine, G. (2007) "The Lower Rio Grande Valley: a case study of a water market area," *Paddy and Water Environment*, **5**(4), pp. 279–284.

[121] Leidner, A. J., M. E. Rister, R. D. Lacewell, and A. W. Sturdivant (2011) "The water market for the middle and lower portions of the Texas Rio Grande Basin," *Journal of the American Water Resources Association*, **47**(3), pp. 597–610.

[122] Characklis, G., B. R. Kirsch, J. Ramsey, K. Dillard, and C. T. Kelley (2006) "Developing portfolios of water supply transfers," *Water Resources Research*, **42**, W05403.

[123] Kirsch, B. R., G. W. Characklis, K. Dillard, and C. T. Kelley (2009) "More efficient optimization of long-term water supply portfolios," *Water Resources Research*, **45**, W03414.

[124] Michelsen, A. M. and R. A. Young (1993) "Optioning Agricultural Water Rights for Urban Water Supplies during Drought," *American Journal of Agricultural Economics*, **75**(4), pp. 1010–1020.

[125] Shortle, J., D. Abler, S. Blumsack, R. Crane, Z. Kaufman, M. McDill, R. Najjar, R. Ready, T. Wagener, and D. Wardrop (2009) *Pennsylvania Climate Impact Assessment, Tech. Rep. 7000-BK-DEP4252*, Pennsylvania Department of Environmental Protection.

[126] Susquehanna River Basin Commission (2000) *Susquehanna River Basin Drought Coordination Plan, Tech. Rep. 212*, Susquehanna River Basin Commission.

[127] Randall, D., L. Cleland, C. S. Kuehne, G. W. Link, and D. P. Sheer (1997) "Water Supply Planning Simulation Model Using Mixed-Integer Linear Programming "Engine"," *Journal of Water Resources Planning and Management*, **123**(2), pp. 116–124.

[128] Pearsall, S. H., B. J. McCrodden, and P. A. Townsend (2005) "Adaptive Management of Flows in the Lower Roanoke River, North Carolina, USA," *Environmental Management*, **35**(4), pp. 353–367.

[129] Sheer, D. P. and A. Dehoff (2009) "Science-Based Collaboration: Finding Better Ways to Operate the Conowingo Pond," *Journal of the American Water Works Association*, **101**(6), pp. 20–22.

[130] Moore, M., M. Pace, J. Mather, P. Murdoch, R. Howarth, C. Folt, C. Chen, H. Hemond, P. Flebbe, and C. Driscoll (1997) "Potential effects of climate change on freshwater systems of the New England/Mid Atlantic Region," *Hydrologic Processes*, **11**, pp. 925–947.

[131] DeWalle, D. R., B. R. Swistock, T. E. Johnson, and K. J. McGuire (2000) "Potential effects of climate change and urbanization on mean annual streamflow in the United States," *Water Resources Research*, **36**(9), pp. 2655–2664.

[132] Neff, R., H. Chang, C. G. Knight, R. G. Najjar, B. Yarnal, and H. A. Walker (2000) "Impact of climate variation and change on Mid-Atlantic Region hydrology and water resources," *Climate Research*, **14**, pp. 207–218.

[133] Najjar, R. G., H. Walker, P. Anderson, E. Barron, R. Bord, J. Gibson, V. Kennedy, C. Knight, J. Megonigal, R. O'Connor, C. Polsky, N. Psuty, B. Richards, L. Sorenson, E. Steele, and R. Swanson (2000) "The potential impacts of climate change on the mid-Atlantic coastal region," *Clim. Res.*, **14**, pp. 219–233.

[134] Najjar, R. G., C. Pyke, M. Adams, D. Breitburg, C. Hershner, M. Kemp, R. Howarth, M. R. Mullholland, M. Paolisso, D. Secor, K. Sellner, D. Wardrop, and R. Wood (2010) "Potential climate-change impacts on the Chesapeake Bay," *Estuarine, Coastal and Shelf Science*, **86**, pp. 1–20.

[135] Pennsylvania Department of Environmental Protection (2009) *Pennsylvania State Water Plan*, *Tech. Rep. 3010-BK-DEP4222*, Pennsylvania Department of Environmental Protection.

[136] ——— (2009) *State Water Plan Principles*, *Tech. Rep. 3010-BK-DEP4227*, Pennsylvania Department of Environmental Protection.

[137] Frederick, K. D. and G. Schwarz (1999) "Socioeconomic Impacts of Climate Change on U.S. Water Supplies," *Journal of the American Water Resources Association*, **35**(6), pp. 1563–1583.

[138] Lane, M. E., P. H. Kirshen, and R. M. Vogel (1999) "Indicators of Impacts of Global Climate Change on U.S. Water Resources," *Journal of Water Resources Planning and Management*, **125**(4), pp. 194–204.

[139] Vorosmarty, C. J., P. Green, J. Salisbury, and R. B. Lammers (2000) "Global Water Resources: Vulnerability from Climate Change and Population Growth," *Science*, **289**, pp. 284–288.

[140] Brekke, L. D., E. P. Maurer, J. D. Anderson, M. D. Dettinger, E. S. Townsley, A. Harrison, and T. Pruitt (2009) "Assessing reservoir operations risk under climate change," *Water Resources Research*, **45**(W04411).

[141] Israel, M. and J. R. Lund (1995) "Recent California Water Transfers: Implications for Water Management," *Natural Resources Journal*, **35**(1), pp. 1–32.

[142] Hadjigeorgalis, E. (2008) "Managing Drought Through Water Markets: Farmer Preferences in the Rio Grande Basin," *Journal of the American Water Resources Association*, **44**(3), pp. 594–605.

[143] Savage, L. J. (1972) *The Foundations of Statistics*, 2 ed., Dover, New York.

[144] van Werkhoven, K., T. Wagener, P. Reed, and Y. Tang (2009) "Sensitivity-guided reduction of parametric dimensionality for multi-objective calibration of watershed models," *Advances in Water Resources*, **32**, pp. 1154–1169.

[145] Stump, G. M., M. Yukish, T. W. Simpson, and E. N. Harris (2003) "Design Space Visualization and its Application to a Design by Shopping Paradigm," in *Proceedings of the ASME 2003 Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, ASME.

[146] Lotov, A. V. (2007) "Visualization of Pareto Frontier in Environmental Decision Making," in *Environmental Security in Harbors and Coastal Areas: Management Using Comparative Risk Assessment and Multi-Criteria Decision Analysis* (I. Linkov, G. A. Kiker, and R. J. Wenning, eds.), Springer, pp. 275–292.

[147] Castelletti, A., A. Lotov, and R. Soncini-Sessa (2010) "Visualization-based multi-objective improvement of environmental decision-making using linearization of response surfaces," *Environmental Modelling and Software*, **25**, pp. 1522–1564.

[148] Teich, J. (2001) *Pareto-Front Exploration with Uncertain Objectives*, no. 1993 in Lecture Notes in Computer Science, Springer-Verlag, Berlin Heidelberg, pp. 314–328.

[149] Deb, K. and H. Gupta (2006) "Introducing Robustness in Multi-Objective Optimization," *Evolutionary Computation*, **14**(4), pp. 463–494.

[150] Fowler, H. J., C. Kilsby, and P. O'Connell (2003) "Modeling the impacts of climatic change and variability on the reliability, resilience, and vulnerability of a water resource system," *Water Resources Research*, **39**(8).

[151] Kenny, J., N. Barber, S. Hutson, K. Linsey, J. Lovelace, and M. Maupin (2009) *Estimated use of water in the United States in 2005*, *Tech. Rep. Circular 1344*, U.S. Geological Survey.

[152] 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), pp. 828–845.

[153] Kollat, J. B. and P. M. Reed (2005) "The Value of Online Adaptive Search: A Performance Comparison of NSGAII, epsilon-NSGAII and epsilon-MOEA," in *EMO 2005: The Third International Conference On Evolutionary Multi-Criterion Optimization* (C. Coello Coello, A. Aguirre, and E. Zitzler, eds.), no. 3410 in Lecture Notes in Computer Science, Springer Verlag, pp. 386–398.

[154] Bromley, D. W. and B. R. Beattie (1973) "On the Incongruity of Program Objectives and Project Evaluation: An Example from the Reclamation Program," *American Journal of Agricultural Economics*, **55**(3), pp. 472–476.

[155] Sophocleous, M. (2000) "From safe yield to the sustainable development of water resources - the Kansas experience," *Journal of Hydrology*, **235**, pp. 27–43.

[156] Cooper, D. C. and G. Sehlke (2012) "Sustainability and Energy Development: Influences of Greenhouse Gas Emission Reduction Options on Water Use in Energy Production," *Environmental Science and Technology*, **46**, pp. 3509–3518.

[157] Haimes, Y. Y. (1977) *Hierarchical Analyses of Water Resources Systems, Modeling and Optimization of Large Scale Systems*, McGraw-Hill.

[158] Cai, X. (2008) "Implementation of holistic water resources-economic optimization models for river basin management - Reflective experiences," *Environmental Modelling and Software*, **23**, pp. 2–8.

[159] Rockstrom, J., W. Steffen, K. Noone, A. Persson, F. Chapin, E. Lambin, T. Lenton, M. Scheffer, C. Folke, H. Schellnhuber, B. Nykvist, C. D. Wit, T. Hughes, S. van der Leeuw, H. Rodhe, S. Sorlin, P. Snyder, R. Costanza, U. Svedin, M. Falkenmark, L. Karlberg, R. Corell, V. Fabry, J. Hansen, B. Walker, D. Liverman, K. Richardson, P. Crutzen, and J. Foley (2009) "Planetary Boundaries: Exploring the Safe Operating Space for Humanity," *Ecology and Society*, **14**(2).

[160] Liu, M. and D. M. Frangopol (2005) "Bridge Annual Maintenance Prioritization under Uncertainty by Multiobjective Combinatorial Optimization," *Computer-Aided Civil and Infrastructure Engineering*, **20**, pp. 343–353.

[161] Medaglia, A. L., S. B. Graves, and J. L. Ringuest (2007) "A multiobjective evolutionary approach for linearly constrained project selection under uncertainty," *European Journal of Operational Research*, **179**, pp. 869–894.

[162] Kourakos, G. and A. Mantoglou (2008) "Remediation of heterogeneous aquifers based on multiobjective optimization and adaptive determination of critical realizations," *Water Resources Research*, **44**, W12408.

[163] Singh, A. and B. S. Minsker (2008) "Uncertainty-based multiobjective optimization of groundwater remediation design," *Water Resources Research*, **44**(W02404, doi:10.1029/2005WR004436).

[164] Salazar Aponte, D. E., C. M. Rocco S., and B. Galvan (2009) "On Uncertainty and Robustness in Evolutionary Optimization-Based MCDM," in *EMO 2009* (M. E. et al., ed.), no. 5467 in Lecture Notes in Computer Science, Springer Verlag, pp. 51–65.

[165] Fu, G. and Z. Kapelan (2011) "Fuzzy probabilistic design of water distribution networks," *Water Resources Research*, **47**, W05538.

[166] Rosenhead, J. (1996) "What's the Problem? An Introduction to Problem Structuring Methods," *Interfaces*, **26**, pp. 117–131.

[167] Palmer, R. N. and G. W. Characklis (2009) "Reducing the costs of meeting regional water demand through risk-based transfer agreements," *Journal of Environmental Management*, **90**.

[168] Yang, Y. E., X. Cai, and D. M. Stipanovic (2009) "A decentralized optimization algorithm for multiagent system-based watershed management," *Water Resources Research*, **45**.

[169] Oberkampf, W. L., J. C. Helton, C. A. Joslyn, S. F. Wojtkiewicz, and S. Ferson (2004) "Challenge problems: uncertainty in system response given uncertain parameters," *Reliability Engineering and System Safety*, **85**(13), pp. 11 – 19.

[170] Roy, B. (1999) "Decision-aiding today: What should we expect?" in *Multicriteria Decision Making: Advances in MCDM Models, Algorithms, Theory, and Applications* (T. Gal, T. Stewart, and T. Hanne, eds.), chap. 1, Kluwer, Dordrecht, pp. 1–35.

[171] Haimes, Y. Y. and W. A. Hall (1977) "Sensitivity, Responsivity, Stability, and Irreversability as Multiple Objectives in Civil Systems," *Advances in Water Resources*, **1**(2), pp. 71–81.

[172] Efstratiadis, A. and D. Koutsoyiannis (2010) "One decade of multiobjective calibration approaches in hydrological modelling: A review," *Hydrological Sciences Journal*, **55**(1), pp. 58–78.

[173] Rudolph, G. (1998) "Evolutionary search for minimal elements in partially ordered sets," in *Evolutionary Programming IV: 7th annual conference on evolutionary programming*, pp. 345–353.

[174] Rudolph, G. and A. Agapie (2000) "Convergence properties of some multi-objective evolutionary algorithms," in *Congress on evolutionary computation (CEC 2000)*, vol. 2, pp. 1010–1016.

[175] McKay, M., R. J. Beckman, and W. J. Conover (1979) "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," *Technometrics*, **21**(2), pp. 239–245.

[176] Popper, S. W., C. Berrebi, J. Griffin, T. Light, E. Y. Min, and K. Crane (2009) *Natural gas and Israel's energy future: Near-term decisions from a strategic perspective*, *Tech. Rep. MG-927-YSNFF*, RAND.

[177] Lempert, R. J. and D. G. Groves (2010) "Identifying and evaluating robust adaptive policy responses to climate change for water management agencies in the American west," *Technological Forecasting and Social Change*, **77**, pp. 960–974.

[178] Reed, P. and J. Kollat (2012) "Save now, pay later? Multi-period many-objective groundwater monitoring design given systematic model errors and uncertainty," *Advances in Water Resources*, **35**, pp. 55 – 68.

[179] Wegman, E. J. (1990) "Hyperdimensional Data Analysis Using Parallel Coordinates," *Journal of the American Statistical Assocation*, **85**(41), pp. 664–675.

[180] Inselberg, A. (1997) "Multidimensional detective," in *Information Visualization, 1997. Proceedings., IEEE Symposium on*, IEEE, pp. 100–107.

[181] Zhu, T., M. W. Jenkins, and J. R. Lund (2005) "Estimated Impacts of Climate Warming on California Water Availability under Twelve Future Climate Scenarios," *Journal of the American Water Resources Association*, **41**(5), pp. 1027–1038.

[182] Ranjan, R. (2010) "Factors Affecting Participations in Spot and Options Markets for Water," *Journal of Water Resources Planning and Management*, **136**(4), pp. 454–462.

[183] Shah, R., P. M. Reed, and T. Simpson (2011) "Many-objective Evolutionary Optimization and Visual Analytics for Product Family Design," in *Multi-objective Evolutionary optimisation for Product Design and Manufacturing* (L. Wang, A. Ng, and K. Deb, eds.), Springer-Verlag, London.

[184] Harou, J. J., J. Medellin-Azuara, T. Zhu, S. K. Tanaka, J. R. Lund, S. Stine, M. A. Olivares, and M. W. Jenkins (2010) "Economic consequences of optimized water management for a prolonged, severe drought in California," *Water Resources Research*, **46**(W05522), W05522.

[185] Loucks, D. P., M. R. Taylor, and P. N. French (1995) *Interactive river-aquifer simulation model, program description and operating manual*, Cornell University, Ithaca, NY.

[186] Matrosov, E. S., J. J. Harou, and D. P. Loucks "A computationally efficient open-source water resource system simulator – Application to London and the Thames Basin," *Environmental Modelling and Software*.

[187] Labadie, J. W. (2004) "Optimal Operation of Multireservoir Systems: State-of-the-Art Review," *Journal of Water Resources Planning and Management*, **130**(2), pp. 93–111.

[188] Jager, H. and B. Smith (2008) "Sustainable resrvoir operation: Can we generate hydropower and preserve ecosystem values?" *River Research and Applications*, **24**, pp. 340–352.

[189] Lins, H. F. and J. R. Slack (1999) "Streamflow trends in the United States," *Geophysical Research Letters*, **26**(2), pp. 227–230.

[190] Lins, H. F. and T. A. Cohn (2011) "Stationarity: Wanted Dead or Alive?" *Journal of the American Water Resources Association*, **47**(3), pp. 475–480.

[191] Kollat, J. B., J. R. Kasprzyk, J. Wilbert O. Thomas, A. C. Miller, and D. Divoky (2012) "Estimating the Impacts of Climate Change and Population Growth on Flood Discharges in the United States," *Journal of Water Resources Planning and Management*, **138**(5), pp. 442–452.

[192] Matalas, N. C. (1967) "Mathematical Assessment of Synthetic Hydrology," *Water Resources Research*, **3**(4), pp. 937–945.

[193] Lawrance, A. J. and N. T. Kottegoda (1977) "Stochastic Modelling of Riverflow Time Series," *Journal of the Royal Statistical Society, Series A (General)*, **140**(1), pp. 1–47.

[194] Salas, J. D., J. A. Ramirez, P. Burlando, and R. A. Pielke, Sr. (2003) "Stochastic Simulation of Precipitation and Streamflow Processes," in *Handbook of Weather, Climate, and Water: Atmospheric Chemistry, Hydrology, and Societal Impacts* (T. D. Potter and B. R. Colman, eds.), pp. 607–640.

[195] Rajagopalan, B., J. D. Salas, and U. Lall (2010) "Stochastic Methods for Modeling Precipitation and Streamflow," in *Advances in Data-Based Approaches for Hydrologic Modeling and Forecasting* (B. Sivakumar and R. Berndtsson, eds.), World Scientific, Singapore, pp. 17–52.

[196] Vogel, R. M. and J. R. Stedinger (1988) "The Value of Stochastic Streamflow Models in Overyear Reservoir Design Applications," *Water Resources Research*, **24**(9), pp. 1483–1490.

[197] Augustin, N. H., L. Beevers, and W. T. Sloan (2008) "Predicting river flows for future climates using an autoregressive multinomial logit model," *Water Resources Research*, **44**, W07403.

[198] Groves, D. G., D. Yates, and C. Tebaldi (2008) "Developing and applying uncertain global change projections for regional water management planning," *Water Resources Research*, **44**(W12413).

[199] Kirsch, B. R., G. W. Characklis, and H. B. Zeff (2012) "Evaluating the Impact of Alternative Hydro-Climate Scenarios on Transfer Agreements: A Practical Improvement for Generating Synthetic Streamflows," *Journal of Water Resources Planning and Management*.

[200] Rajagopalan, B., K. Nowak, J. Prairie, M. Hoerling, B. Harding, J. Barsugli, A. Ray, and B. Udall (2009) "Water supply risk on the

Colorado River: Can management mitigate?" *Water Resources Research*, **45**, W0801.

[201] Stedinger, J. R. and M. R. Taylor (1982) "Synthetic Streamflow Generation 2. Effect of Parameter Uncertainty," *Water Resources Research*, **18**(4), pp. 919–924.

[202] Brandao, C. and R. Rodrigues (2000) "Hydrological Simulation of the Interntional Catchment of Guadiana River," *Phys. Chem. Earth (B)*, **25**(3), pp. 329–339.

[203] Matrosov, E. S., I. Huskova, J. R. Kasprzyk, J. J. Harou, and P. M. Reed (In-Review) "Many-Objective Optimization and Visual Analytics Reveal Key Planning Tradeoffs for London's Water Supply," *Journal of Hydrology*.

[204] USGS (2013), "National Water Information System data available on the World Wide Web," .
URL `http://waterdata.usgs.gov/nwis/`

[205] Exelon Generation (2012) *Application for New License for Major Water Power Project-Existing Dam, Conowingo Hydroelectric Project FERC Project Number 405*, Exelon Generation.

[206] Fredrich, A. J. and E. F. Hawkins (1969) *Hydrologic Engineering Techniques for Regional Water Resources Planning*, *Tech. Rep. TP-17*, U.S. Army Corps of Engineers Hydrologic Engineering Center.

[207] Hsu, N., W. C. Cheng, W. M. Cheng, C. C. Wei, and W. W. G. Yeh (2008) "Optimization and capacity expansion of a water distribution system," *Advances in Water Resources*, **31**(5), pp. 776–786.

[208] Thomas, Jr., H. A. and M. B. Fiering (1962) "Mathematical Synthesis of Streamflow Sequences for the Analysis of River Basins by Simulation," in *Design of Water-Resource Systems: New Techniques for Relating Economic Objectives, Engineering Analysis, and Governmental Planning* (A. Maass, M. M. Hufschmidt, R. Dorfman, H. A. Thomas, Jr., S. A. Marglin, and G. M. Fair, eds.), Harvard University Press, Cambridge.

[209] Box, G. E. P. and G. M. Jenkins (1970) *Time series analysis: Forecasting and control*, Holden-Day, San Francisco.

[210] Lall, U. and A. Sharma (1996) "A nearest neighbor bootstrap for resampling hydrologic time series," *Water Resources Research*, **32**(3), pp. 679–693.

[211] Yates, D., S. Gangopadhyay, B. Rajagopalan, and K. Strzepek (2003) "A technique for generating regional climate scenarios using a nearest-neighbor algorithm," *Water Resources Research*, **39**(7).

[212] Prairie, J., B. Rajagopalan, U. Lall, and T. Fulp (2007) "A stochastic nonparametric technique for space-time disaggregation of streamflows," *Water Resources Research*, **43**(W03432).

[213] Nowak, K., J. Prairie, B. Rajagopalan, and U. Lall (2010) "A nonparametric stochastic approach for multisite disaggregation of annual to daily streamflow," *Water Resources Research*, **46**, W08529.

[214] Perera, B. J. C. and G. P. Codner (1996) "Reservoir Targets for Urban Water Supply Systems," *Journal of Water Resources Planning and Management*, **122**(4), pp. 270–279.

[215] Oliveira, R. and D. P. Loucks (1997) "Operating rules for multireservoir systems," *Water Resources Research*, **33**(4), pp. 839–852.

[216] Lund, J. and J. Guzman (1999) "Derived Operating Rules for Reservoirs in Series or in Parallel," *Journal of Water Resources Planning and Management*, **125**(3), pp. 143–153.

[217] Cui, L.-J. and G. Kuczera (2003) "Optimizing Urban Water Supply Headworks Using Probabilistic Search Methods," *Journal of Water Resources Planning and Management*, **129**(5), pp. 380–387.

[218] Hobbs, B. F. and A. Hepenstal (1989) "Is Optimization Optimistically Biased?" *Water Resources Research*, **25**(2), pp. 152–160.

[219] Deb, K. and R. B. Agrawal (1994) *Simulated binary crossover for continuous search space*, *Tech. Rep. IITK/ME/SMD-94027*, Indian Institute of Technology, Kanpur.

[220] Storn, R. and K. Price (1997) "Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, **11**(4), pp. 341–359.

[221] Deb, K., D. Joshi, and A. Anand (2002) "Real-coded evolutionary algorithms with parent-centric re-combination," in *Proceedings of the World Congress on Computational Intelligence*, pp. 61–66.

[222] Kita, H., I. Ono, and S. Kobayashi (1999) "Multi-parental extension of the unimodal normal distribution crossover for real-coded genetic algorithms," in *Congress on Evolutionary Computation*, pp. 1581–1588.

[223] Tsutsui, S., M. Yamamura, and T. Higuchi (1999) "Multi-parent recombination with simplex crossover in real coded genetic algorithms," in *Genetic and Evolutionary Computation Conference (GECCO 1999)*.

[224] Wilby, R. L. and S. Dessai (2010) "Robust adaptation to climate change," *Weather*, **65**(7), pp. 180–185.

[225] R Development Core Team (2010) *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.
URL `http://www.R-project.org`

[226] Akaike, H. (1974) "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, **19**(6), pp. 716–723.

[227] Roy, B. (1990) "Decision-aid and decision-making," *European Journal of Operational Research*, **45**, pp. 324–331.

[228] Fu, G., Z. Kapelan, J. Kasprzyk, and P. Reed (2012) "Optimal Design of Water Distribution Systems Using Many-Objective Visual Analytics," *Journal of Water Resources Planning and Management*.

[229] Schaller, R. (1997) "Moore's law: past, present and future," *Spectrum, IEEE*, **34**(6), pp. 52 –59.

[230] Lundstrom, M. (2003) "Moore's Law Forever?" *Science*, **299**(5604), pp. 210–211.

[231] Bayer, P., M. de Paly, and C. M. Burger (2010) "Optimization of high-reliability-based hydrological design problems by robust automatic sampling of critical model realizations," *Water Resources Research*, **46**, W05504.

[232] Weirich, S. R., J. Silverstein, and B. Rajagopalan (2011) "Effect of average flow and capacity utilization on effluent water quality from US municipal wastewater treatment facilities," *Water Research*, **45**, pp. 4279–4286.

[233] Mays, D. C. and R. M. Neupauer (2012) "Plume spreading in groundwater by stretching and folding," *Water Resources Research*, **48**, W07501.

[234] Piscopo, A. N., J. R. Kasprzyk, R. M. Neupauer, and D. C. Mays (2013) "Many-Objective Design of Engineered Injection and Extraction Sequences for In Situ Remediation of Contaminated Groundwater," in *EWRI World Water Congress 2013*, ASCE.

[235] Holling, C. S. and A. D. Chambers (1973) "Resource Science: The Nurture of an Infant," *BioScience*, **23**(1), pp. 13–20.

[236] Baker, J. P., D. W. Hulse, S. V. Gregory, D. White, J. Van Sickle, P. A. Berger, D. Dole, and N. H. Schumaker (2004) "Alternative Futures for the Willamette River Basin, Oregon," *Ecological Applications*, **14**(2), pp. 313–324.

[237] Groves, D. G., D. Knopman, R. J. Lempert, S. H. Berry, and L. Wainfan (2007) *Presenting Uncertainty About Climate Change to Water-Resource Managers: A Summary of Workshops with the Inland Empire Utilities Agency*, *Tech. Rep. TR-505-NSF*, RAND.

Vita

Joseph Robert Kasprzyk

Education:

- Ph.D. in Civil Engineering, Penn State University, May 2013
- M.S. in Civil Engineering, Penn State University, May 2009
- B.S. in Civil Engineering, Penn State University (Honors in Civil Engineering), May 2007

Professional Experience:

- Assistant Professor, Department of Civil, Environmental, and Architectural Engineering, The University of Colorado Boulder, starting Fall 2013. Adjunct appointment, Summer 2012-Fall 2013.
- Graduate Research Assistant, The Pennsylvania State University, Fall 2007-Spring 2013.
- Intern, AECOM, State College, PA, December 2009-May 2013.
- Intern, S and G Gas and Oil, Butler, PA, Term Breaks 2005-2007.

Publications:

- Kasprzyk, JR, S Nataraj, PM Reed, and RJ Lempert. 2013. "Many Objective Robust Decision Making for Complex Environmental Systems Undergoing Change" Environmental Modelling and Software. 42: 55-71. doi:10.1016/j.envsoft.2012.12.007.
- Fu, G, Z Kapelan, JR Kasprzyk, and PM Reed. 2012. "Optimal design of water distribution systems using many-objective visual analytics." Journal of Water Resources Planning and Management. Online ahead of print. doi:10.1061/(ASCE)WR.1943-5452.0000311.
- Reed, PM, D Hadka, JD Herman, JR Kasprzyk, and JB Kollat. 2012. "Evolutionary Multiobjective Optimization in Water Resources: The Past, Present, and Future." Advances in Water Resources 51: 438-456.
- Kollat, JB, JR Kasprzyk, WO Thomas, AC Miller, and D Divoky. 2012. "Estimating the Impacts of Climate Change and Population Growth on Flood Discharges in the United States" Journal of Water Resources Planning and Management 138(5): 442-452.
- Kasprzyk, JR, PM Reed, GW Characklis, and BR Kirsch. 2012. "Many-Objective de Novo Water Supply Portfolio Planning Under Deep Uncertainty." Environmental Modelling and Software 34: 87-104.
- Kasprzyk, JR, PM Reed, BR Kirsch, and GW Characklis. 2009. "Managing population and drought risks using many-objective water portfolio planning under uncertainty." Water Resources Research 45, W12401.
- Reed, PM and JR Kasprzyk. 2009. "Water Resources Management: The Myth, the Wicked, the Future." Journal of Water Resources Planning and Management 35(6): 411-413.
- Kollat, JB, PM Reed, and JR Kasprzyk. 2008. "A New Epsilon-Dominance Hierarchical Bayesian Optimization Algorithm for Large Multi-Objective Monitoring Network Design Problems." Advances in Water Resources 31: 828-845.