ERIN BARTHOLOMEW

ROBUST DECISION SUPPORT METHODS

A COMPARATIVE ANALYSIS

DELFT UNIVERSITY OF TECHNOLOGY

ROBUST DECISION SUPPORT METHODS

A COMPARATIVE ANALYSIS

by

ERIN BARTHOLOMEW

Student number:	4627237	
Project duration:	9 February 2018 – 23 August 2018	
Thesis committee:	Prof. dr. A. Verbraeck,	TU Delft, Chair
	Dr. J. Kwakkel,	TU Delft, first supervisor
	Dr. M. Warnier,	TuU Delft, second supervisor

Dedication

Summary

The executive summary of the thesis

Contents

Todo list

I

Introducing the Problem

1

Introduction

In 2000, the United Nations established the Millennium Development Goals (United Nations, 2015), eight goals that all United Nations members and several international organizations would work to achieve by 2015. In 2015, these goals were renewed and expanded under the name Sustainable Development Goals to a total of 17 objectives (General Assembly Resolution 70/1, 2015). Included in the list of objectives is eliminating poverty and hunger, ensuring clean water and sanitation, developing affordable and clean energy, improving infrastructure, addressing threats brought by climate change, encouraging sustainable living, and more (General Assembly Resolution 70/1, 2015). Each of these sustainable development goals can be considered wicked problems, as defined by Rittel and Webber (1973).

Characteristics of wicked problems include no definitive problem formulation, no immediate or ultimate test of a solution, and no single explanation for the cause of the problem or solution to address the problem (Rittel & Webber, 1973). Wicked problems also frequently demonstrate irreversible tipping point behaviors, where a rapid and significant change in a system can result from small initial changes to that system, and no way to go back after a tipping point is reached (Gladwell, 2006; Lenton, 2013). Tipping points in policy problems mean that solutions to wicked problems can be one-shot operations with no right for decision makers to be wrong (Rittel & Webber, 1973).

Based on these features, wicked problems are subject not only to traditional types of uncertainty that are traditionally addressed with probabilities and statistics, but also to cognitive, strategic, and institutional uncertainties (van Bueren, Klijn, & Koppenjan, 2003). These sources of uncertainty cannot be treated as stochastic functions, but are the result of a lack of information or agreement among stakeholders; this is known as deep uncertainty (Walker, Lempert, & Kwakkel, 2013).

Given the deeply uncertain nature of the sustainable development goals and other major policy problems, traditional methods of analysis and decision support, which seek policies that are best in the context of a specific and well-defined model (Weaver et al., 2013), almost always fall short. To address this, several methods of analysis have been developed that under the umbrella of exploratory modeling and analysis (Bankes, 1993). Instead of relying on a single model, exploratory modeling uses scenario exploration and consideration of an ensemble of models that represent different potential futures to support decision making processes (Bankes, 2002). As wicked problems have no one right solution, these ensembles can be used to produce a set of potential solutions that all meet specified performance goals (Bankes, 2002). There have been many methods developed that all attempt to discover the set of potential solutions, and many of those methods seek to find those policies that perform well over as wide a variety of potential futures as possible. These can be called robust decision support methods, and each provides a different approach to recommending satisfactory solutions for a policy problem. The question, then, becomes how to determine which method is most appropriate for a specific policy problem, which is what this research will address.

1.1 Robust Decision Support

Whereas traditional predict-then-act methods that focus on developing a policy based on optimal outcomes of predictive models, robust decision support seeks a satisficing policy that achieves good enough performance across multiple potential futures (Walker, Haasnoot, & Kwakkel, 2013). More specifically, a robust policy is one that performs well across a variety of possible future states of the world (Herman, Reed, Zeff, & Characklis, 2015; Popper, Lempert, & Bankes, 2005; Walker, Lempert, & Kwakkel, 2013). Robustness has be operationalized in many ways, and each method prioritizes different policy properties, including flexibility, minimizing risk, avoiding regret, or satisficing (McPhail et al., 2018).

Several frameworks for robust decision support have been developed, including decision scaling, info-gap, robust decision making, and robust optimization (Herman et al., 2015). The info-gap framework focuses on quantifying how far future conditions must deviate from a base future state before a selected strategy is considered to be performing poorly (Ben-Haim, 2006). However, when dealing with wicked policy problems that are deeply uncertain, it is unlikely that a base future state from which to compare will be accurate enough to assess robustness (Maier et al., 2016). Decision scaling leverages stakeholder feedback to reveal key uncertainties and improve projections made by existing models (Brown, Ghile, Laverty, & Li, 2012). However, decision scaling requires a system that is well-quantified (Brown et al., 2012) and pre-specified alternative policy solutions (Herman et al., 2015), which is rarely possible when addressing policy problems with deep uncertainty. Each of these methods, therefore, require analysis built on theoretical assumptions made about the system under consideration and solutions for that system before reaching a conclusion.

Unlike info-gap and decision scaling, both robust decision making and robust optimization support decision making for policy problems that include deep uncertainty by avoiding a priori assumptions of deeply uncertain factors.

Note: In this research, the name robust decision support methods is used to refer to the category of methods considered for comparison here. This is to avoid confusion with one of the primary methods under consideration, which is called robust decision making. However, it is common for literature to use robust decision making to refer to the entire category of methods that support the decision-making process when considering deeply uncertain problems.

1.1.1 Robust Decision Making

Early efforts to address the problems brought by wicked policy problems and deep uncertainty used ensembles of models and sensitivity analysis to explore a wider spectrum of potential futures (Bankes, 1993). This method of scenario focused planning does highlight the inherent variability in deeply uncertain policy problems. However, it does not include guidance on how to use that new knowledge to rank policy choices and make decisions (Lempert, 2002). What has become known as Robust Decision Making (RDM) builds on scenario analysis to evaluate robustness of potential policy solutions over a wide range of plausible futures (Lempert, 2002). Though RDM provides a structure for comparing policy alternatives and for discovering how changes in model properties affect each alternative's performance, it does not support the consideration of multiple conflicting objectives (Kasprzyk, Nataraj, Reed, & Lempert, 2013). As the presence of conflicting objectives is a common characteristic of wicked problems (Rittel & Webber, 1973), methods of decision support that support consideration of multiple objectives are essential to the analysis of wicked problems. Combined with multi-objective evolutionary algorithms (MOEA), which aim to solve many objective problems with four or more objectives, RDM becomes capable of handling conflicting objectives. This is known as multi-objective robust decision making (Kasprzyk et al., 2013).

Recently, multi-objective robust decision making (MORDM) was extended to address a recognized weaknesses: that the analysis must stem from a baseline scenario that is represented with a single set of input data (Watson & Kasprzyk, 2017). Doing so may yield invalid optimizations if the data changes significantly when compared to the original baseline scenario. Multi-scenario MORDM lowers that risk by optimizing under multiple baseline scenarios (Watson & Kasprzyk, 2017).

1.1.2 Robust Optimization

Different optimization methods may consider a single objective function or multiple objectives. Single-objective optimization is rarely sufficient to address policy problems with deep uncertainty, as there are almost always several conflicting objectives that must be considered. Therefore, when supporting the decision making process of deeply uncertain policy problems, multi objective optimization is preferred. Several methods have been developed to support multiple objectives, including the weighted global criterion method, goal programming, Successive Pareto Optimization, and evolutionary algorithms (Coello Coello, 2006; Marler & Arora, 2004). Each of the first three methods have significant shortcomings, including generating only one solution at a time or producing invalid results (Coello Coello, 2006). Evolutionary algorithms are able to find the set of optimal solutions with one run and aren't affected by the shape of the Pareto front (Coello Coello, 2006), and so will be the focus of multi-objective optimization methods that this research considers.

These traditional multi-objective optimization methods look for the Pareto optimal set of solutions, where each solution in the set is non-dominated compared with the others (Deb & Gupta, 2006). However, when considering deep uncertainty, optimal solutions under one potential future can, following another possible path, lead to unacceptable outcomes (Deb & Gupta, 2006; McInerney, Lempert, & Keller, 2012). Several studies have focused on single-objective optimization not for the optimal solution, but robust solutions (Branke, 1998; Parmee & Bonham, 2002; Tsutsui & Ghosh, 1997). Ideas from these approaches were then extended to develop methods for multi-objective robust optimization (MORO). The MORO method determines robustness by examining how each solution in the discovered Pareto optimal set of solutions responds to changes in key model variables (Deb & Gupta, 2006).

1.1.3 Comparing MORDM and MORO

Both MORO and MORDM have been well established in literature. Though each method has the common goal of determining policy alternatives that are optimally robust across many potential futures, the path each follows to reach this goal differs greatly. MORDM seeks robust solutions by first searching for alternative robust solutions for one or more baseline scenarios and then performs optimization to determine which alternatives perform the best (Kasprzyk et al., 2013). In contrast, MORO first determines the set of optimal solutions across the entire solution space and then performs a robustness analysis on those solutions (Hamarat, Kwakkel, Pruyt, & Loonen, 2014). There are several articles that compare concepts held by MORDM and MORO to other decision support methods. Hall et al. (2012) and Matrosov, Woods, and Harou (2013) compare RDM with the Info-Gap methodology and Roach, Kapelan, and Ledbetter (2015) compares robust optimization with Info-Gap. Kwakkel, Haasnoot, and Walker (2016) compares RDM with a method called Dynamic Adaptive Policy Pathways, which leverages robust optimization techniques to build its recommendations. These articles all compare methods by leveraging real-world case studies. Additional literature compares concepts held by several methods (Dittrich, Wreford, & Moran, 2016; Herman et al., 2015; Maier et al., 2016). However, this existing body of work does not yet compare these two methods directly or with respect to policy problems characterized by tipping points, as many wicked problems are. This represents the first gap that this research will address.

Research Gap One

There is a lack of comparative analysis of key robust decision support methods (MORDM, multi-scenario MORDM, and MORO) that may help decision-makers determine which method is most suitable to their needs.

1.2 Policies Developed with Robust Decision Support Methods

Each of these robust decision support methods have been used to develop policies with different implementation structures, especially with respect to policy adaptation. These methods have each been used to recommend sets of robust policy options that are static and do not change over time (Sözüer & Thiele, 2016; Kasprzyk et al., 2013). These methods have also been used to develop sets of robust policies that are adaptive, both through automatic responses to adaptation triggers and from manual adjustments at predefined points in time (Hamarat et al., 2014; Kwakkel, Haasnoot, & Walker, 2015; Trindade, Reed, Herman, Zeff, & Characklis, 2017).

The comparative literature listed in Section 1.1.3 that leverage real-world case studies for comparison have focused on only one policy implementation structure for the identified problem. This is the second gap in literature that this research will address.

Research Gap Two

Existing work that compares robust decision support methods focus on specific case and policy implementation formulations.

1.3 Thesis Structure

This thesis will be structured in the following way. First, the remainder of Part I: Introducing the Problem will establish the goals and methods of this research (Chapter 2), and will provide a review of the concepts that are fundamental to answering the identified research question (Chapter 3). Part II: Design and Development will establish the implementation details for the methods and problem formulations identified in Part I. Finally, Part III: Analysis and Conclusions, will discuss the results of analysis and will discuss conclusions that can be made based on these results. This part will also discuss future avenues of research related to the problem identified and results that were discovered.

2

Research Definition

2.1 Research Question

The gaps identified in Chapter 1 provide the foundation for this masters thesis. The goal will be to perform a comparative analysis that considers the tradeoffs of three methods of robust decision support: MORDM, multi-scenario MORDM, and MORO. Considered will be tradeoffs in computation, communication to stakeholders, similarity of results, and method complexity. Therefore, the research question that will be answered is the following:

Research Question

Given three alternative methods for robust decision support: MORDM, Multi-Scenario MORDM, and RO, what are the tradeoffs between methods when addressing wicked policy problems with tipping point characteristics when considering different policy implementation structures?

2.2 Approach

Of interest in this research is how each of the three selected robust decision support methods respond to wicked policy problems that feature tipping point characteristics. To investigate this, a multiple case study approach will be used, which provides a structure for developing a deeper understanding of a theoretical framework or methodology, in this case the robust decision support methods, through the analysis of a case or cases of interest (Edwards, 1998). The goal of following a structured case study approach and of considering a small number of cases will be to develop generalized conclusions of method tradeoffs that can be applied to wicked policy problems with tipping point characteristics (Seawright & Gerring, 2008). Furthermore, using multiple cases will provide a higher level of confidence in the tradeoffs discovered than a single case alone (Zainal, 2007).

A comparative approach will be incorporated into the case study to ensure rigor in when comparing methods. The comparative method, as defined by Pennings, Keman, and Kleinnijenhuis (2006) provides several guidelines to ensure rigor in a comparative analysis. This includes defining the important concepts and points of comparison, establishing the cases that will be used for comparison, and carefully developing causal statements from the established comparisons (Pennings et al., 2006).

2.2.1 Case Selection

Case selection involves seeking a representative sample of cases that include variation along the key dimension under consideration (Seawright & Gerring, 2008). In this case, the key dimension being considered is the policy implementation structure of a solution set. To facilitate method comparison and to ensure a higher level of confidence in the results obtained, two stylized policy problems will be used: the lake problem and fishery management problem. Both problems feature tipping point behavior and are commonly used in policy analysis research. To test the response of each robust decision support method to different policy implementation structures, each stylized problem will be varied to support three levels of policy intervention: a static policy that is implemented at the start of execution, a planned adaptive policy that is updated after a predefined period of time passes, and a dynamic adaptive policy that is updated after every time period.

2.2.2 Caveats

There are a few recognized pitfalls of using a case study approach to research that must be considered and guarded against. Case studies are often accused of a lack of rigor (Yin, 2012), so it will be imperative to monitor the data collection, method application, and results analysis to ensure that rigor is maintained. Second, generalization of results from the selected cases to the wider population (Zainal, 2007). Careful selection of the cases considered and clear establishment of the research objectives can increase confidence of generalizing the results found (Seawright & Gerring, 2008; Yin, 2012).

2.3 Supporting Questions

To support the primary research question, several sub-questions will be addressed. These sub-questions follow the commonly accepted IMRAD research framework of Introduction, Methods, Results, And Discussion (Nair & Nair, 2014). Leveraging such an established framework for research, provides a structure and helps to ensure that this work is understood and accepted by other researchers.

1. **Introduction**
   Beyond the research definition, this will include establishing a thorough background for the key concepts identified in the research definition
   - What are recognized methods of defining robustness?
   - What are the origins of each of the selected robust decision support methods?
   - How can robustness be determined for each selected robust decision support method?
   - What policy implementation structures are commonly recognized by policy analysis literature?
   - How have stylized policy problems been leveraged in previous research?
   - How do the lake problem and fisheries management represent tipping point characteristics of interest in this research?
2. **Methods**
   - How does the selection of MOEA affect the outcome of MORDM?
   - How should the baseline scenario(s) for MORDM and multi-scenario MORDM be determined?
   - What are definitions of each stylized policy model that represent the three essential types of policy implementation?
   - How can the conclusions for each robust decision support method be compared?
3. **Results**
   - What are the results of method application for each combination of robust decision support method and iteration of the stylized policy problem?

- Based on the points of comparison and definitions of robustness already developed, which method is shown to be most effective for each policy problem iteration?
- Are the results across both stylized problems consistent?

4. **Discussion**
   - Which method of robust decision making can be considered most suitable for each category?
   - How can the results for the specific cases be generalized to wicked problems with tipping point characteristics?

2.4 Research Methods

2.4.1 Introduction

An extensive review of the literature about each of the robust decision support methods selected will establish the intended purpose and recognized weaknesses of each method. Additional literature review will explore how robustness can defined generally and for each robust decision support method. Literature and desktop research will also be leveraged to determine how the results of each method will be compared.

2.4.2 Method

Development of iterations of each stylized policy problem will use Cython, a high-performance language written in Python. Using such a high-performance language will reduce computing time as much as possible for these computation-intensive decision support methods.

This step will also define how each robust decision support method will be executed. The research required to identify proper algorithms for each step of a method will be completed with further literature review of the relevant concepts. Flow charts will be leveraged to identify the specific order of execution for each method. Finally, methods will be implemented by leveraging open-source libraries wherever possible to avoid duplicated work and to leverage already well-tested functionality, and in code using Cython when necessary. By using Cython for coding needs whenever possible, additional computational power beyond what is available for student use should not be necessary.

Furthermore, as this research intends to compare multiple methods of robust decision support, it is important to establish how each method is evaluated and how those evaluation methods can be compared. Therefore, literature will be reviewed to determine and quantify points of comparison that will be tracked during execution of each pairing.

2.4.3 Results

The results for each stylized policy problem will first be determined in isolation, both through data analysis and visual interpretation of results. Visualizations will be built using python-based graphing libraries like Matplotlib. Comparison points established earlier will be quantified individually, using mathematical formulas and visualizations where applicable. Then, the set of comparison points will be examined across policy problem iterations and methods to determine applicable tradeoffs. After each stylized problem is compared in isolation, the results for each stylized problem will be compared to determine consistency of results and to develop a generalized theory about using robust decision support methods for problems displaying tipping point characteristics.

2.4.4 Discussion

The final step will involve several methods of communication. First, a chart that indicates which method of robust decision support is preferred given the comparison points established. Second, the development of a comprehensive thesis that describes the results of each step of research as described in this proposal will be developed. It is also possible that a scientific article that summarizes the results of comparison and presents the aforementioned chart as a method for selecting the best robust decision support method for wicked policy problems will be written. All word processing activities will use Latex and the TU Delft template for thesis documentation. The thesis will be submitted to the TU Delft repository and will be made available for download. Finally, code developed for both policy problem iterations and for method implementations will be made available through GitHub whenever possible.

2.5 Research Flow and Schedule

Describe research process and flow

Missing figure

Research Flow

3
Review of Concepts

The purpose of this chapter is to provide background knowledge and definitions of key concepts that are essential for addressing the identified research question. First, the concept of deep uncertainty is codified in Section Section 3.1. Section Section 3.2 investigates recognized definitions of robustness in literature. Then, Section Section 3.3 explores the foundations and key concepts for each of the selected robust decision support methods. Finally, the stylized policy problem identified in the research definition is discussed in Section Section 3.4, along with the relevant policy implementation structures.

3.1 Deep Uncertainty

Uncertainty is generally defined as the state of something not being known or not being completely certain (“uncertain,” 2013). The basis for uncertainty as a result of a lack of information despite growing scientific knowledge has long been discussed in philosophy and scientific research (Tannert, Elvers, & Jandrig, 2007). Modern research into the concept began with Frank Knight, who was one of the first to distinguish types of uncertainty: knowable uncertainty or Knightian risk, which can be converted easily into effective certainty through, for example, probabilities, and true uncertainty or Knightian uncertainty, which is not able to be measured (Knight, 1921). Similarly, Edward Quade developed two categories of uncertainty: stochastic and real, Stochastic uncertainty is similar to Knightian risk and can be quantified using probabilities, while real uncertainty is similar to Knightian uncertainty and is the result of unknowable futures or actions (Quade & Carter, 1989). This Knightian or real uncertainty provides the foundation for deep uncertainty. The formal definition of deep uncertainty that has become commonly accepted in policy analysis research was developed by Lempert, Popper, and Bankes (2003) to be the following:

> Deep uncertainty exists in policy problems when analysts do not know or cannot agree on one or more of the following three elements:
>
> 1. The model(s) that correctly describe relationships between key system elements that will shape the future (Knightian uncertainty)
> 2. The probability distributions that best represent uncertainties in key system elements
> 3. The manner in which to rank the desirability of potential outcomes

Given the nature of wicked policy problems, established in Chapter 1, deep uncertainty is an unavoidable obstacle to the decision support process. Whereas traditional policy analysis focuses on using a predict-then-act model to find the optimal solution, the presence of deep uncertainty means that accurate prediction and determination of what is an optimal solution become extremely difficult, if not impossible. When the focus for problems with deep uncertainty is on the search for an optimal solution, assumptions must be made in several areas:

- In the determination possible future scenarios and of their likelihood
- To define the probabilities and value ranges that describe identified uncertain variables
- To understand and operationalize what criteria an optimal solution must meet.

Each of the elements required for optimal search directly contravene the properties of deeply uncertain problems as defined above. Because the correctness of a model to describe potential futures cannot be agreed upon, there is no way to concretely determine future scenarios and their likelihood. With respect to the second point, the presence of deep uncertainty means that is no agreement on specification of uncertainties and their quantitative values, there can be no certainty about their values. And finally, because there is a lack of agreement on what makes a potential solution more desirable than another, there can be no concrete definition of what makes a solution optimal. Instead, policy making must focus on a different evaluation method that looks to satisfy stated goals instead of optimize the system at hand.

3.2 Robustness

As the search for an optimal solution is recognized as an impossible task when faced with deeply uncertain problems, policy makers have instead looked to an alternate mechanism to analyze the goodness of potential solutions: robustness (Maier et al., 2016). The concept of robustness has a long history in engineering, economics, supply chain management, and many other fields (Capano & Woo, 2017; Maier et al., 2016).

How robustness is used in the fields listed, to set up introduction in policy analysis

Robustness in policy analysis can be defined in multiple ways. First, it is possible to define robustness at a system level, where the system is most often represented as a model that describes the uncertainties, policy levers, relationships, and desired outcomes that is used to analyze the problem of interest. Robustness of a system can be defined similarly to robustness of natural systems. A robust system has the ability to maintain its functions despite changes in internal or external factors (Capano & Woo, 2017). Defined in this way, robustness becomes distinct from other possible measures of system performance: resilience, stability, and adaptability. A robust system is able to maintain functionality but is not required to maintain the same state, where a resilient or stable system is able to maintain the same state. And in contrast to robustness, adaptability can be considered a property of a lever or policy that helps a system maintain robustness (Capano & Woo, 2017)

Robustness can also be defined for a policy, which is generally represented as a set of possible lever values. This definition of robustness will be the focus of this research. The following is a general definition of robustness as defined for potential policies of a specific system:

> A robust policy is one that performs well across a variety of possible future states of a system, due to both internal and external changes (Herman et al., 2015; Kasprzyk et al., 2013; Matrosov, Padula, & Harou, 2013; Walker, Lempert, & Kwakkel, 2013).

Instead of searching for the optimum solution, by seeking a set of solutions based on robustness, the search process will better avoid finding solutions that are overly sensitive to changes in uncertain parameter values. It is possible for the optimum solution of a system to belong to the set of solutions that are considered robust (which is known as a super-robust solution). However, it is much more common that the robust solution to have lower performance than the optimum solution given a set of uncertain parameter and decision lever values (Sniedovich, 2016). This is known as *the price of robustness* (Bertsimas & Sim, 2004).

Robustness of a policy can be analyzed from multiple perspectives, including resistance to change, avoidance of change, recovery from change and adaptability in response to change (Durach, Wieland, & Machuca, 2015; de Goede, Gremmen, & Blom-Zandstra, 2013; Twomey, 2012). Because there are several elements that may indicate a policy's robustness, there exists many established robustness metrics, each of which prioritize a different element of robustness, as listed earlier. Calculation of each of these metrics generally involve the same three elements: determination of the different decisions that could be made, outcomes of interest or performance metrics, and the scenarios or possible future states of the world that will be considered. Robustness metrics may determine performance as an absolute calculation or relatively to other policies. Each metric also employs differing levels of risk aversion: include more extreme scenarios in calculations to have a higher level of risk aversion. Finally, each metric has a different method of combining robustness calculations across scenarios for a specific policy option, including mean, standard deviation, skewness, or kurtosis (McPhail et al., 2018).

The following are robustness metrics that have been identified in previous policy and system analysis literature. The first group of methods listed below can be considered classical robustness metrics who use the absolute value of a performance measure to determine robustness, and report a robustness value with the same units as the performance measure under consideration (McPhail et al., 2018). The benefit of methods that directly communicate system performance mean that it is extremely simple to communicate the elements that make a policy alternative robust.

Minimax (strict robustness): a worst-case approach that seeks the best performance under the worst case analyzed. Worst case scenarios are often black swan types of event (Taleb, 2007), where a black swan event is one that is rare, unexpected, and has a significant impact on a system. Because black swan events are inherently rare, they do not represent a good estimate of actual performance. Minimax is therefore an extremely conservative approach to robust optimization. The value of such a conservative approach depends on the problem under consideration. If costs of a worst case scenario are extremely high, then it makes sense to consider that worst case scenario in robustness calculations. However, under many other conditions, a worst-case scenario does not lead to catastrophically high costs, so determining a policy's robustness based on performance under the worst-case scenario may lead to unnecessarily expensive or conservative solutions.

Maximax: follows a similar principle as the minimax criterion, but focuses instead on best case performance instead of worst case McPhail et al., 2018; Rosenhead, Elton, and Gupta, 1972. By focusing the extreme positive range of values for an outcome, the maximax metric may ignore catastrophic conditions in the worst case scenario, leading to significant negative problems for decision makers.

Hurwicz optimism-pessimism rule: representing the middle ground between minimax and maximax, Hurwicz rule involves a weighted average between the two metrics, with the weighting up to the decision maker to determine robustness (Rosenhead et al., 1972). Because the weighting is left up to decision makers, they are able to customize this metric according to their desired level of risk aversion (McPhail et al., 2018).

Laplace's principle of insufficient reason: by assigning equal weight to all possible scenarios, robustness is determined as the mean of all values for an outcome (Rosenhead et al., 1972). This results in a robustness metric that has an average level of risk aversion.

Other methods in this category include maximin, percentile-based skewness and peakedness, and mean-variance.

The methods listed next calculate robustness based on relative performance, but produce outputs with the same units as the considered performance metric, making communication of robustness with decision makers a straightforward process. In this case, the robust option is one in which minimizes the maximum regret (McPhail et al., 2018)

Minimax Regret and 90th percentile minimax regret: seeks to minimize regret with respect to the worst-case performance (or the 90th percentile of the worst case). Similar to the minimax metric, these metrics are conservative and have high risk avoidance.

Undesirable deviations: unlike minimax regret, this metric determines robustness as the deviation of the mean of the lower 50th percent of performance values from th median performance, given the set of performance measures over an ensemble of scenarios (Kwakkel, Eker, & Pruyt, 2016). This method results in a level of risk aversion that is lower than that found in minimax regret and 90th percentile regret, but still maintains a relatively high level of risk aversion (McPhail et al., 2018)

Instead of determining robustness that indicate actual system performance, the final robustness calculation considered here indicates only whether the system is performing satisfactorily or not. Satisficing metrics use the value of performance measures directly, but will report a relative robustness value that indicates whether performance is satisfactory (McPhail et al., 2018). There are many satisficing metrics available, the most common of which is described below:

Starr's domain criterion: defines robustness based on the number of scenarios in which a performance measure meets a decision maker's defined threshold (Hadka, Herman, Reed, & Keller, 2015). Because the threshold for performance is determined by the decision maker, she is able to customize this metric according to her preferred level of risk aversion.

Common among these metrics is that each is defined with respect to some value or set of conditions that must be established based on the problem that is being analyzed. Therefore, a policy's robustness value is only valid under conditions specific to the system being analyzed and to the definition of robustness implemented. At the same time, selection of the appropriate metric to use is based primarily on decision maker objectives, but may also depend on limitations of the method that is used for analysis. Combining the wide variety of possible robustness metrics, each of which measure robustness in a different manner, the subjectivity of metric selection, and the fact that robustness values are only valid in the specific problem and analysis context under which they are calculated, it becomes difficult to asses the real-world robustness of a system, given a specific policy implementation.

Several studies have been completed that compare different robustness metrics given a specific problem (Giuliani & Castelletti, 2016; Herman et al., 2015; Kwakkel, Eker, & Pruyt, 2016; Roach, Kapelan, Ledbetter, & Ledbetter, 2016). These studies have generally concluded that each robustness metric indicates the robustness of a different aspect of a potential policy. This makes it difficult to compare robustness across metrics and can lead to confusion between decision makers, who must determine how to use the analysis and robustness values in their decision-making process. McPhail et al. (2018) proposes a framework that categorizes the different robustness metrics based on several factors, including required inputs, method of calculation, unit of output, and level of risk aversion. This taxonomy aims to assist decision makers in determining the most appropriate metric to use when analyzing their own problems. And while this guidance may provide a stronger foundation with which to select a specific robustness metric, selection of a specific metric will still have a significant impact on the policy recommendation process.

3.3 Robust Decision Support Methods

3.3.1 Robust Decision Making

Analysis of deeply uncertain wicked problems brings along with it several requirements. First, that policies should be analyzed with respect to robustness and not optimality (as discussed in Section 3.2). Second, that the set of potential futures cannot be represented as a small number of possibilities (given the large amount of uncertainty that is frequently influenced by multiple input variables, it is generally impossible to codify a short list of possible futures for a problem), but has to instead be described using large ensembles of potential futures, with the number of scenarios stretching anywhere from a few hundred to several million. Analysis of these problems must also lead to results that can be clearly communicated to decision makers, to ensure that any conclusions are not just used to inform decisions, but are interpreted correctly. Furthermore, wicked problems characterized by deep uncertainty commonly involve multiple decision makers and always include conflicting views on model and input specification, and on output ranking. Because of these factors, any decision making process must be iterative, providing the opportunity for feedback from decision makers and model or policy refinement based on that feedback (Tsoukiàs, 2008).

Lempert et al. (2003) proposes a prescriptive and systematic method to support decision making of problems with characteristics similar to a wicked problem known as robust decision making (RDM) that attempts to address many of the factors listed above. The proposed method involves an iterative process of model and policy specification, computer aided computational experimentation that involves the generation and execution of a large ensemble of scenarios that span the defined uncertainty space, development of interactive visualizations, and decision maker input and refinement based on the results of computational experimentation and generated visualization Lempert, Groves, Popper, and Bankes, 2006. For the purposes of this research, a description of the method flow can be found as Fig. 3.1.

The first step of any formal analysis must be to organize any relevant information into a usable form. In the case of RDM, this step is referred to as model specification. To accomplish this, Lempert et al. (2003) proposes a 4-category structure with which to organize model elements, known as the XLRM framework. ”X” refers to the exogenous uncertainties that are outside of the control of decision makers but still impact behavior of the system being considered. Variables that are controlled by decision makers are categorized under ”L”, also known as policy levers. The desired outcomes of interest are categorized as measures, or ”M”. Finally, relations (”R”) describe how each of the elements in the ”X”, ”L”, and ”M” categories relate to one another. Together, the items described by each of the 4 categories become the model that will be used for the remainder of analysis.

As a part of the structuring of data and information, the second step of the RDM process involves identifying policy alternatives, also known as candidate strategies. The selection process can manifest in many different ways. Decision makers may contribute a set of initial strategies that will be considered (Lempert et al., 2006). Alternatively, analysis can leverage the specified model to determine policies based on sensitivities to identified sources of uncertainty or decision levers through a traditional sensitivity analysis.

In step three, the XLRM specification is used to build a diverse ensemble of possible future states of

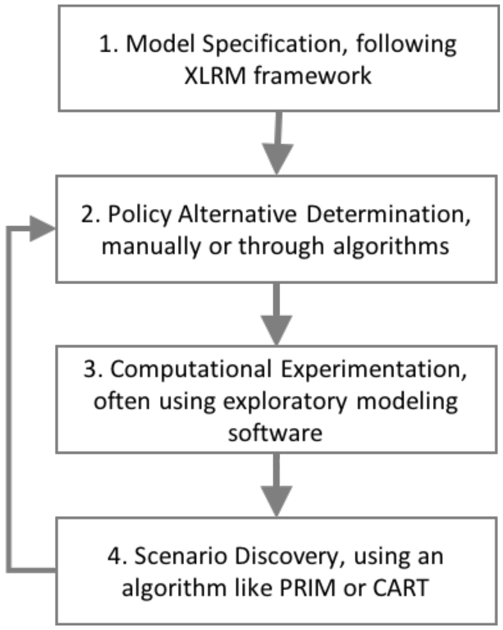

Figure 3.1: Robust Decision Making Process

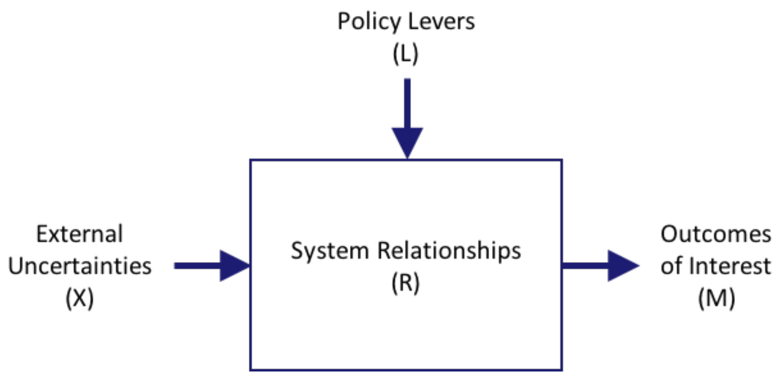

Figure 3.2: The XLRM Framework (Kwakkel, 2017)

the world (SOWs), or scenarios, that span the uncertainty space specified in the first step. The set of future SOWs are used to determine how the defined policies may react to a wide variety of possible futures. This stage generally involves the use of software to both build the ensemble of scenarios and to apply that ensemble to the set of policy options to build a dataset that evaluates the potential effectiveness of a policy given the XLRM problem definition.

Once the evaluation data is built through computational experimentation in step 3, that information is used to calculate policy robustness and to discover vulnerabilities in the existing policy options based on the identified robustness measure (possible robustness metrics were discussed in Section 3.2, but RDM traditionally favors a satisficing measure (Matrosov, Woods, & Harou, 2013)). The process of searching for vulnerabilities is known as Scenario Discovery. Prim (the Patient Rule Induction Method) is typically identified as the scenario discovery method that is most suitable for RDM, as it uses simple rules to accurately determine the factors (both uncertainties and policy settings) that are contributing to poor performance (Lempert et al., 2006). This information is used to refine both the model and policy options in an iterative process until the analyst and other decision makers reach a policy option or set of options that they can implement to address the problem.

Together, these four steps formed an iterative process that is one of the first attempts to develop a prescriptive method to guide the decision making process under conditions of deep uncertainty. The RDM method as it stands falls short in one key way. First, when there is deep uncertainty present in an analysis, there are going to be many decision makers involved who don't agree on the proper XLRM specification and who will have conflicting objectives (Walker, Lempert, & Kwakkel, 2013). Though the RDM process does support decision maker interaction throughout the process, it does not codify a formal mechanism for determining reasonable policy options when faced with conflicting objectives (Kasprzyk et al., 2013). Other methods of decision support have been developed in order to address this shortcoming, which will be reviewed next.

3.3.2 Multi-Objective Robust Decision Making (MORDM)

Building on the foundation of RDM, Kasprzyk et al. (2013) proposes a decision support framework called multi-objective robust decision making (MORDM), which provides a structure for managing multiple views from decision makers and conflicting objectives. Fig. 3.3 indicates how the MORDM process has been adapted from RDM. The most significant change, highlighted in green, is the introduction of a formal process to determine potential policy options in step 2 through the application of a multi-objective evolutionary algorithm (MOEA). The MOEA searches for potential policy solutions and ranks them based on performance of the system at a base reference point (which is determined with decision maker input). An alternative is stored if it belongs to the Pareto optimal set of all alternatives, based on that performance. By leveraging an MOEA, the MORDM process is able to quantify conflicting objectives and account for them in the policy selection phase.

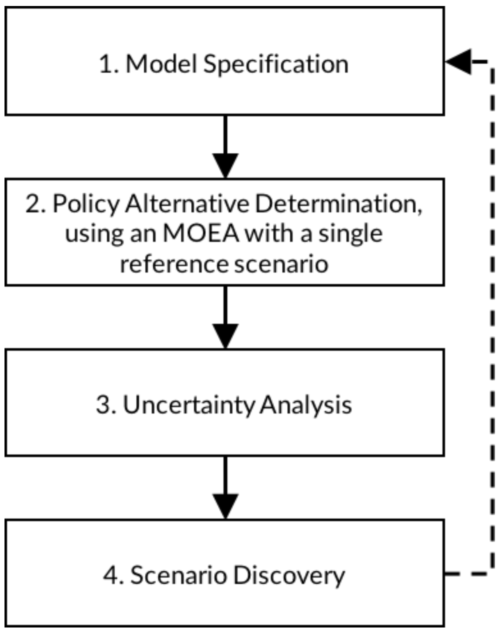

Figure 3.3: Multi Objective Robust Decision Making Process

The MORDM framework also codifies the process with which to help select a preferred solution from the set of solution alternatives generated with the MOEA, through uncertainty analysis, scenario discovery, and interactive visualizations (Kasprzyk et al., 2013). After model specification and an MOEA search for policy alternatives, the performance of the list of alternatives is tested under the recognized deep uncertainties through a process known as uncertainty analysis. This involves building a set of alternative future states of the world by sampling across the set of uncertainty parameters, a specified samapling method. Though there are many techniques available, from Monte Carlo to full and partial factorial, Kasprzyk et al. (2013) recommends using Latin Hypercube Sampling, which ensures that each member of the uncertainty set is represented evenly across the sampled set of SOWs (McKay, Beckman, & Conover, 1979). The data gathered from exercising each alternative policy with the set of SOWs provides the required information to perform a robustness analysis, in which a robustness metric is selected in a similar manner to the RDM and robustness of each policy is calculated on a per-objective basis. This information is communicated through interactive visualizations that decision makers can leverage examine the robusntess of policy alternatives and to better understand the tradeoffs that exist between conflicting objectives. At this point, candidate strategies are selected by the decision maker for further analysis, through the scenario discovery process. Scenario discovery in the MORDM framework is similar to that in the RDM framework, where candidate policy alternatives are tested to determine potential vulnerabilities (Bryant & Lempert, 2010). The MORDM framework also supports an iterative structure wherein the information gathered in the uncertainty analysis and scenario discovery process can be used to refine potential policy alternatives. However, as MORDM leverages an MOEA to determine alternatives, any refinements occur at the model specification level, where new information can be used to adjust the uncertainty space or to influence the set of available decision levers, instead of the policy alternative determination phase (see Fig. 3.1 and Fig. 3.3).

Because of the changes MORDM has made to the RDM process, the space for decision makers to impact the decision making process has also changed. Whereas RDM involves the decision maker in the early stages of the process to determine initial policy alternatives, by combining an MOEA with software-enabled uncertainty analysis and scenario discovery, the policy alternative selection process is not influenced by assumptions from decision makers until after the initial set of policy alternatives have been discovered and tested against a large ensemble of potential future SOWs. Given that problems analyzed with either RDM or MORDM frameworks are characterized by deep uncertainty, where there is conflict about the best way to achieve success, removing conflicting biases from the initial policy alternative selection process and focusing on optimizing over the set of conflicting objectives defined in the XLRM specifications directly can lead to the discovery of more robust solutions that might not have been considered otherwise.

Applications of MORDM

Since its inception in 2013, the MORDM framework has been tested and applied to several cases in policy analysis literature. As the MORDM method was designed to address challenges faced by problems relating to environmental systems management (Kasprzyk et al., 2013), many of the problems analyzed also fall within this domain (though this is not a requirement to apply MORDM). Though this is not an exhaustive list of all applications of MORDM, it provides an overview of the uses of MORDM and some of the extensions to the framework that have been developed since its inception.

In the initial proposal of MORDM, Kasprzyk et al. (2013) demonstrates the MORDM framework through a case that considers options for dealing with the water supply in the Lower Rio Grand Valley in Texas, USA and applied a robustness metric that focuses on performance in the worst-case SOW (minimax). In this case, the application of the MORDM framework was able to recommend a small and manageable set of policy alternatives, each of which includes robustness calculations for the conflicting outcomes of interest.

Herman, Zeff, Reed, and Characklis (2014) uses a case about water management in the Research Triangle region of North Carolina to demonstrate and extend the MORDM framework. Proposed extensions include to more explicitly handle analysis of deeply uncertain problems with multiple interacting decision makers and to demonstrate how to use the lessons learned in scenario discovery to improve robustness by managing uncertainties that affect policy robustness. The analysis in this research led to the discovery of key vulnerabilities of the system under consideration, and indicated which elements of a comprehensive are common among the identified robust methods and will ensure the greatest chance for success (Herman et al., 2014). This application leverages a satisficing robustness metric that seeks the best performance over a range of possible futures, as recommended in the initial RDM literature (Lempert & Collins, 2007). Satisficing in this case is defined as the ”fraction of sampled states of the world in which a solution satisfies all performance requirements” (Herman et al., 2014).

Finally, Trindade et al. (2017) uses an application of the same water management problem to extend the policy alternative search phase of the MORDM framework; a satisficing definition of robustness similar to Herman et al. (2014) is also used. This application tests potential policy alternatives against a random scenario culled from set built by sampling the uncertainty space, instead of a constant base reference scenario. The goal of this effort is to provide a broader reference space with which to determine a policy's robustness earlier in the analysis. Policy alternatives discovered using a broader set of test scenarios in the search phase of the MORDM framework were discovered to have a higher level of robustness. The set of policy alternatives was also more diverse with the modified MOEA search (Trindade et al., 2017). Together, these elements can provide decision makers with more detailed analysis of vulnerabilities and trends in recommended policies.

The use of different robustness metrics while applying MORDM indicate that the framework is not tied to a specific metric. In fact, the MORDM framework does not provide any guidance about the most appropriate robustness metric, leaving it to the decision makers and analysts to determine for themselves. Given the wide variety of robustness metrics that exist which test so many different facets of robustness (see Section 3.2), the chosen metric can have a significant impact on the recommendations made by the MORDM framework. The enhancement made to the search phase of MORDM by Trindade et al. (2017) indicate that the proposed application of MOEA search in MORDM is not always sufficient to discover many valid robust policy alternatives.

3.3.3 Multi-Scenario MORDM

The MORDM framework made significant strides toward developing a more effective process for handling policy problems characterized by deep uncertainty where there are multiple decision makers who are unable to agree on system details and solution objectives. The key to MORDM is the use of multi-objective evolutionary algorithms, which are responsible for determining potential robust policy alternatives. Based on the MORDM framework as established by Kasprzyk et al. (2013), the outcomes of a potential policy is compared with others based on a single base reference scenario. Policies that perform better than those belonging to the existing non-dominated set under the uncertainty settings established by the base scenario will be added to the non-dominated set, with conditions of deep uncertainty being included later in the analysis (Eker & Kwakkel, 2018). Watson and Kasprzyk (2017) recognize that this means the non-dominated set of options will be based on a single reference point and may not lead to policies that are robust if conditions change. This conclusion is similar to that of Trindade et al. (2017), discussed in Section 3.3.2. Remember that Trindade et al. (2017) addresses this by introducing randomness to the selection of the scenario used to compare success of policy options, which changes every generation. In contrast, the Multi-Scenario MORDM method proposed by Watson and Kasprzyk (2017) proposes to perform multiple iterations of the MOEA search process using distinct reference scenarios built based on the results of a sensitivity analysis in the traditional MORDM framework, with the goal of building a set of policy alternatives that cover a more diverse range of decision levers, in an effort to discover policy alternatives that perform well under even the most extreme conditions.

As Fig. 3.4 demonstrates, the key difference between MORDM and Multi-Scenario MORDM lies in the search for policy alternatives. The fact that multi-scenario MORDM involves repeating the MORDM process multiple times carries through the remainder of the framework's process, with both computational experimentation and scenario discovery being repeated for each reference scenario applied in the search phase.

Given the proposal to repeat the search for policy alternatives with multiple reference scenarios, the most influential decision in this method will be the selection of reference scenarios. The initial proposal, which is illustrated in Fig. 3.4 is to use the results of sensitivity analysis performed during the course of MORDM-based analysis as boundaries with which to select the reference scenarios from (Watson & Kasprzyk, 2017). Through this framework, Watson and Kasprzyk (2017) are able to conclude that the use of multiple reference scenarios lead to the discovery of a more diverse set of policy alternatives that perform well under deep uncertainty to consider.

The number of scenarios to select is left to the analyst to determine. By selecting a wide range of scenarios to use as reference points, it is more likely to discover a more diverse set of robust policy

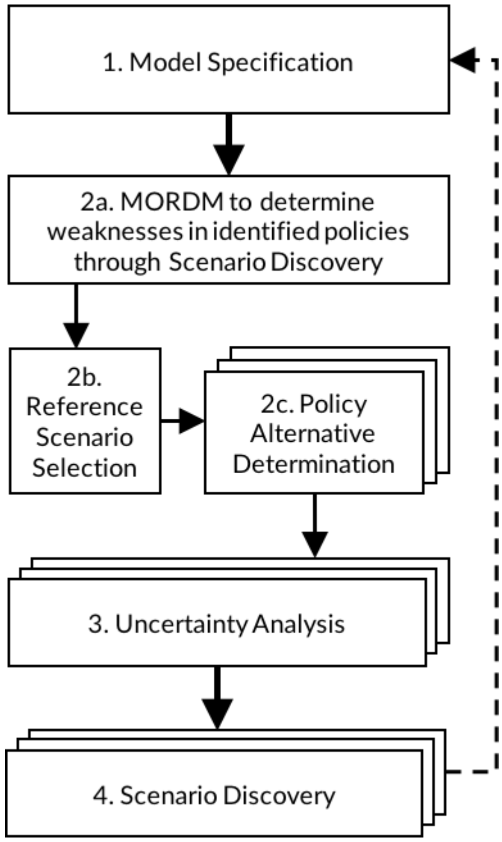

Figure 3.4: Multi-Scenario MORDM Process

options. However, analysis is often limited by computational capabilities and other factors, which means that analysts should search for only a small number of reference scenarios (Eker & Kwakkel, 2018). There are several criteria that can be considered during the selection process, including internal consistency, diversity of outcome indicators, extremeness, and policy relevance, which may be combined or considered separately depending on goals of the analysis or decision maker preference. Eker and Kwakkel (2018) use this information to develop an alternative method of reference scenario selection that is designed for the purpose of finding reference scenarios that, when used in the MOEA search, will lead to more robust solutions, focusing on policy relevance and maximum diversity (where policy relevance is defined as scenarios that lead to outcomes in the lower half of the scenario space and the diversity criterion is based on that which was defined by Carlsen, Lempert, Wikman-Svahn, and Schweizer (2016)). Eker and Kwakkel (2018) also apply a random approach by selecting a number of scenarios randomly across the uncertainty space, with no consideration for policy relevance or diversity.

After applying each mechanism to an intertemporal version of the lake problem (for more details about the lake problem, see Section 3.4), the results using policy relevant reference scenario selection mechanism were shown to lead to a more diverse set of policy alternatives, as well as larger variety of tradeoffs in robustness metrics (Eker & Kwakkel, 2018). The same analysis did not find a signficant difference between the results for the policy relevant scenario selection mechanism and the random mechanism, with both leading to similar increases policy alternative diversity and robust outcome tradeoffs (Eker & Kwakkel, 2018).

As Multi-Scenario MORDM is such an new framework, there are currently no other applications in literature. Much remains to be learned with respect to the impact of differing scenario selection mechanisms or the effect of using a different number of scenarios on the diversity and robustness of discovered policy alternatives.

3.3.4 Multi-Objective Robust Optimization (MORO)

Robust optimization has its roots in mathematical optimization of functions and systems. Traditional optimization algorithms seek the maximum (or minimum) solution(s) generated from a mathematical system and specified input space. In decision making, system optimization commonly refers to determining optimum values for levers which decision makers can use to create the conditions required for a system to reach the targeted outcome. As discussed in Section 3.2 these optimums can be unstable and sensitive to even small changes in parameter values. When the system under consideration includes uncertainty, let alone the deep uncertainty present in wicked problems with tipping point charactarisitcs, an optimum that is sensitive to changes in parameter values may easily lead to outcomes that do not hit the desired target. In these cases, a robust solution is desired over an optimum one, which must be accounted for in the optimization process. Early attempts to combine a desire for robustness with a search for optimality occurred in the 1960s (Dorato & Drenick, 1966) with the consideration of insensitive optimums, and 1970s with the development of an algorithm that seeks the optimum values of levers under worst-case performance settings, given a set of uncertain input parameters (Soyster, 1973). It was not until the 1990s, however, that what is now known as robust optimization began to take shape (Sözüer & Thiele, 2016).

A robust optimization algorithm seeks solutions to a problem that perform reasonably across all potential future states of the world and performs best in worst-case scenarios. Two primary forms of robust optimization have been developed. Deterministic robust optimization uses numerical techniques to determine function optimums explicitly. In situations with deep uncertainty, however, it is not possibly to explicitly define all robustness constraints and function variables. In this case, a randomized approach or simulation optimization approach is used (Beyer & Sendhoff, 2007). Given that wicked problems with tipping point characteristics will always include deep uncertainty, the consideration of robust optimization in this research will focus on simulation optimization.

Robust optimums in uncertain models are found through direct evaluation. Robust solutions to the objective function are determined by sampling across the uncertainty space, which can be accomplished in three primary ways (Beyer & Sendhoff, 2007):

1. Monte-Carlo strategies: this involves averaging the objective function values across a sample set of the uncertainty space. Sets of uncertainty values are determined through one of a variety of sampling techniques, including Latin hypercube, Monte-Carlo, and full- and partial-factorial. This method of simulation optimization quickly becomes computationally expensive.

2. Meta-model approach: generally used to reduce the computational cost of optimization, a meta-model is carefully constructed to represent the model of interest. Results of optimization using the meta-model are then generalized to estimate optimization results of the original model (Zhou, Zhou, Liu, & Zhou, 2017)

3. Optimization using objective function directly: instead of calculating robustness values from the objective function, this method proposes to use the values of the objective function directly to compare policy alternatives (Beyer & Sendhoff, 2007)

Evolutionary algorithms have been used in combination with Monte-Carlo strategies and noisy optimization to more efficiently obtain robust solutions to a model's objective function. Evolutionary algorithms are discussed in more detail in **??**.

Previous literature has applied robust optimization techniques as a small step in a larger framework of designing policy solutions for a problem characterized by deep uncertainty. Hamarat, Kwakkel, and Pruyt (2013) uses multi-objective robust optimization as a method of fine-tuning tipping points during adaptive policy making. Kwakkel et al. (2015) makes use of MORO in the early stages of the development of Dynamic Adaptive Policy Pathways (DAPP) by analyzing the robustness of a set of pre-specified policy pathways across a sample of the defined uncertainty space for a model. Extensions of the MORDM framework have approached multi-objective robust optimization in the search phase by including multiple reference scenarios in the search process (Trindade et al., 2017; Watson & Kasprzyk, 2017).

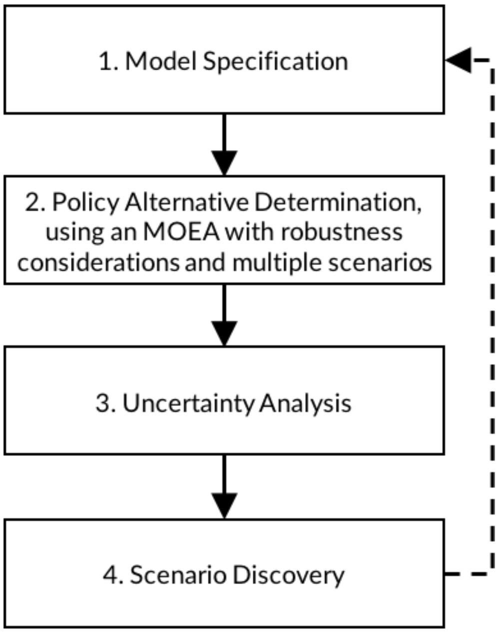

Figure 3.5: Multi-Objective Robust Optimization Process

This research proposes multi-objective robust optimization framework that follows the structure of RDM and can be used as the primary framework for analysis to determine the set of robust policy options to a deeply uncertain wicked problem with multiple conflicting objectives. Model specification follows the same XLRM format as was described in Section 3.3.1. Robust optimization takes place in the MOEA-based policy alternative determination phase of the framework, as illustrated in Fig. 3.5. In this stage, policies are compared based on the outcomes of defined robustness metrics, leading to the Pareto optimized set of policies based on robustness. The multi-objective robust search is most suited for the domain criterion robustness metric, which determines robustness as the volume of scenarios in which each outcome meets a specified threshold. To support this calculation, each policy tested in the MOEA search is evaluated against a set of pre-specified scenarios. The outcomes calculated are then used to determine robustness for each outcome. An alternative implementation of the policy alternative determination process is described by Beh, Zheng, Dandy, Maier, and Kapelan (2017), in which a meta model is used to evaluate robustness, which is determined using a single robustness function. However, by using the model under analysis directly, and by considering each outcome of interest as an independent element of robustness, the search phase is better able to both fully capture the dynamics of the model and consider conflicting objectives, making the robust optimization method proposed in this research a more direct method of incorporating robustness into the search phase of analysis. After the search phase is complete, the result is a set of Pareto non-dominated policy solutions that are already considered robust against the set of evaluation scenarios specified for the search. These policies are further tested in uncertainty analysis and scenario discovery to both determine robustness against a significantly larger set of scenarios and find remaining vulnerabilities that can be used to refine the model specification in a new iteration of analysis.

3.3.5 The Fundamental Difference between MORDM, Multi-Scenario MORDM, and MORO

Each of the methods described, MORDM, Multi-Scenario MORDM, and MORO, are based on the foundations of robust decision support: the RDM framework. Though there are many similarities between the three methods, there is one fundamental difference. As highlighted in Fig. 3.6, each method takes a different approach to the policy alternative determination phase of the framework.

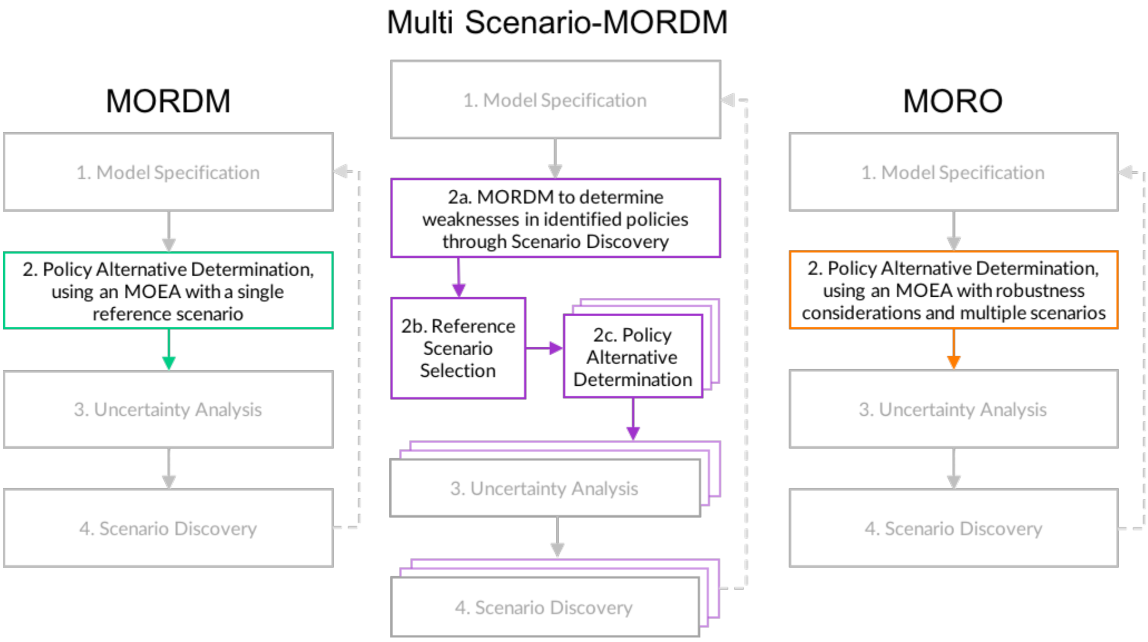

Figure 3.6: Comparison of robust decision support method structures

From left to right, the frameworks increasingly consider robustness in the search phase. MORDM selects policies based on outcomes of interest alone. While multi-scenario MORDM also selects policies based on outcome values, it performs the search using multiple reference scenarios identified from within the vulnerable uncertainty space as determined in scenario discovery of MORDM, thereby introducing small robustness considerations into the search phase. And finally, instead of using outcome values, MORO selects policies by testing each against a pre-defined set of scenarios and using the robustness metric for each outcome, which even more directly brings robustness considerations into the search phase of the framework.

As the model specification, uncertainty analysis, and scenario discovery processes are all constant across the three methods, this research will be examining the impact of considering robustness in the policy alternative determination step of an RDM-based framework for decision support.

3.4 Problem and Policy Configuration Details

3.4.1 The Lake Problem

- History of the lake problem including origins.

- Why the lake problem is a good representation of a wicked problem with tipping point characteristics.

- Benefit of using benchmarking problems in comparison analysis

2-3 pages

3.4.2 Policy Implementation Structure

Policy structure types:

- Static. A fixed strategy that is established at the beginning of the time period and is not adjusted despite changes in the system being affected. This is seen as impractical in the context of a wicked problem where behavior over time is so largely unknown.

- Adaptive. A flexible strategy that responds to changing conditions. Adaptive policies can be implemented in many different ways:

  Action that is preset to be made every n timesteps. Action can vary across timesteps but is preset. Considered in this analysis as the intertemporal model. (Can also consider intertemporal with multiple timesteps - 1 and 10)

  Decisions are set by a rule (function of multiple variables) that is examined every n policy timesteps. Considered in this analysis through both the direct policy search on 1 and n timesteps. Define what is meant by direct policy search.

  Adaptive, based on trigger points which are developed in regards to outcome values. Similar to policies defined by rules, though instead of the rules being a function of multiple variables, the rules are a function of outcome variables. The adaptive policy can take on many forms, and as it is similar in behavior to the direct policy search structure, it will not be considered.

1-2 pages

II

Design and Development

4

Methods of Comparisons

Several dimensions, both quantitative and qualitative,will be used to compare the results of each method and problem iteration pairing.

4.1 Setup Complexity

The first set of measures relates to the complexity of setup and use of each method. These are both qualitative measures that are related to each method and are independent of the problem iteration under consideration.

************ I am aware these are currently not specified in enough detail. My intention is to use number of instances of method and problem use in literature and existence of software packages that execute each method automatically. ************

Comparison 1 : Complexity of problem setup
Based on earlier work comparing methods of decision making, this metric focuses on the effort required to prepare the problem under consideration for analysis (Kwakkel, Haasnoot, & Walker, 2016; Roach et al., 2015). This metric considers the following elements: if policies, must be specified before execution; how the model must format uncertainties, decision levers, and outcomes of interest.

Comparison 2 : Complexity of method setup
This metric considers any method-specific setup that is required for the robust decision support method to be executed. Similar to metrics used in comparative research completed previously, this includes the availability and number of tools and software packages required to execute each method (Gersonius, Ashley, Jeuken, Pathinara, & Zevenbergen, 2015; Kwakkel, Haasnoot, & Walker, 2016). Under this metric, the amount of work (through additional code and analysis beyond what is provided directly by any available tools), will also be considered.

4.2 Communication

Both quantitative and qualitative, these metrics describe the strength of method and results communication.

Comparison 3 : Results communication This metric refers to whether methods communicate results throughout execution or only upon completion. If results are reported progressively, the similarity of the initial recommended policy set communicated to the final set recommended by the method under consideration will be considered. This will provide an indication of the value of the earliest results communicated. See Section 4.4 for details on the method of determining similarity.

Comparison 4 : Robustness communication
This metric, specific to the methods of robustness used in each method's execution during this analysis, indicates whether the robustness metric is an abstract or direct representation of policy measures. Similar measures have been used in the past to indicate the ability of a method to effectively recommend robust policy solutions (Gersonius et al., 2015; Roach et al., 2015).

Comparison 5 : Ease of results updating if uncertainty space changes
Wicked problems with tipping point logic are characterized by deep uncertainty, so the models that describe these problems can never be exact. These models are subject to frequent changes based on new data or additional stakeholder input, whether it be to uncertainty ranges, model structure, or available policy levers. Therefore, it is important to consider how decision support methods respond to these changes (Gersonius et al., 2015). This metric examines how each method is able to respond to changes in the model under analysis.

Comparison 6 : Ease of results updating if desired robustness measure changes
As discussed in Section 3.2, the selected robustness measure has a significant impact on the recommendations made by each decision support method. A change in robustness measure may occur due to changing stakeholder interests, or a desire to consider multiple robustness perspectives for the same problem. This metric supports the previous by indicating the ability for each method to quickly and easily respond to a change in robustness measure.

4.3 Results

The final set of metrics are quantitative in nature and focus on comparing the resulting recommended policy sets for each method and problem iteration pairing.

Comparison 7 : Computational cost
Computation cost will be measured in two related ways: time of execution and number of model executions required. This is also known as the number of function executions.

Comparison 8 : Convergence of results
This metric tracks the effect of altering the number of functional executions or number of scenarios tested on the convergence to a policy set with similar robustness values to what was originally determined. **??** describes method implementations with large numbers of scenarios and function executions to ensure that maximally robust policies are found. By examining the effects of changing these method parameters on convergence toward the original set of recommended robust policies, the minimum number of function executions and scenarios can be approximated. This, in combination with computational cost of the original values (determined in *Comparison 7: Computational cost*), can provide an estimated minimum computational cost for each method.

Comparison 9 : Robustness of recommended policy sets
Given that each of the methods use similar robustness measures to evaluate policy options (see **??**), this metric examines the differences in robustness calculated for each set of recommended solutions.

Describe why comparing robustness is good

Comparison 10 : Similarity of recommended policy sets
Once final policy sets are determined for each method, this metric compares the

similarity of policy options in each set. This provides a mechanism for determining whether each method reaches similar insights independent of the robustness calculation for each policy in the final set (Hall et al., 2012). Section 4.4 describes the mechanism used to determine similarity for this and other comparison methods described.

Comparison 11 : Similarity of robustness
This metric compares the robustness value of policies that are considered similar (based on the similarity mechanism defined in Section 4.4). This communication, paired with the analysis of overall robustness of recommended policy sets (see *Comparison 9*), provides an guideline for how selection of a specific method may affect final robustness values. For example, this may indicate that one method leads to lower robustness values, despite similar policies being recommended. Given that the robustness values, like the policy values, are a vector of multiple distinct values, a mechanism must be defined to determine similarity. A description of this mechanism can be found in Section 4.4.

4.4 Determining Policy and Robustness Similarity

Each recommended policy and the robustness determination for each is represented as a vector of values. Therefore, similarity must be determined through a clearly defined measure. The selection of similarity measure will have a large impact on the comparisons made in this study, as it will be used in several of the comparison metrics defined in Sections 4.1 to 4.3.

Policy and robustness data are both quantitative and dimensionless vectors. Based on two studies that compares several similarity and dissimilarity metrics, both qualitatively and quantitatively, the similarity measure identified as best for this research is Euclidean distance Buttigieg and Ramette, 2014; Shirkhorshidi, Aghabozorgi, Wah, and Dalby, 2015. Euclidean distance is one of the most well-known methods for determining the distance of numerical data sets (Shirkhorshidi et al., 2015) and is best used when considering quantitative value sets of homogeneous data (Buttigieg & Ramette, 2014), which is what is found in both the policy and robustness vectors.

Euclidean distance has two recognized disadvantages. First, that it may indicate two vectors of distinct values are closer together than another pair of vectors that share one or more common values (Shirkhorshidi et al., 2015). Similarity comparisons in this research involve continuous data. Because of this, vectors sharing common values are not necessarily more similar if other values have larger differences. For example, the vectors <0.05, 0.20, 0.34> and <0.05, 0.015, 0.34>, with two common values, should not be considered as more similar than vectors <0.05, 0.20, 0.34> and <0.04, 0.21, 0.33>, with no common values. Therefore, this first disadvantage will not negatively affect the similarity calculations in this research.

Second, Euclidean distance can be dominated by the variable with the largest potential value and so values used should be normalized when vectors include values of different scales (Shirkhorshidi et al., 2015). Based on the defined mechanism for determining robustness (see **??**), the robustness vector contains values of the same scale and do not need to be normalized. However, the policy vector is defined with variables that have varying ranges of possible values. Therefore, policy similarity will be determined using normalized policy vectors.

Euclidean distance will be calculated practically with the Python-based SciPy library and using the method `scipy.spatial.distance.euclidean`. This method will accept two normalized policy vectors or two unmodified robustness vectors, and will return the Euclidean distance. In Eq. (4.1), the two normalized policy vectors are represented by X and Y.

$$distance(X,Y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^n} \tag{4.1}$$

Each policy and robustness vector will be compared with every other vector, resulting in a two-dimensional matrix of Euclidean distance values. In the Euclidean distance function, `distance(X,Y) == distance(Y,X)`, the resulting matrix will be bisymmetric, where the upper triangle will be equal to the lower triangle, separated by a diagonal of zeros. Because of this, an upper triangular matrix will be constructed where the lower triangular values are set to zero, which will be further processed in the targeted comparison method.

III

Analysis and Conclusions

References

Bankes, S. C. (1993). Exploratory modeling for policy analysis. *Operations Research*, *41*(3), 435–449. https://doi.org/10.1287/opre.41.3.435

Bankes, S. C. (2002). Tools and techniques for developing policies for complex and uncertain systems. *Proceedings of the National Academy of Sciences*, *99*(suppl 3), 7263–7266. https://doi.org/10.1073/pnas.092081399

Beh, E. H., Zheng, F., Dandy, G. C., Maier, H. R., & Kapelan, Z. (2017). Robust optimization of water infrastructure planning under deep uncertainty using metamodels. *Environmental Modelling & Software*, *93*, 92–105. https://doi.org/10.1016/j.envsoft.2017.03.013

Ben-Haim, Y. (2006). Chapter 1 - overview. In Y. Ben-Haim (Ed.), *Info-gap decision theory (second edition)* (2nd ed., Chap. 1, pp. 1–8). Oxford: Academic Press. https://doi.org/10.1016/B978-012373552-2/50002-3

Bertsimas, D. & Sim, M. (2004). The price of robustness. *Operations Research*, *52*(1), 35–53. https://doi.org/10.1287/opre.1030.0065

Beyer, H.-G. & Sendhoff, B. (2007). Robust optimization - a comprehensive survey. *Computer Methods in Applied Mechanics and Engineering*, *196*(33), 3190–3218. https://doi.org/10.1016/j.cma.2007.03.003

Branke, J. (1998). Creating robust solutions by means of evolutionary algorithms. In A. E. Eiben, T. Bäck, M. Schoenauer, & H.-P. Schwefel (Eds.), *Parallel problem solving from nature — ppsn v* (pp. 119–128). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/bfb0056855

Brown, C., Ghile, Y., Laverty, M., & Li, K. (2012). Decision scaling: Linking bottom-up vulnerability analysis with climate projections in the water sector. *Water Resources Research*, *48*(9), 1–12. https://doi.org/10.1029/2011WR011212

Bryant, B. P. & Lempert, R. J. (2010). Thinking inside the box: A participatory, computer-assisted approach to scenario discovery. *Technological Forecasting and Social Change*, *77*(1), 34–49. https://doi.org/10.1016/j.techfore.2009.08.002

Buttigieg, P. L. & Ramette, A. (2014). A guide to statistical analysis in microbial ecology: A community-focused, living review of multivariate data analyses. *FEMS Microbiology Ecology*, *90*(3), 543–550. https://doi.org/10.1111/1574-6941.12437

Capano, G. & Woo, J. J. (2017). Resilience and robustness in policy design: A critical appraisal. *Policy Sciences*, *50*(3), 399–426. https://doi.org/10.1007/s11077-016-9273-x

Carlsen, H., Lempert, R., Wikman-Svahn, P., & Schweizer, V. (2016). Choosing small sets of policy-relevant scenarios by combining vulnerability and diversity approaches. *Environmental Modelling & Software*, *84*, 155–164. https://doi.org/10.1016/j.envsoft.2016.06.011

Coello Coello, C. A. (2006). Evolutionary multi-objective optimization: A historical view of the field. *IEEE Computational Intelligence Magazine*, *1*(1), 28–36. https://doi.org/10.1109/mci.2006.1597059

de Goede, D., Gremmen, B., & Blom-Zandstra, M. (2013). Robust agriculture: Balancing between vulnerability and stability. *NJAS - Wageningen Journal of Life Sciences*, *64–65*, 1–7. https://doi.org/10.1016/j.njas.2012.03.001

Deb, K. & Gupta, H. (2006). Introducing robustness in multi-objective optimization. *Evolutionary Computation*, *14*(4), 463–494. https://doi.org/10.1162/evco.2006.14.4.463

Dittrich, R., Wreford, A., & Moran, D. (2016). A survey of decision-making approaches for climate change adaptation: Are robust methods the way forward? *Ecological Economics*, *122*, 79–89. https://doi.org/10.1016/j.ecolecon.2015.12.006

Dorato, P. & Drenick, R. (1966). Optimality, insensitivity, and game theory. In L. Radanovic (Ed.), *Sensitivity methods in control theory* (pp. 78–102). Pergamon. https://doi.org/10.1016/B978-1-4831-9822-4.50009-1

Durach, C. F., Wieland, A., & Machuca, J. A. (2015). Antecedents and dimensions of supply chain robustness: A systematic literature review. *International Journal of Physical Distribution & Logistics Management*, *45*(1/2), 118–137. https://doi.org/10.1108/IJPDLM-05-2013-0133

Edwards, D. J. A. (1998). Types of case study work: A conceptual framework for case-based research. *Journal of Humanistic Psychology*, *38*(3), 36–70. https://doi.org/10.1177/00221678980383003

Eker, S. & Kwakkel, J. H. (2018). Including robustness considerations in the search phase of many-objective robust decision making. *Environmental Modelling & Software*, *105*, 201–206. https://doi.org/10.1016/j.envsoft.2018.03.029

General Assembly Resolution 70/1. (2015). Transforming our world: The 2030 agenda for sustainable development. 21 October 2015. United Nations. Retrieved from http://undocs.org/A/RES/70/1

Gersonius, B., Ashley, R., Jeuken, A., Pathinara, A., & Zevenbergen, C. (2015). Accounting for uncertainty and flexibility in flood risk management: Comparing real�in�options optimisation and adaptation tipping points. *Journal of Flood Risk Management*, *8*(2), 135–144. https://doi.org/10.1111/jfr3.12083

Giuliani, M. & Castelletti, A. (2016). Is robustness really robust? how different definitions of robustness impact decision-making under climate change. *Climatic Change*, *135*(3), 409–424. https://doi.org/10.1007/s10584-015-1586-9

Gladwell, M. (2006). *The tipping point: How little things can make a big difference*. Little, Brown.

Hadka, D., Herman, J., Reed, P., & Keller, K. (2015). An open source framework for many-objective robust decision making. *Environmental Modelling & Software*, *74*, 114–129. https://doi.org/10.1016/j.envsoft.2015.07.014

Hall, J. W., Lempert, R. J., Keller, K., Hackbarth, A., Mijere, C., & McInerney, D. J. (2012). Robust climate policies under uncertainty: A comparison of robust decision making and info�gap methods. *Risk Analysis*, *32*(10), 1657–1672. https://doi.org/10.1111/j.1539-6924.2012.01802.x

Hamarat, C., Kwakkel, J. H., & Pruyt, E. (2013). Adaptive robust design under deep uncertainty. *Technological Forecasting and Social Change*, *80*(3), 408–418. https://doi.org/10.1016/j.techfore.2012.10.004

Hamarat, C., Kwakkel, J. H., Pruyt, E., & Loonen, E. T. (2014). An exploratory approach for adaptive policymaking by using multi-objective robust optimization. *Simulation Modelling Practice and Theory*, *46*, 25–39. Simulation-Optimization of Complex Systems: Methods and Applications. https://doi.org/10.1016/j.simpat.2014.02.008

Herman, J. D., Reed, P. M., Zeff, H. B., & Characklis, G. W. (2015). How should robustness be defined for water systems planning under change? *Journal of Water Resources Planning and Management*, *141*(10), 04015012-1–04015012-14. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000509

Herman, J. D., Zeff, H. B., Reed, P. M., & Characklis, G. W. (2014). Beyond optimality: Multistakeholder robustness tradeoffs for regional water portfolio planning under deep uncertainty. *Water Resources Research*, *50*(10), 7692–7713. https://doi.org/10.1002/2014WR015338

Kasprzyk, J. R., Nataraj, S., Reed, P. M., & Lempert, R. J. (2013). Many objective robust decision making for complex environmental systems undergoing change. *Environmental Modelling & Software*, *42*, 55–71. https://doi.org/10.1016/j.envsoft.2012.12.007

Knight, F. H. (1921). *Risk, uncertainty and profit*. Library of Economics and Liberty. Retrieved from http://www.econlib.org/library/Knight/knRUP.html

Kwakkel, J. H. (2017). The exploratory modeling workbench: An open source toolkit for exploratory modeling, scenario discovery, and (multi-objective) robust decision making. *Environmental Modelling & Software*, *96*, 239–250. https://doi.org/10.1016/j.envsoft.2017.06.054

Kwakkel, J. H., Eker, S., & Pruyt, E. (2016). How robust is a robust policy? comparing alternative robustness metrics for robust decision-making. In M. Doumpos, C. Zopounidis, & E. Grigoroudis (Eds.), *Robustness analysis in decision aiding, optimization, and analytics* (1st ed., Chap. 10, Vol. 241, pp. 221–237). International Series in Operations Research & Management Science. Springer International Publishing. https://doi.org/10.1007/978-3-319-33121-8_10

Kwakkel, J. H., Haasnoot, M., & Walker, W. E. (2015). Developing dynamic adaptive policy pathways: A computer-assisted approach for developing adaptive strategies for a deeply uncertain world. *Climatic Change*, *132*(3), 373–386. https://doi.org/10.1007/s10584-014-1210-4

Kwakkel, J. H., Haasnoot, M., & Walker, W. E. (2016). Comparing robust decision-making and dynamic adaptive policy pathways for model-based decision support under deep uncertainty. *Environmental Modelling & Software*, *86*, 168–183. https://doi.org/10.1016/j.envsoft.2016.09.017

Lempert, R. J. (2002). A new decision sciences for complex systems. *Proceedings of the National Academy of Sciences*, *99*(suppl 3), 7309–7313. https://doi.org/10.1073/pnas.082081699

Lempert, R. J. & Collins, M. T. (2007). Managing the risk of uncertain threshold responses: Comparison of robust, optimum, and precautionary approaches. *Risk Analysis*, *27*(4), 1009–1026. https://doi.org/10.1111/j.1539-6924.2007.00940.x

Lempert, R. J., Groves, D. G., Popper, S. W., & Bankes, S. C. (2006). A general, analytic method for generating robust strategies and narrative scenarios. *Management Science*, *52*(4), 514–528. https://doi.org/10.1287/mnsc.1050.0472

Lempert, R. J., Popper, S. W., & Bankes, S. C. (2003). *Shaping the next one hundred years: New methods for quantitative, long-term policy analysis* (1st ed.). RAND Corporation. Retrieved from http://www.jstor.org/stable/10.7249/mr1626rpc

Lenton, T. (2013). Tipping elements from a global perspective. In *Addressing tipping points for a precarious future*. Oxford: British Academy. https://doi.org/10.5871/bacad/9780197265536.003.0002

Maier, H., Guillaume, J., van Delden, H., Riddell, G., Haasnoot, M., & Kwakkel, J. (2016). An uncertain future, deep uncertainty, scenarios, robustness and adaptation: How do they fit together? *Environmental Modelling & Software*, *81*, 154–164. https://doi.org/10.1016/j.envsoft.2016.03.014

Marler, R. T. & Arora, J. S. (2004). Survey of multi-objective optimization methods for engineering. *Structural and Multidisciplinary Optimization*, *26*(6), 369–395. https://doi.org/10.1007/s00158-003-0368-6

Matrosov, E. S., Padula, S., & Harou, J. J. (2013). Selecting portfolios of water supply and demand management strategies under uncertainty—contrasting economic optimisation and 'robust decision making' approaches. *Water Resources Management*, *27*(4), 1123–1148. https://doi.org/10.1007/s11269-012-0118-x

Matrosov, E. S., Woods, A. M., & Harou, J. J. (2013). Robust decision making and info-gap decision theory for water resource system planning. *Journal of Hydrology*, *494*, 43–58. https://doi.org/10.1016/j.jhydrol.2013.03.006

McInerney, D., Lempert, R., & Keller, K. (2012). What are robust strategies in the face of uncertain climate threshold responses? *Climatic Change*, *112*(3), 547–568. https://doi.org/10.1007/s10584-011-0377-1

McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, *21*(2), 239–245. https://doi.org/10.2307/1268522

McPhail, C., Maier, H. R., Kwakkel, J. H., Giuliani, M., Castelletti, A., & Westra, S. (2018). Robustness metrics: How are they calculated, when should they be used and why do they give different results? *Earth's Future*, *6*(2), 169–191. https://doi.org/10.1002/2017EF000649

Nair, P. K. R. & Nair, V. D. (2014). Organization of a research paper: The imrad format. In *Scientific writing and communication in agriculture and natural resources* (Chap. 2, pp. 13–25). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-03101-9_2

Parmee, I. C. & Bonham, C. R. (2002). Improving cluster oriented genetic algorithms for high-performance region identification. In I. C. Parmee & P. Hajela (Eds.), *Optimization in industry* (pp. 189–202). London: Springer London. https://doi.org/10.1007/978-1-4471-0675-3_16

Pennings, P., Keman, H., & Kleinnijenhuis, J. (2006). Meaning and use of the comparative method: Research design. In *Doing research in political science* (2nd ed., Chap. 4, pp. 30–52). London: SAGE Publications Ltd. https://doi.org/10.4135/9781849209038.n4

Popper, S. W., Lempert, R. J., & Bankes, S. C. (2005). Shaping the future. *Scientific American*, *292*(4), 66–71. https://doi.org/10.1038/scientificamerican0405-66

Quade, E. S. & Carter, G. M. (1989). *Analysis for public decisions*. MIT Press.

Rittel, H. W. J. & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, *4*(2), 155–169. https://doi.org/10.1007/bf01405730

Roach, T., Kapelan, Z., & Ledbetter, R. (2015). Comparison of info-gap and robust optimisation methods for integrated water resource management under severe uncertainty. *Procedia Engineering*, *119*, 874–883. https://doi.org/10.1016/j.proeng.2015.08.955

Roach, T., Kapelan, Z., Ledbetter, R., & Ledbetter, M. (2016). Comparison of robust optimization and info-gap methods for water resource management under deep uncertainty. *Journal of Water Resources Planning and Management*, *142*(9), 04016028-1–04016028-13. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000660

Rosenhead, J., Elton, M., & Gupta, S. K. (1972). Robustness and optimality as criteria for strategic decisions. *Journal of the Operational Research Society*, *23*(4), 413–431. https://doi.org/10.1057/jors.1972.72

Seawright, J. & Gerring, J. (2008). Case selection techniques in case study research: A menu of qualitative and quantitative options. *Political Research Quarterly*, *61*(2), 294–308. https://doi.org/10.1177/1065912907313077

Shirkhorshidi, A. S., Aghabozorgi, S. R., Wah, T. Y., & Dalby, A. R. (2015). A comparison study on similarity and dissimilarity measures in clustering continuous data. *PLoS One*, *10*(12), 20. https://doi.org/10.1371/journal.pone.0144059

Sniedovich, M. (2016). From statistical decision theory to robust optimization: A maximin perspective on robust decision-making. In M. Doumpos, C. Zopounidis, & E. Grigoroudis (Eds.), *Robustness analysis in decision aiding, optimization, and analytics* (Chap. 4, Vol. 241, pp. 59–87). International Series in Operations Research & Management Science. Cham, Swizerland: Springer International Publishing. https://doi.org/10.1007/978-3-319-33121-8_4

Soyster, A. L. (1973). Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research*, *21*(5), 1154–1157. https://doi.org/10.1287/opre.21.5.1154

Sözüer, S. & Thiele, A. C. (2016). The state of robust optimization. In M. Doumpos, C. Zopounidis, & E. Grigoroudis (Eds.), *Robustness analysis in decision aiding, optimization, and analytics* (Chap. 5, Vol. 241, pp. 89–112). International Series in Operations Research & Management Science. Cham, Switzerland: Springer International Publishing. https://doi.org/10.1007/978-3-319-33121-8_5

Taleb, N. N. (2007). *The black swan: The impact of the highly improbable* (1st ed.). Random House Group.

Tannert, C., Elvers, H.-D., & Jandrig, B. (2007). The ethics of uncertainty. *EMBO reports*, *8*(10), 892–896. https://doi.org/10.1038/sj.embor.7401072

Trindade, B., Reed, P., Herman, J., Zeff, H., & Characklis, G. (2017). Reducing regional drought vulnerabilities and multi-city robustness conflicts using many-objective optimization under deep uncertainty. *Advances in Water Resources*, *104*, 195–209. https://doi.org/10.1016/j.advwatres.2017.03.023

Tsoukiàs, A. (2008). From decision theory to decision aiding methodology. *European Journal of Operational Research*, *187*(1), 138–161. https://doi.org/10.1016/j.ejor.2007.02.039

Tsutsui, S. & Ghosh, A. (1997). Genetic algorithms with a robust solution searching scheme. *IEEE Transactions on Evolutionary Computation*, *1*(3), 201–208. https://doi.org/10.1109/4235.661550

Twomey, P. (2012). Resilience lessons for australian climate change mitigation policy. In *35th annual IAEE international conference*. Perth, Western Australia. Retrieved March 26, 2018, from http://www.ceem.unsw.edu.au/staff/paul-twomey

uncertain. (2013), In *Cambridge advanced learner's dictionary & thesaurus* (4th). Cambridge University Press. Retrieved April 2, 2014, from https://dictionary.cambridge.org/dictionary/english/uncertain

United Nations. (2015). *The millennium development goals report*. United Nations. New York. Retrieved from http://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG%202015%20rev%20(July%201).pdf

van Bueren, E. M., Klijn, E., & Koppenjan, J. F. M. (2003). Dealing with wicked problems in networks: Analyzing an environmental debate from a network perspective. *Journal of Public Administration Research and Theory*, *13*(2), 193–212. https://doi.org/10.1093/jpart/mug017

Walker, W. E., Haasnoot, M., & Kwakkel, J. H. (2013). Adapt or perish: A review of planning approaches for adaptation under deep uncertainty. *Sustainability*, *5*(3), 955–979. https://doi.org/10.3390/su5030955

Walker, W. E., Lempert, R. J., & Kwakkel, J. H. (2013). Deep uncertainty. In S. I. Gass & M. C. Fu (Eds.), *Encyclopedia of operations research and management science* (2013th ed., pp. 395–402). Springer US. https://doi.org/10.1007/978-1-4419-1153-7_1140

Watson, A. A. & Kasprzyk, J. R. (2017). Incorporating deeply uncertain factors into the many objective search process. *Environmental Modelling & Software*, *89*, 159–171. https://doi.org/10.1016/j.envsoft.2016.12.001

Weaver, C. P., Lempert, R. J., Brown, C., Hall, J. A., Revell, D., & Sarewitz, D. (2013). Improving the contribution of climate model information to decision making: The value and demands of robust decision frameworks. *Wiley Interdisciplinary Reviews: Climate Change*, *4*(1), 39–60. https://doi.org/10.1002/wcc.202

Yin, R. K. (2012). A (very) brief refresher on the case study method. In *Applications of case study research* (3rd ed., Chap. 1, pp. 3–20). Thousand Oaks, CA: SAGE Publications. Retrieved from https://www.sagepub.com/sites/default/files/upm-binaries/41407_1.pdf

Zainal, Z. (2007). Case study as a research method. *Jurnal Kemanusiaan*, *5*(1), 1–6. Retrieved from https://jurnalkemanusiaan.utm.my/index.php/kemanusiaan/article/view/165

Zhou, H., Zhou, Q., Liu, C., & Zhou, T. (2017). A kriging metamodel-assisted robust optimization method based on a reverse model. *Engineering Optimization*, *50*(2), 253–272. https://doi.org/10.1080/0305215X.2017.1307355