Delft University of Technology

Including robustness considerations in the search phase of Many-Objective Robust Decision Making

Eker, Sibel; Kwakkel, Jan H.

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Including robustness considerations in the search phase of Many-Objective Robust Decision Making

Highlights

- MORDM aims at developing robust solutions but identifies them only under a reference scenario.
- We propose an extension to the search phase of MORDM.
- We generate the candidate solutions under multiple scenarios.
- We select these scenarios based on diversity and policy relevance.
- We obtain a wider variety of robustness tradeoffs compared to the reference case.

Abstract

Many-Objective Robust Decision Making (MORDM) is a prominent model-based approach for dealing with deep uncertainty. MORDM has four phases: a systems analytical *problem formulation*, a *search* phase to generate candidate solutions, a *trade-off analysis* where different strategies are compared across many objectives, and a *scenario discovery* phase to identify the vulnerabilities. In its original inception, the search phase identifies optimal strategies for a single reference scenario for deep uncertainties, which may result in missing locally near-optimal, but globally more robust strategies. Recent work has addressed this issue by generating candidate strategies for multiple policy-relevant scenarios. In this paper, we incorporate a systematic scenario selection procedure in the search phase to consider both policy relevance and scenario diversity. The results demonstrate an increased tradeoff variety besides higher robustness, compared to the solutions found for a reference scenario. Future research can routinize multi-scenario search in MORDM with the aid of software packages.

1. Introduction

Environmental decision problems are often characterized by multiple conflicting objectives and deep uncertainties. Deep uncertainty refers to situations where decision makers and stakeholders do not know or cannot agree on a system delineation and description, the probability distributions of the uncertainties, and the valuation of alternative outcomes (Lempert et al., 2003; Walker et al., 2013). In other words, a variety of possibilities (e.g. futures, model formulations) can be enumerated for deep uncertainties, but they cannot be rank ordered or associated with probability distributions (Kwakkel et al., 2010). Robust Decision Making (Lempert et al., 2006) is an approach to decision making under deep uncertainty, where decision options are evaluated over a large set of scenarios. These scenarios are generated using analytic models for exploring future possibilities regardless of their likelihood, rather than relying on a best-estimate future. RDM emphasizes the evaluation of decision options based on the trade-offs among them, and introduces an iterative stress-testing framework using scenario discovery (Bryant and Lempert, 2010; Kwakkel and Jaxa-Rozen, 2016) to enhance the robustness of decision options.

As an extension to RDM, Many-Objective Robust Decision Making (MORDM) is a prominent model-based framework developed to deal with multi-objective decision problems under deep uncertainty (Herman et al., 2014; Kasprzyk et al., 2013). MORDM extends RDM in terms of the generation of alternative strategies. While the latter relies on a set of discrete, pre-specified alternatives which are

1

iteratively refined, the former generates a large set of alternative decisions with computational search (Herman et al., 2015). Multi-objective evolutionary algorithms employed in the search phase of MORDM align with the emphasis of RDM on tradeoffs between the options. MORDM has 4 iterative steps that incorporates various decision analytic methods:

(i) *Problem formulation:* Formulating the problem based on a systems analytical problem definition framework such as a system diagram (Thissen, 2013; Van der Lei et al., 2011; Walker, 2000), or the XLRM framework (Lempert et al., 2006). In the XLRM problem formulation framework, the components of a decision problem are grouped as uncertainties (X), decision levers (L), relationships (R) that represent the system structure, e.g. a simulation model, and performance metrics (M) that relate to the policy goals and measure the success of a solution, e.g. outcome indicators or objectives in an optimization problem.

(ii) *Generating candidate solutions:* Searching for candidate solutions that optimize multiple objectives by using multi-objective evolutionary algorithms (Coello et al., 2007; Reed et al., 2013).

(iii) *Tradeoff analysis:* Generating an ensemble of scenarios to explore the effects of deep uncertainties (Bankes, 1993; Bankes et al., 2013) and evaluating the robustness tradeoffs among multiple objectives for each candidate solution across these scenarios. A 'scenario' in this context refers to a computational experiment with an analytical model, where an alternative representation of each uncertain model parameter or structure is taken into account in each experiment. 'States of the world' or 'cases' are alternative terms used in the literature to refer to these computational experiments.

(iv) *Scenario discovery:* Improving the candidate solutions by using scenario discovery to detect the vulnerabilities, i.e. the combinations of uncertainties which cause the candidate solutions to perform poorly (Bryant and Lempert, 2010; Kwakkel and Jaxa-Rozen, 2016; Lempert, 2013).

In the original MORDM framework, candidate strategies are determined for a reference scenario, and the effect of deep uncertainties are taken into account later. In other words, deeply uncertain model elements are set to their reference values, and multi-objective optimization is run given this model parameterization in order to determine the candidate solutions. Hence, candidate solutions are identified for a baseline scenario without considering deep uncertainties. However, these candidate solutions are expected to be robust against uncertainties across a variety of scenarios. Therefore, searching for them only under a baseline scenario confines the potential to identify highly robust solutions.

To enhance the robustness of candidate solutions, it is promising to incorporate uncertainties in the search phase of MORDM and generate candidate solutions for a variety of scenarios. A recent study by Watson and Kasprzyk (2017) has drawn attention to this absence of uncertainty consideration in the search phase. Watson and Kasprzyk (2017) searched for candidate solutions in multiple scenarios, and demonstrated how the performance of candidate solutions depends on the scenario for which they were generated. Therefore, extending the search phase of MORDM with optimization for multiple scenarios has the potential to identify more robust solution alternatives.

An alternative to MORDM that ensures a (Pareto) optimal robustness of the final solutions is multi-objective robust optimization, where the objective functions are robustness metrics defined on a large set of scenarios similarly for each outcome indicator (Hamarat et al., 2014; Kwakkel et al., 2015), or heterogeneously by taking different characteristics of the objectives into account (Trindade et al., 2017). Yet, robust optimization is computationally intense since it generates a large number of scenarios at each step of the search process to calculate the robustness of each alternative solution. Therefore, the multi-scenario search extension to MORDM introduced by Watson and Kasprzyk (2017) is one step towards enhancing the robustness of final solutions with a lower computational requirement than robust optimization.

2

Using multiple scenarios in the search phase raises the question of how to determine these scenarios so that the robustness potential of alternatives can be increased. In Watson and Kasprzyk (2017)'s study, these scenarios are selected based on scenario discovery results. Namely, a set of parameter values is picked from the uncertainty ranges that constitute policy vulnerabilities. This approach to scenario selection is strong at identifying policy-relevant scenarios, and the candidate solutions generated under these scenarios have a high potential to address vulnerabilities. However, this approach also has some drawbacks. First, such a reduction of the uncertainty ranges to a few distinct values is rather subjective. Since scenario discovery results in a subset of multiple scenarios, indicated by particular ranges of uncertainties, selecting a discrete set of values from these ranges, and hence a certain number of scenarios, is not straightforward. Therefore, systematic approaches can rationalize this selection process. Second, scenario selection based on vulnerabilities focuses only on severe cases and ignores the diversity of scenarios. Optimization under these severe and policy relevant scenarios is expected to yield solutions that can mitigate vulnerabilities. However, since the solutions resulting from MORDM process are expected to be robust against a wide variety of alternative futures, accounting for this variety in the search phase can be beneficial, too. Third, different policies can have significantly different vulnerabilities. Therefore, selecting scenarios according to the vulnerabilities of only one policy can lead to a pitfall of overlooking highly robust candidate solutions.

This paper follows Watson and Kasprzyk (2017) in generating the candidate solutions for multiple scenarios, yet presents an alternative scenario selection procedure to deal with the abovementioned drawbacks. In other words, this paper demonstrates a revised MORDM application with a formal scenario selection procedure that takes into account both diversity and policy relevance. This revised approach constitutes a sub-process in the search phase of MORDM, which results in the selection of a particular number of policy-relevant and diverse scenarios from a much larger ensemble. To investigate the added-value of this approach, the candidate solutions generated for systematically selected scenarios are compared to the solutions generated for randomly selected scenarios. An alternative way to demonstrate the added-value of this approach is to compare it to the scenario selection method of Watson and Kasprzyk (2017). This comparison, however, is not included in this study because selecting a particular number of scenarios from the scenario discovery results is rather arbitrary as mentioned before, therefore a solid common ground for comparison cannot be found.

To demonstrate this approach to multi-scenario search in MORDM, we use a stylized and hypothetical environmental management problem known as the lake problem where the inhabitants of a city have to decide on the annual pollution amount they release into a lake (Carpenter et al., 1999). The lake problem is often used for the demonstration of the MORDM and related decision making under deep uncertainty tools and techniques (Hadka et al., 2015; Kwakkel, 2017; Quinn et al., 2017; Singh et al., 2015; Ward et al., 2015). Therefore, it provides a common ground for the demonstration of the multi-scenario approach to the search phase of MORDM.

In the remainder of the paper, Section 2 elaborates on the Many-Objective Robust Decision Making approach, the revised MORDM procedure we follow in this study for multi-scenario search, and the scenario selection technique we used. Section 3 introduces the case study, namely the lake problem. Section 4 presents the results of scenario selection and robustness analyses of the candidate strategies found under these scenarios, as well as a comparison between systematically selected scenarios and randomly selected scenarios. Section 5 discusses the proposed approach and future research potential in light of the results. The paper ends with conclusions in Section 6.

2. Methods

178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236

In this section, we elaborate on the scenario selection procedure we followed in light of various existing scenario selection approaches. We then describe the revised MORDM framework we adopted with multi-scenario search.

2.1. Scenario selection for multi-scenario search

Model-based scenario development enables the analysts to include a large number of uncertain factors and to explore a wide variety of future possibilities by systematic variation of uncertain model elements. This is one of the major advantages of model-based scenario development over intuitive approaches based on qualitative knowledge and expert insights (Van Notten et al., 2003). However, it is often necessary to prioritize a small number of these scenarios because (i) users can pay close attention only to a few, (ii) policy evaluation poses computational constraints (Carlsen et al., 2016b), and (iii) limited resources can be allocated more effectively for further scenario development (Comes et al., 2015).

Possible criteria to select a small number of scenarios from a large set are internal consistency, diversity of outcome indicators, extremeness, and policy relevance (Trutnevyte et al., 2016). Existing studies often use a combination of these criteria, depending on the purpose of prioritizing a small number of scenarios. For instance, Tietje (2005) follows an optimization-based selection procedure where internal consistency and diversity of scenarios are maximized. Trutnevyte (2013) chooses a small set of energy-mix scenarios by maximizing a diversity metric defined for each scenario as the Euclidean distance to a reference scenario. Comes et al. (2015) also focus on diversity and construct scenario equivalence classes where two scenarios are considered similar if the weighted average of differences between all outcome indicators (i.e. weighted Manhattan distance) is smaller than a certain threshold. They then select a representative from each class based on reliability. Carlsen et al. (2016a) also maximize diversity for scenario selection, yet they define the diversity of a scenario set, rather than an individual scenario. The diversity metric is then defined as the weighted sum of the minimum and mean values of pairwise distances within the set, where the distance between two scenarios is calculated as the weighted Manhattan distance. Alternatively, Auping et al. (2016) select a number of scenarios based on their extremeness, i.e. whether they result in extreme values, without following an optimization procedure. Carlsen et al. (2016b) combine diversity and policy relevance criteria in an optimization procedure similar to Carlsen et al. (2016a), where policy relevance refers to the extreme values of outcome indicators that make potential policies vulnerable. Watson and Kasprzyk (2017) focus on policy-relevant scenarios, and employ scenario discovery to identify the vulnerabilities of a policy and pick a few values from the uncertainty ranges that create these vulnerabilities.

In this study, our purpose for scenario selection is to find scenarios that potentially lead to solutions that are more robust when the search phase of MORDM is conducted for each of these scenarios. Therefore, among the possible scenario selection criteria listed above, we consider policy relevance and diversity as the relevant ones. Policy-relevance is a problem-specific concept that should reflect the decision-making concerns and preferences. In this study, we choose to define policy-relevance with respect to undesirable scenario conditions specified by the median values of the scenario space. Dividing the scenario space at the median values focuses the diversity analysis on a smaller number of policy-relevant scenarios regardless of their distribution. However, we do not propose this as a general way to define policy relevance. For the diversity criterion, we propose to use the specific diversity maximization method introduced by Carlsen et al. (2016b).

To combine these two selection criteria, we follow a two-step scenario selection procedure: First, we filter the policy-relevant scenarios from a large set based on the threshold conditions specified with the median values. This is similar to specifying scenarios of interest in scenario discovery. Equation 1 denotes this filtering step, where S_N is the entire set of N scenarios, S_M is the subset with selected M

4

scenarios, s refers to scenario indices, $f_{i,s}$ is the value of outcome indicator i in scenario s, and m_i is the threshold value for outcome indicator i that specifies the undesired region. Note that the undesired region is above the threshold if an outcome indicator is to be minimized, and below the threshold if it is to be maximized. Equation 1 denotes only the former case.

$$S_M = \left\{ s \in S_N \mid (\forall i) \left[f_{i,s} \geq m_i \right] \right\} \tag{1}$$

In the second step, we select a small number of scenarios from this subset S_M following the diversity maximization approach of Carlsen et al. (2016b). Namely, from S_M, we choose the subset S_K^* of K scenarios, which has the maximum diversity value among all K-th combinations of the M scenarios as shown in Equation 2, where $D_{K,l}$ is the diversity measure of the subset l and $\binom{M}{K}$ is the number of K-th combinations of M. We define diversity following Carlsen et al. (2016b) as in Equation 3, where w is the weight given to mean distance, *mean* function refers to the arithmetic mean of pairwise distances within set K_l, and $d_{j,k}$ is the distance between scenarios j and k in the set K_l. We use Euclidean distances between the normalized values of outcome indicators as Equation 4 shows, where $\bar{f}_{i,j}$ is the normalized value of $f_{i,j}$, with respect to maximum and minimum values among all M scenarios. We employ an exhaustive search to find the set(s) of K scenarios with maximum diversity among $\binom{M}{K}$ such sets.

$$S_K^* = \max_{l=1..\binom{M}{K}} \left\{ D_{K_l} \right\} \tag{2}$$

$$D_{K_l} = (1-w) \min_{\forall j,k \in K_l} \left\{ d_{j,k} \right\} + w \; mean_{\forall j,k \in K_l} \left\{ d_{j,k} \right\} \tag{3}$$

$$d_{j,k} = \sqrt{\sum_i \left(\bar{f}_{i,j} - \bar{f}_{i,k} \right)^2} \tag{4}$$

2.2. MORDM with multi-scenario search

In this study, to incorporate multiple scenarios in the search phase, we follow a modified version of the MORDM framework of Kasprzyk et al. (2013) of which the main steps are listed in Section 1. This modified version can be described as follows:

Step 1 - Problem formulation: This step involves structuring the decision problem with respect to the XLRM framework as in the original MORDM process.

Step 2 - Scenario selection: The scenario selection procedure described in Section 2.1 is performed in this step, which therefore comprises the following three sub-steps:

2.1) Generate N scenarios with randomly sampled values of deep uncertainties and decision levers

2.2) Filter M policy-relevant scenarios from the ensemble of N scenarios. In this study, we opt to divide the scenario space at medians, and filter the scenarios in which the value of each outcome indicator is in the undesired half of its range across all scenarios.

2.3) From M scenarios, select K maximally diverse scenarios

Step 3 – Generating candidate solutions: In this step, candidate solutions are identified for the reference scenario and for each of the K selected scenarios. The reference scenario is kept in the analysis for a

comparison to the selected scenarios. To search for the candidate solutions, any Multi-Objective Evolutionary Algorithm (MOEA) could be used. We particularly use the ε-NSGAII algorithm (Kollat and Reed, 2006) since it is shown to outperform many MOEA's in terms of scalability, diversity of solutions, and convergence over function evaluations measured by e-indicator and hypervolume success (Reed et al., 2013).

Step 4 – Tradeoff analysis: In this step we examine the robustness tradeoffs of each candidate solution across a scenario ensemble of deep uncertainties and investigate if the candidates generated under different scenarios perform differently. Robustness can be operationalized by using many different metrics (Giuliani and Castelletti, 2016; Herman et al., 2015; Kwakkel et al., 2016a; McPhail et al., 2018), which represent different decision making preferences. For instance, statistical metrics compute a robustness value based on the distribution of the outcomes of interest over the set of scenarios, whereas regret-based metrics consider the regret in each possible scenario if a candidate solution is selected. These metrics can be computed separately for each objective variable, or can be aggregated across the objectives to obtain a single value for each candidate solution.

In this study, to evaluate the robustness of candidate solutions, we use the fraction *mean/standard deviation* and *maximum regret* as the commonly used representatives of the abovementioned two metric groups. We avoid aggregation of the objectives and compute each metric for each objective variable, in alignment with the MORDM framework to explicitly demonstrate robustness tradeoffs between the objectives.

The fraction *mean/standard deviation* represents the concept of a good average result with very limited dispersion around it. Inspired by the signal-to-noise ratio in control theory, the mathematical form of this metric is shown in Equation 5, where $\mu\left(f_{i,j,s}^{*}\right)$ is the mean over the set of scenarios for outcome indicator f_i in the case of implementing the candidate solution j, and σ is the standard deviation. The formulation is converted into a multiplication for objectives to be minimized. In addition, to avoid division by zero, the two terms of the metric are increased by 1, as in previous studies (Hamarat et al., 2014; Kwakkel et al., 2016a).

$$
R_{i,j} =
\begin{cases}
\dfrac{\mu\left(f_{i,j,s}^{*}\right)+1}{\sigma\left(f_{i,j,s}^{*}\right)+1} & ; \text{ if } f_i \text{ is to be maximized; } s=1...N \\[4mm]
\left(\mu\left(f_{i,j,s}^{*}\right)+1\right) \times \left(\sigma\left(f_{i,j,s}^{*}\right)+1\right) & ; \text{ if } f_i \text{ is to be minimized; } s=1...N
\end{cases}
\tag{5}
$$

The *maximum regret* of a candidate solution is the maximal difference across the scenarios between this candidate solution's performance and the zero-regret option's performance in each scenario. Equation 6 shows the formulation of *maximum regret* for candidate solution j and outcome indicator i. In this formulation, $Z_{i,s}$ is the zero-regret value of outcome indicator i in scenario s, which is assumed to be the best value obtained in this scenario with the optimal policy.

$$
R_{i,j} = \max_{s=1...N}\left\{\left|f_{i,j,s}^{*} - Z_{i,s}\right|\right\}
$$
$$
Z_{i,s} =
\begin{cases}
\max\limits_{j=1...J}\left\{f_{i,j,s}^{*}\right\} & ; \ \textit{if } f_i \text{ is to be maximized} \\[3mm]
\min\limits_{j=1...J}\left\{f_{i,j,s}^{*}\right\} & ; \ \textit{if } f_i \text{ is to be minimized}
\end{cases}
\tag{6}
$$

6

The tradeoff analysis step is to be followed by scenario discovery to identify the vulnerabilities of candidate solutions in the applications of MORDM with multi-scenario search, too. However, we do not proceed with scenario discovery in this study, because the primary objective is to demonstrate an extension in the search phase and its effects on the robustness tradeoffs among candidate solutions. Scenario discovery does not directly contribute to this objective, yet it helps to investigate whether the solutions obtained in different scenarios have different vulnerabilities or not, as will be discussed later.

Computational experiments of MORDM can be conducted with the aid of open source packages that support a variety of modelling techniques. In this study, we use Exploratory Modelling and Analysis (EMA) Workbench (Kwakkel, 2017) to generate and visualize the scenarios in Steps 2 and 4. We use the Platypus library (Hadka, 2017) to generate the candidate solutions with the ε-NSGAII algorithm in Step 3. For scenario selection and solution evaluation in Steps 2 and 4, we use our own Python scripts to which a link can be found in the Software and Data Availability section.

3. Case Study: The Lake Problem

We demonstrate our MORDM approach with multi-scenario search on the lake problem. The lake problem is a hypothetical decision problem developed by Carpenter et al. (1999). It has been used in several studies to showcase different ecosystem management approaches and MORDM applications (Hadka et al., 2015; Lempert and Collins, 2007; Peterson et al., 2003; Quinn et al., 2017; Singh et al., 2015; Ward et al., 2015). In the lake problem, the inhabitants of a town decide on annual pollution control strategies. Polluting the lake, by releasing phosphorus in particular, leads to a higher economic utility, but it can also lead to irreversible eutrophication if a threshold is passed. The lake model formulates the phosphorus amount in the lake as in Equation 7, where X_t and a_t are the phosphorus level and the rate of anthropogenic pollution at time t respectively, and b and q are the parameters controlling the rate at which phosphorus is recycled from the sediment and lost from the lake. ε_t is the rate of natural pollution at time t, and following Singh et al. (2015) it is formulated as a log normal distribution with mean μ and standard deviation σ.

$$X_{t+1} = X_t + a_t + \frac{X_t^q}{1 + X_t^q} - bX_t + \varepsilon_t \tag{7}$$

Aligning the lake problem with the XLRM framework for the problem formulation step of MORDM, four performance metrics (M) or outcome indicators can be listed. The first is the *maximum phosphorus concentration* in the lake over time, which is formulated as in Equation 8 where T is the number of time points. The objective of the inhabitants is to minimize this outcome indicator.

$$f_{pollution} = \max_{t=1..T} \{X_t\} \tag{8}$$

The second outcome indicator is the economic *utility* obtained from polluting the lake. This is defined as the discounted net present value of the benefits, as in Ward et al. (2015), without considering the costs of having a polluted lake. In Equation 9 that denotes this outcome indicator, αa_t is the utility derived from polluting, δ is the discount rate and α is constant at 0.04 (Ward et al., 2015). This outcome indicator is to be maximized.

$$f_{utility} = \sum_{t=1}^{T} \alpha a_t \delta^t \tag{9}$$

The *inertia* of anthropogenic pollution is the third outcome indicator and it is defined as the fraction of years where the absolute annual change in anthropogenic pollution is greater than a threshold value.

7

This outcome indicator is formulated in Equation 10, where τ is the threshold value and I is an indicator function which is equal to 1 when the statement is true, and 0 otherwise. *Inertia* indicator ($f_{inertia}$) is to be minimized, since rapid reductions in anthropogenic pollution require large infrastructure investments and the decision makers favor stability. We set τ to 0.02 following Ward et al. (2015).

$$f_{inertia} = \frac{1}{T-1}\sum_{t=1}^{T} I\left(\left|a_t - a_{t-1}\right| > \tau\right) \tag{10}$$

The fourth outcome indicator is *reliability*, defined as the fraction of years where the pollution in the lake is below the critical threshold. In Equation 11, X_{crit} is the critical threshold of pollution and is a function of b and q, I is an indicator function that is equal to 1 if the statement is true, and 0 otherwise. The reliability objective is to be maximized.

$$f_{reliability} = \frac{1}{T}\sum_{t=1}^{T} I\left(X_t < X_{crit}\right) \tag{11}$$

The lake problem involves both stochastic and deep uncertainties. As mentioned before, the natural pollution inflow (ε_t) involves stochastic uncertainty and is formulated with a lognormal distribution. Deep uncertainties are associated with the mean μ and standard deviation σ of the lognormal distribution characterizing the natural inflow, the natural removal rate of the lake b, the natural recycling rate of the lake q, and the discount rate δ. Table 1 specifies the ranges for these deeply uncertain factors as used also in earlier work (Kwakkel, 2017; Quinn et al., 2017), as well as their best estimate or default values.

Table 1: Deeply uncertain parameters

Parameter	Uncertainty range	Default value
μ	0.01 – 0.05	0.02
σ	0.001 – 0.005	0.0017
b	0.1 – 0.45	0.42
q	2 – 4.5	2
δ	0.93 – 0.99	0.98

Each scenario we generate is a computational experiment with a combination of the values of deeply uncertain factors sampled from these ranges. To address the stochastic uncertainty in the rate of natural pollution, multiple replications are performed for each scenario, and the average values of outcome indicators over these replications are used. For instance, the outcome indicator *reliability* in each scenario is formulated as in Equation 12, where R is the number of replications and $f_{reliability}^{r}$ is the *reliability* value in each replication.

$$f_{reliability}^{*} = \frac{1}{R}\sum_{r=1}^{R} f_{reliability}^{r} \tag{12}$$

The lake problem is formulated with two alternative decision mechanisms in previous studies. In several applications (Hadka et al., 2015; Singh et al., 2015; Ward et al., 2015), it involves T decision levers corresponding to the amount of anthropogenic pollution at each time point, namely a_t. An alternative is presented by Quinn et al. (2017), where the control strategy is formulated as a closed loop, i.e. anthropogenic pollution is a function of the current phosphorus level in the lake, leading to highly adaptive decision strategies. In this study, we use the former, non-adaptive version with independent T decision levers, since it is expected to better demonstrate the implications of multi-scenario search for the robustness of candidate solutions.

8

4. Results

Following the problem formulation in Section 3, this section presents the results of an MORDM application with multi-scenario search in the order of steps described in Section 2.2.
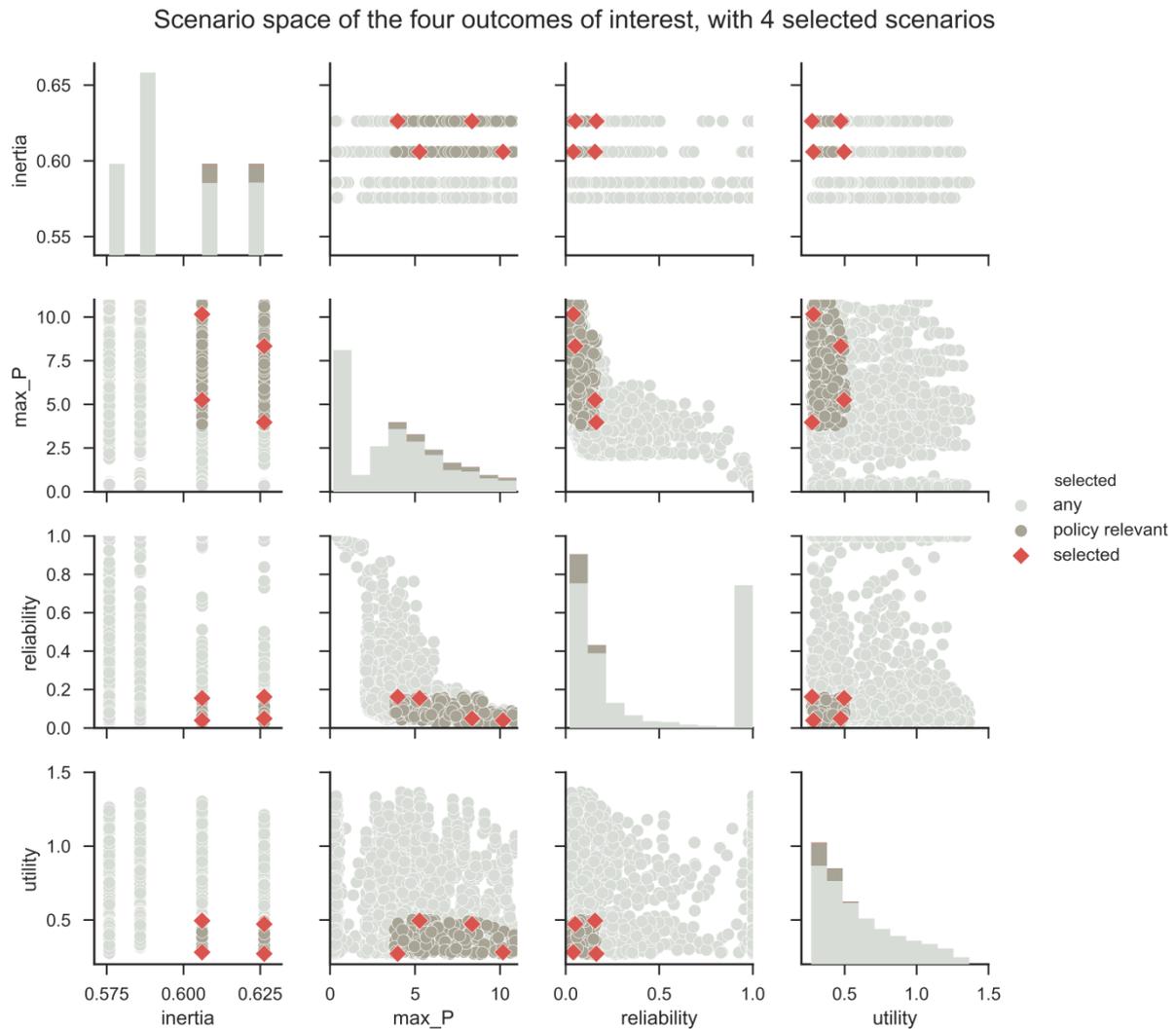
4.1.　　　Step 2: Scenario selection

For multi-scenario search, we first generate *N=2500* scenarios, formed with 5 randomly formed policy strategies and the same 500 scenarios for each strategy. The scenarios are generated by Latin Hypercube Sampling over the ranges of deeply uncertain parameters listed in Table 1, and we perform 100 replications in each scenario with 100 time series samples of the stochastic parameter natural pollution rate. From these 2500 scenarios, we select a subset of policy relevant scenarios, determined according to the median values of each outcome indicator across these 2500 scenarios, as formulated in Equation 13. This subset contains *M=206* scenarios.

$$S_M = \left\{ s \in S_N \left| \begin{array}{l} f_{inertia,s}^* \geq \underset{s=1...N}{\text{median}}\left(f_{inertia,s}^* \right) \\[2mm] and \ f_{pollution,s}^* \geq \underset{s=1...N}{\text{median}}\left(f_{pollution,s}^* \right) \\[2mm] and \ f_{reliability,s}^* \leq \underset{s=1...N}{\text{median}}\left(f_{reliability,s}^* \right) \\[2mm] and \ f_{utility,s}^* \leq \underset{s=1...N}{\text{median}}\left(f_{utility,s}^* \right) \end{array} \right. \right\} \tag{13}$$

From the set of M=206 policy relevant scenarios, we select K=4 maximally diverse ones in terms of their outcomes, with an Euclidean distance calculation and the weight *w=0.5* given to the *mean* of pairwise distances in the diversity metric. As mentioned before, this selection is done by an exhaustive search procedure that evaluates around 72 million subsets with size 4 of these 206 scenarios in terms of the diversity within the subset.

The results of this selection procedure are illustrated in the pairwise scatter plots in Figure 1, where the entire scenario space of 2500 experiments, the subset of policy relevant scenarios and the selected four scenarios are distinctly shown. Each marker represents a scenario, and the histograms on the diagonal axis show how the density varies across the scenario space. The four scenarios selected based on their pairwise distances and colored in red are indeed located in distant corners of the policy-relevant scenario space. In addition, Figure 10 shows the scenario space resulting from scenario discovery, when the scenarios of interest are defined as the policy-relevant failure scenarios with median thresholds. The scenarios in blue are the ones corresponding to the uncertainty ranges of a box found by the PRIM algorithm. This PRIM box covers the 98.5% of the 206 scenarios of interest, and 65% of the scenarios in this box are of interest. From this figure, it can be concluded that the four scenarios we select align with scenarios that could be selected via scenario discovery. Besides, a diversity-based selection is also complementary to a selection based on scenario discovery, since it allows identifying a certain number of scenarios within the uncertainty ranges obtained from PRIM.

Figure 1: Scenario selection results with N=2500 scenarios in light grey, M=206 policy relevant scenarios in darker grey and K=4 diverse scenarios in red

Figure 2 provides an alternative visualization of the selected scenarios on a parallel coordinate plot, where the vertical axes belong to the five deep uncertainties and the four outcome indicators. Each line in this figure corresponds to a scenario and crosses the axes at the uncertainty values and the resulting outcome indicator values in that scenario. The four selected scenarios are highlighted with different colors, and they demonstrate various tradeoffs. Scenario A in green has the most favorable values in terms of *maximum pollution* and *reliability*, yet the worst ones in terms of *inertia* and *utility*. Such a low value for *utility* is attributed to the relatively low value of parameter *delta*, which is the discount rate that determines economic utility. Scenario B is favorable only in terms of *utility*. Scenario C is also an undesirable situation due to very low *reliability* and *utility* values, and a high *maximum pollution*. Scenario D in pink is the best case among these four, with a low *inertia* and *maximum pollution*, as well as a high *reliability* and *utility*. The parameters *b* and *q*, which control the recycling of phosphorus as in Equation 7, have high values in this scenario, as well as the parameter *delta*..

591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
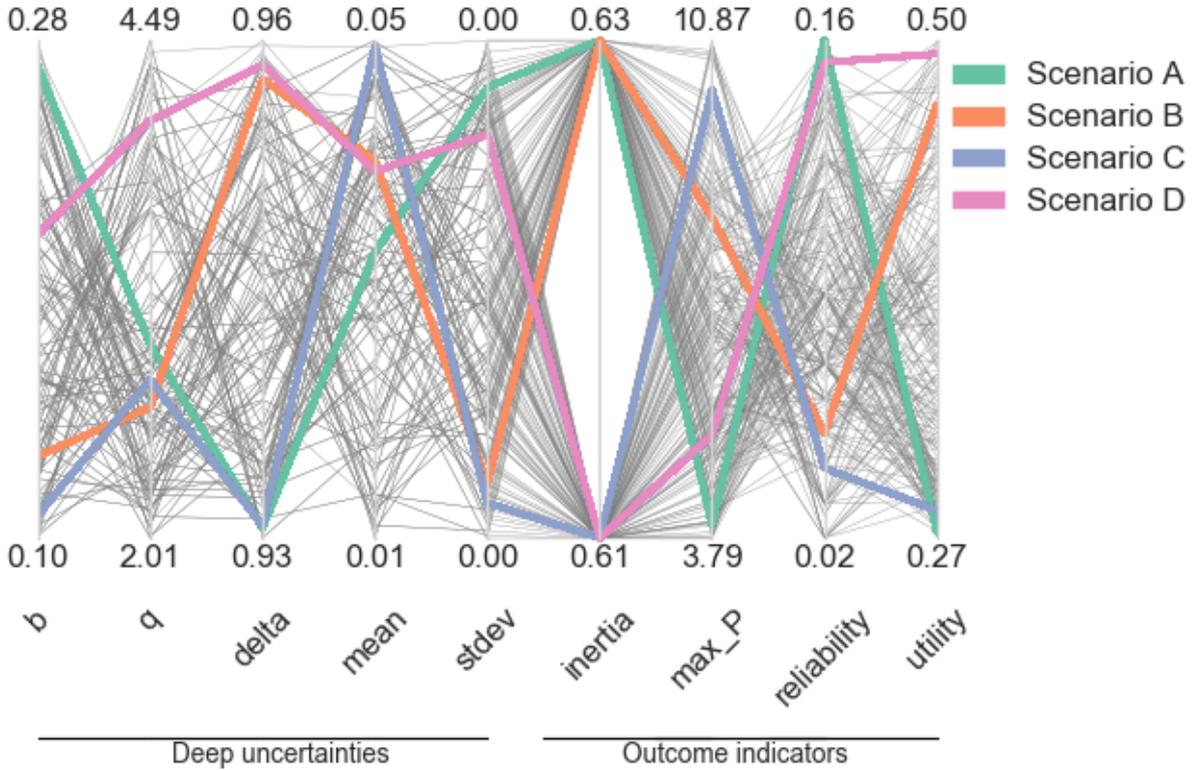638
639
640
641
642
643
644
645
646
647
648
649

Figure 2: The selected four scenarios among the policy-relevant scenarios shown in gray

4.2. Step 3: Generating candidate solutions

In this step, we generate candidate solutions for each of the selected four scenarios and the reference scenario indicated by the default values of uncertainties listed in Table 1. These candidate solutions are the non-dominated tradeoffs in the Pareto fronts identified by the ε-NSGAII algorithm for each scenario. The values of the algorithm parameters used in these search processes can be found in Appendix I, as well as the convergence of the search process for each scenario with these parameters.

The search algorithm generated 105 candidate solutions under the reference scenario, and 96, 89, 88 and 85 solutions for the scenarios *A, B, C and D*, respectively. We restrict this solution space by filtering the ones that result in *reliability* values greater than 80%, assuming that decision makers consider high values of *reliability* crucial and impose this constraint on the solutions. This brushing results in 13, 8, 16, 7 and 6 candidate solutions for the reference scenario and scenarios A, B, C, D respectively.

Figure 3 visualizes these brushed candidate solutions. Each line in this figure corresponds to a candidate solution, and the color refers to the scenario under which it is generated. The four vertical axes show the four objectives. There are considerable differences between the solution sets that can be related to the differences in scenarios under which they are generated, although they are not in the completely distinct regions of the solution space. For instance, in almost all solutions generated under Scenario D, which is considered the best case among the four selected scenarios, very low *pollution*, very high *reliability*, and relatively favorable *inertia* and *utility* values are obtained. Many candidate solutions generated in the reference scenario lead to relatively undesired *inertia*, *pollution* and *reliability* values. Yet, there are still a few solutions generated in this scenario with favorable values of these objectives. Scenario B was a

11

compromise between the four objectives, with moderate values of all, and the candidate solutions generated in this scenario reflect this characteristic, too. However, the candidates generated under Scenario C, which is an undesirable scenario, are highly similar to those generated under Scenario D, which was considered the best among the selected scenarios.
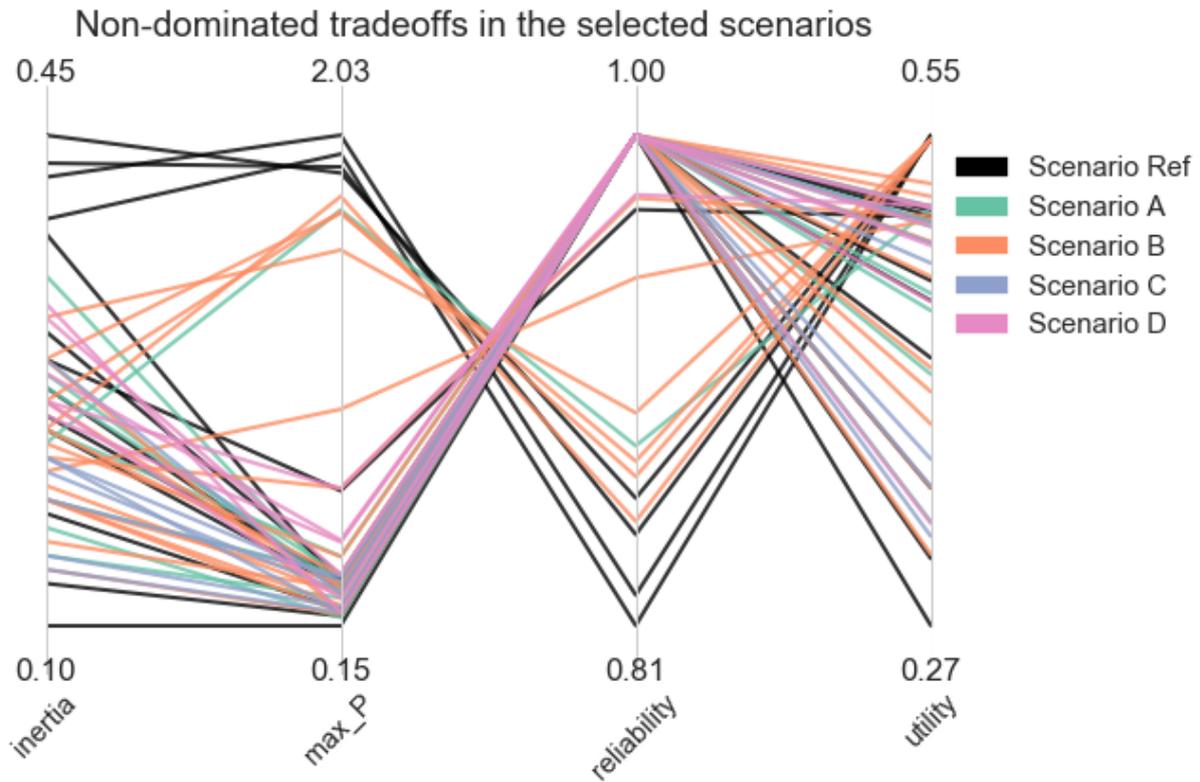

Figure 3: Candidate solutions generated under each selected scenario

Following these observations, we conclude that searching for candidate solutions under multiple scenarios generate a wider variety of tradeoffs than searching only under a reference scenario. Since the goal of the search phase in MORDM is to inform the decision-making debate about potentially good solutions and their consequences, a wider variety of candidate solutions provide a room for maneuver in this debate, for instance to prioritize the objectives or to agree on an acceptable degree of tradeoffs. Therefore, it provides more insight and more options to the decision makers. A direct relationship, however, cannot be drawn between the tradeoffs and the scenario under which they are generated, based only on the results of this case study.

4.3.   Step 4: Tradeoff analysis

In Step 4, the candidate solutions generated under the reference scenario and the four selected scenarios are evaluated in terms of their robustness against deep uncertainties. For this purpose, firstly the performance of each candidate solution is explored across a variety of scenarios. Namely, N=1000 scenarios are generated for each candidate solution, by sampling for the alternative values of the five deep uncertainties listed in Table 1.Then, the robustness of each candidate solution is evaluated based on this ensemble of 1000 scenarios. Appendix II demonstrates the densities of these scenarios, which can support understanding the robustness performance of the candidate solutions.

Figure 4 show the robustness tradeoffs of the candidate solutions evaluated across 1000 scenarios, with respect to the *mean/std. deviation* metric. The candidates generated under multiple scenarios result in a

wider variety of robustness tradeoffs and better robustness values than the ones generated under the reference scenario. Higher values of the *mean/std. deviation* metric are preferred for *utility* and *reliability*, but lower values for *inertia* and *maximum pollution*. Therefore, the first two axes are inverted to indicate that the higher a line on each axis, the higher the robustness of the corresponding solution. Considering these preferences, many solutions generated under the 4 scenarios outperform the solutions generated under the reference scenario, except for the *maximum pollution* objective. Besides, most of the solutions highlight a tradeoff between the *maximum pollution* and *reliability* robustness objectives. Still, some candidate solutions, especially the ones found under scenarios C and D, balance this tradeoff and offer more favorable robustness values both for *reliability* and *maximum pollution*. The extent of tradeoffs generated under Scenario D is low, indicating that solutions generated under this scenario offer compromises between the robustness values for all objectives.
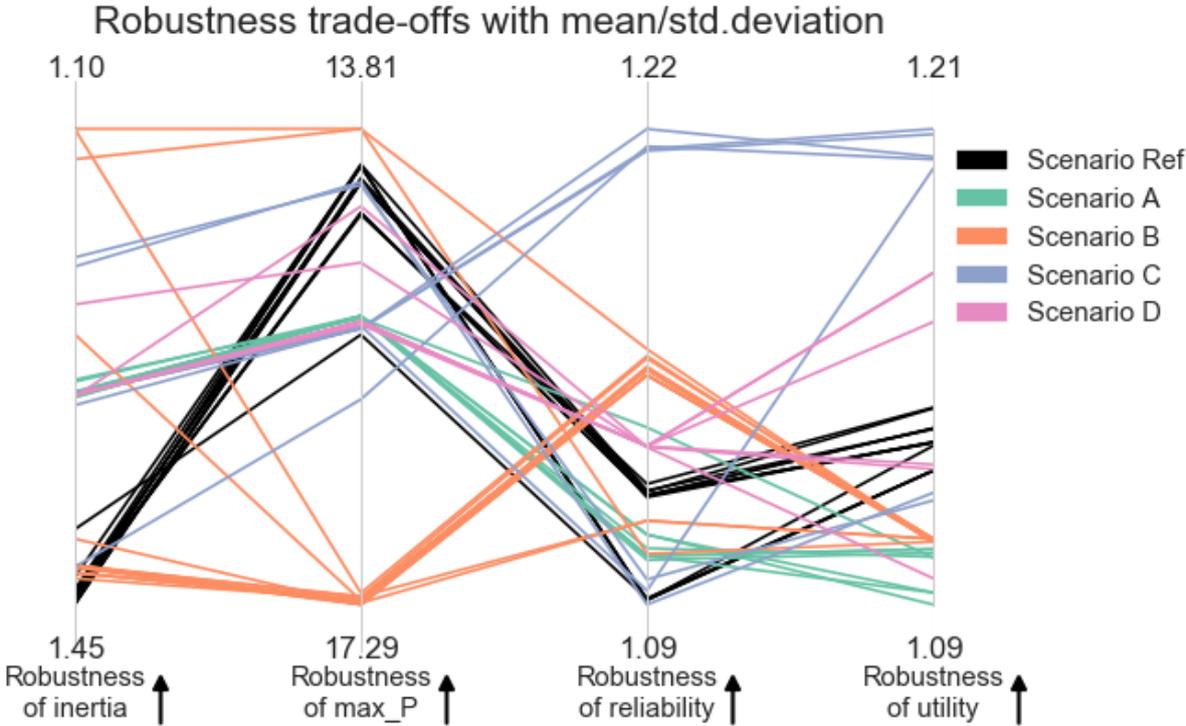

Figure 4: The robustness tradeoffs of the candidate solutions generated under each selected scenario with the mean/standard deviation metric (arrows indicate the preferred direction of values)

When the candidate solutions' robustness is evaluated by the *maximum regret* metric, the results are similar in terms of leading to a wider variety of tradeoffs and higher number of robustness values as Figure 5 illustrates. Considering that the solutions with a low *maximum regret* are preferred for any outcome indicator, most solutions generated for the reference scenario show a tradeoff similar to the results of the *mean/std. deviation* metric, with low robustness for *inertia* and high robustness for *maximum pollution*. The candidates generated for Scenario B also result in highly unfavorable robustness values for the *inertia* and *pollution* objectives. With relatively low values for almost all objectives, the candidates generated for Scenarios A and D present compromise solutions with relatively low maximum regret values for all objectives.
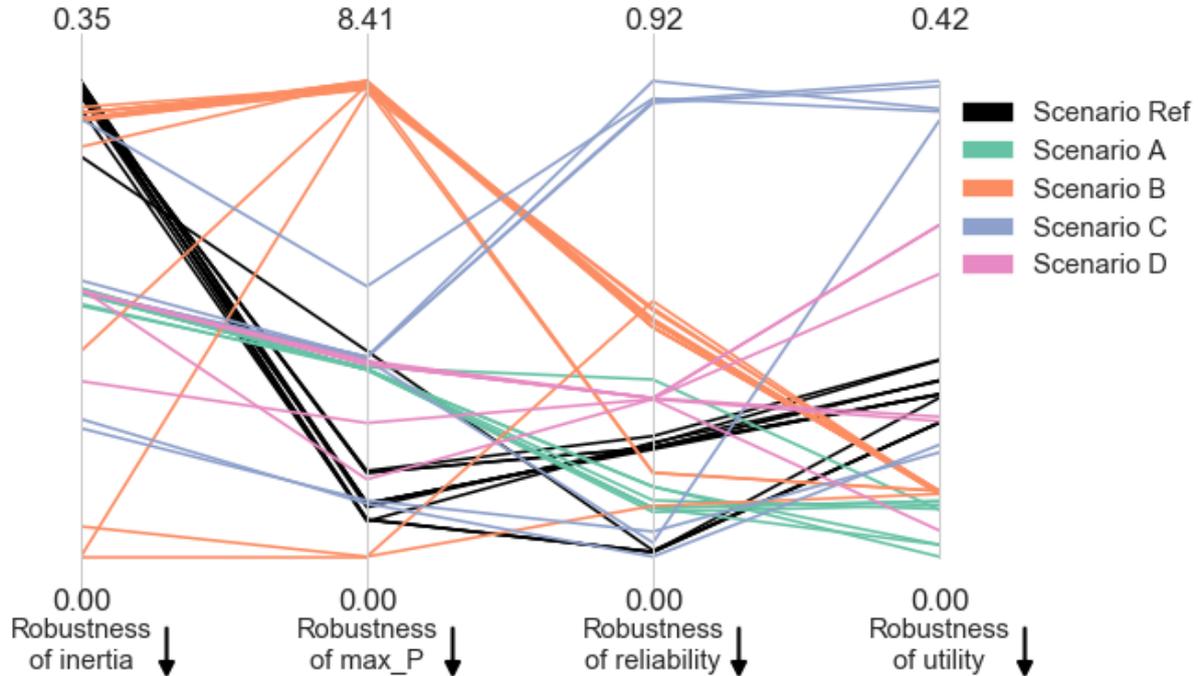
13

Figure 5: The robustness tradeoffs of the candidate solutions generated under each selected scenario with the maximum regret metric (arrows indicate the preferred direction of values)

Given that multi-scenario search has resulted in various robustness tradeoffs, and MORDM application can proceed with a subset of these candidate solutions. This subset can be chosen depending on decision-making preferences, for instance the compromise solutions, or the solutions that enhance the prioritized objectives. These candidate solutions can then be improved according to their vulnerabilities identified by scenario discovery.

4.4. Comparison of using systematically selected scenarios to using randomly selected scenarios in the search phase

This paper introduced a systematic scenario selection procedure to be incorporated into the search phase of MORDM to enhance the robustness potential of the candidate solutions. In this section, we investigate whether using systematically selected scenarios in the search phase of MORDM differs from using randomly selected scenarios in terms of the robustness of the candidate solutions the two scenario groups lead to. Namely, we repeat the MORDM process with multi-scenario search using four randomly selected scenarios instead of the four policy relevant and maximally diverse scenarios identified in Section 4.1. We compare the robustness tradeoffs of the candidate solutions generated for these randomly selected scenarios to those resulted from the systematically selected scenarios presented in Section 4.3.

An arbitrary set of four scenarios is selected from the initial ensemble of 2500 scenarios (See Section 4.1) Figure 11 and 11 in Appendix III show these four randomly selected scenarios. The search phase under these four scenarios yielded 399 candidate solutions in total, which are reduced to 53 after brushing according to the reliability criterion. The performance of each of these candidate solutions is explored in 1000 scenarios, and their robustness scores are calculated. Figure 6 below compares the robustness tradeoffs of these 53 candidate solutions to the 37 candidate solutions generated for the four

827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885

systematically selected scenarios (namely the scenarios A, B, C, D), in terms of the robustness metric *mean/standard deviation*. Note that the results shown in the 'systematic' group of the candidate solutions are an aggregation of those shown in Figure 5, whereas the ones in the 'random' group are newly obtained. A similar comparison between the systematically selected and random scenarios is illustrated in Figure 7 using the *maximum regret* robustness metric.

From the visual comparison in Figures 7 and 8, it can be concluded that, regardless of the robustness metric used, systematically selected scenarios and randomly selected scenarios lead to different robustness tradeoffs. However, there are candidate solutions that are generated for randomly selected scenarios that still lead to very favorable robustness outcomes. Similarly, there are candidate solutions from the systematic group that underperform the ones from the random group.

These two solutions groups can be compared, for instance, in terms of the number of solutions that lead to favorable results, specifically above or below the median objective values within the group. Such favorable results correspond to the upper half of the solution space for the *mean/std. deviation* metric in Figure 6, and the lower half for the *maximum regret* metric in Figure 7. Considering the former, there is only one such solution both in the systematic and random group, i.e. a solution curve which is in the upper half of Figure 6 for all four objectives. Similarly, there are only two such solutions in both groups when the robustness metric maximum regret is considered. Allowing one tradeoff and considering only three of the four objectives, the two solutions groups have again similar numbers of solutions that lead to favorable results. However, considering the total number of solutions in the two groups, a higher portion leads to such favorable groups in the systematic group. Still, the solutions in the two groups do not cover mutually exclusive regions of the solution space. Therefore, in the case of the lake model, selecting the scenarios based on policy relevance and diversity does not lead to significantly more favorable results in terms of robustness or variety of robustness tradeoffs, compared to an arbitrary set of scenarios.

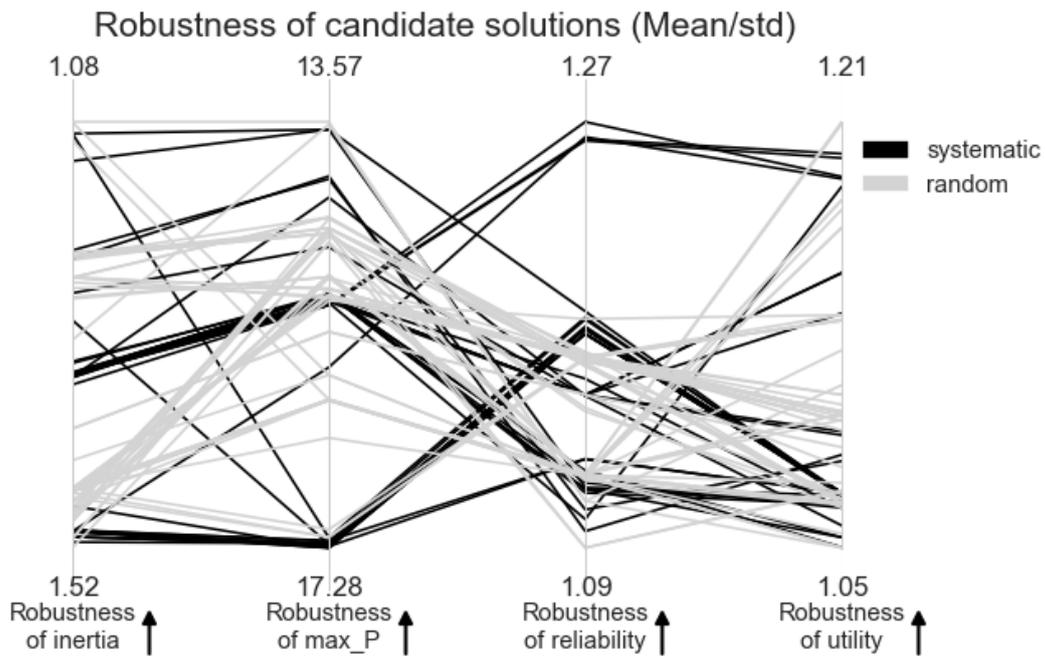## Robustness of candidate solutions (Mean/std)

Figure 6: The robustness tradeoffs of the candidate solutions generated under the systematically and randomly selected scenarios, with the *mean/std. deviation* metric (arrows indicate the preferred direction of values)

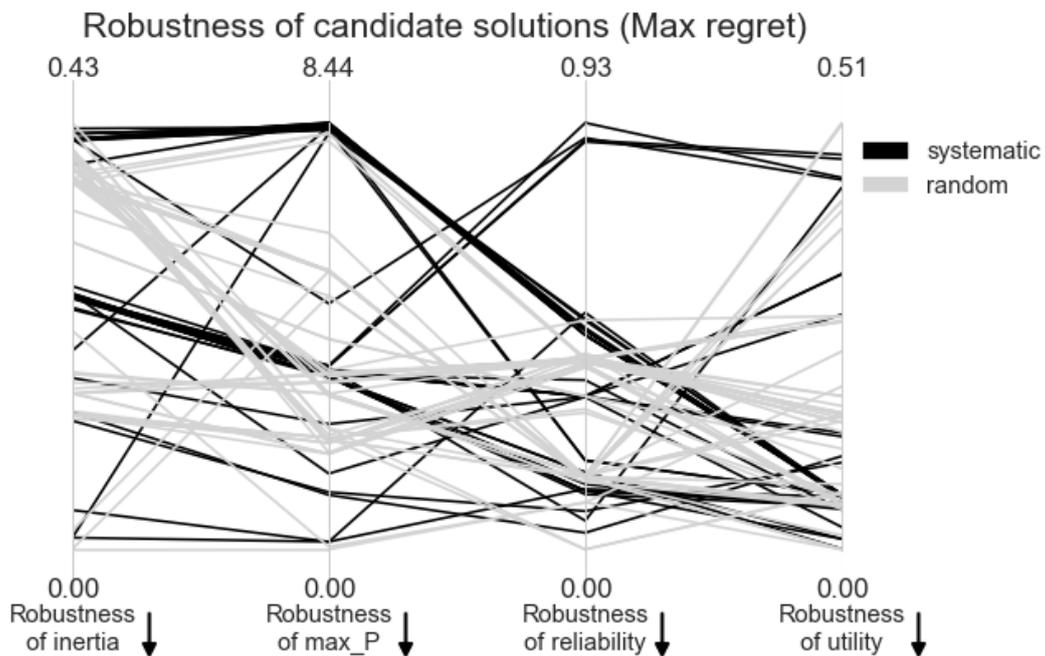## Robustness of candidate solutions (Max regret)

Figure 7: The robustness tradeoffs of the candidate solutions generated under the systematically and randomly selected scenarios, with the *maximum regret* metric (arrows indicate the preferred direction of values)

5. Discussion

This paper incorporated a systematic scenario selection mechanism into the search phase of MORDM in order to optimize a decision problem for many different scenarios and to find potentially more robust solutions. The results show that optimization for different scenarios indeed lead to a wider variety of

16

trade-offs both for the performance indicator values and for their robustness scores compared to optimization only for a reference scenario, and provide the decision makers with more options that can be aligned with different decision-making preferences. Below, we discuss these results and the proposed approach with respect to the scenario selection method and the extended MORDM approach.

5.1. Scenario selection

As discussed in Section 2.1, scenario selection can be based on different criteria such as consistency, extremeness, diversity or policy relevance, and these criteria can be operationalized in various ways. In this study, we combined two of these criteria, policy relevance and diversity. These two criteria are important in a robust decision making context, because they refer to the undesired conditions that pose vulnerabilities for policies and to the variety of conditions that can be encountered due to deep uncertainties.

The selection procedure we choose to use involves an exhaustive search process from maximum diversity with over 70 million evaluations, which is computationally intense. This computational requirement is expected to be higher when the search is conducted on a larger ensemble of scenarios. Therefore, to reduce the computation time and make this approach more accessible, heuristic search algorithms can be used in future applications. In particular, evolutionary algorithms used for multi-objective optimization (Reed et al., 2013) can also be used for finding a maximally diverse set of scenarios. A possible formulation of this maximization problem can include a single objective function, namely the diversity metric, and decision variables corresponding to scenario indices.

In this study, as a part of an MORDM application, scenario selection had the purpose of identifying multiple scenarios under which candidate solutions could be generated. However, scenario selection can be motivated by various other reasons, such as computational constraints, narrative elaboration or policy evaluation as discussed in Section 2.1. Therefore, systematic scenario selection tools based on diversity can support future studies on decision making under deep uncertainty beyond the purpose of a multi-scenario search in MORDM. Thus, it is potentially beneficial to incorporate such tools, in addition to the current scenario discovery tools, in the software packages that are developed for decision making under deep uncertainty, such as the Exploratory Modelling Workbench (Kwakkel, 2017), open MORDM (Hadka et al., 2015) and Project Platypus (Hadka, 2017).

5.2. MORDM with multi-scenario search

Generating candidate solutions under multiple scenarios led to a wider variety of Pareto optimal tradeoffs for the multi-objective decision problem in this study, and provided more decision options compared to optimization only for a reference scenario. Besides, in Section 4.2, we discussed the relation between each scenario and the characteristics of the solution set generated under it. For instance, Scenario D has a low *inertia,* relatively low *maximum pollution* and high *reliability* and *utility* values. The solutions obtained in this scenario result also in very low *pollution*, very high *reliability*, and relatively favorable *inertia* and *utility* values. However, such observations cannot be generalized for all outcome indicators and for all scenarios. Therefore, based only on the results of this study, we cannot derive a direct relation between the scenarios and the solution sets generated under them. This can be attributed to the nonlinearity in the lake model, and to a stronger influence of anthropogenic pollution (policy lever) than the uncertainties (external factors) on the lake problem's objectives. Furthermore, relating the scenarios and the candidate solutions can help to investigate the sensitivity of outcome indicators to the scenarios for which they are optimized. In other words, it reveals whether particular objective values can ever be achieved or not, if the starting point of optimization is a particular scenario.

17

Future applications can benefit from such an analysis, which is not demonstrated in this study since it is beyond the steps of the MORDM framework.

Since optimization for multiple different scenarios led to different results than the optimization for the reference scenario, the selection of these scenarios is indeed important for the eventual outcomes. When compared to randomly selected scenarios, a higher percentage of the solutions generated for systematically selected scenarios resulted in favorable robustness tradeoffs than the ones generated for randomly selected scenarios. Still, the solutions generated for policy-relevant and diverse scenarios did not all result in a higher robustness performance than the solutions generated for randomly selected scenarios. In other words, randomly selected scenarios covered the variety in the scenario space reasonably (see Figure 10 in Appendix III) and resulted in similarly favorable outcomes as the systematically selected scenarios. We suggest that this observation is due to the features of the lake model, i.e. the scenarios not considerably affecting the characteristics of the solutions found, nonlinearities in the model, or the extent of tradeoffs between the objectives, as well as the ability of 4 scenarios to capture the variety of such tradeoffs. Future research can investigate a suitable number of scenarios to be included in a multi-scenario search, with respect to the number of objectives in the problem and the nature of tradeoffs between them.

In this study, we treated the solutions generated under different scenarios independently and we did not compare them to each other in terms of dominance. The motivation behind this choice was to demonstrate how the performance of candidate solutions generated under different scenarios differ from each other. However, a solution found under one of the scenarios can be dominated by a solution found under another scenario, if the latter results in equal or better values for all objectives in the second scenario. Therefore, further analysis can investigate how a solution generated for a particular scenario performs if re-evaluated in another scenario, i.e. whether it would be dominated or not. This type of analysis strays from the general MORDM framework, since the robustness analysis we presented in this study already covers the performance of a solution in other scenarios. Still, this additional analysis can provide disaggregated insights for particular scenarios of interest if needed.

Furthermore, as robust decision making, MORDM focuses on static robustness, and does not consider the dynamic nature of uncertainties and dynamic decision responses to their realizations (Kwakkel et al., 2016b). In the lake problem example used in this study, we optimized a dynamic decision strategy, yet did not consider adaptation pathways, i.e. switches between different policy options over time according to the realizations of uncertainties. To enhance the robustness of final solutions, MORDM can also benefit from considering dynamic adaptation in future studies.

The method proposed in this study is an extension to MORDM in order to generate candidate solutions that are more robust without the expense of computational power required by many objective robust optimization. This method also inherits the usability and ease of interpretation of the original MORDM framework, since the multi-scenario search is only a repetition of the original search step, and once a combined set of candidate solutions is generated for multiple scenarios, the rest of the MORDM application proceeds as in the original framework. Whether scenario selection is performed as proposed in this study or differently, an MORDM application with multi-scenario search is a promising extension in order to provide solutions with enhanced robustness and a wider variety of tradeoffs to a decision-making debate. Therefore, for the problems where robustness is important, and a wider variety of alternative solutions is required to reach a consensus on the conflicting objectives, it is recommended to use MORDM with multi-scenario search.

18

6. Conclusions

Many objective robust decision making (MORDM) is one of several model-based decision support approaches for designing strategies with robust performance over a large ensemble of scenarios. MORDM generates candidate solutions using many objective optimization given a reference scenario, and subsequently evaluates the robustness of these solutions, i.e. how well they perform under a variety of scenarios. However, there is no guarantee that solutions that are optimal in the reference scenario will have good robustness over a larger set of scenarios. A recent extension to MORDM sought to address this by performing many objective optimization for a number of scenarios, and select these scenarios from that part of the uncertainty space that most clearly negatively influences plan performance (Watson and Kasprzyk, 2017). In this paper, we sought to further improve upon this, by putting forward a systematic approach for scenario selection.

The particular MORDM extension proposed in this paper is to systematically select the scenarios under which the search step is repeated. Scenario selection is based on diversity and policy relevance. Diversity refers to how these scenarios differ from each other in terms of the outcome indicator values, and policy-relevance refers to whether these scenarios are undesirable for decision makers. These two criteria of scenario selection align with the purpose of generating candidate solutions that can perform well under a variety of conditions to which potential solutions are vulnerable.

The results show that, compared to the candidate generation only under a reference scenario, searching for candidate solutions under the selected scenarios not only increases the variety of robustness tradeoffs between the objectives and provide more decision options, but also lead to more robust solutions. In the case of the lake model, the results did not demonstrate considerable differences between using policy relevant and diverse scenarios and using randomly selected scenarios in the search phase. The comparison did, however, highlight that the robustness of candidate strategies is sensitive to the scenario for which it was found to be optimal.

In future MORDM applications, the importance of a systematic scenario selection can be investigated on different cases. The computational expense of the scenario selection procedure described in this paper can be reduced by using heuristic search algorithms for diversity maximization. The software packages developed for decision making under deep uncertainty can benefit from incorporating systematic scenario selection tools either for multi-scenario search in MORDM or for other purposes such as policy evaluation or narrative elaboration. Another departure point for future research is a detailed investigation of the sensitivity of robustness outcomes to the initial scenarios for which optimization is performed. Furthermore, a comparison of multi-scenario MORDM to multi-objective robust optimization helps elaborate on advantages and disadvantages of the two methods.

19

References

Auping, W.L., Pruyt, E., de Jong, S., Kwakkel, J.H., 2016. The geopolitical impact of the shale revolution: Exploring consequences on energy prices and rentier states. Energy Policy 98 390-399.

Bankes, S., 1993. Exploratory Modeling for Policy Analysis. Operations Research 41(3) 435-449.

Bankes, S., Walker, W.E., Kwakkel, J.H., 2013. Exploratory modeling and analysis, Encyclopedia of operations research and management science. Springer, pp. 532-537.

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.

Carlsen, H., Eriksson, E.A., Dreborg, K.H., Johansson, B., Bodin, Ö., 2016a. Systematic exploration of scenario spaces. Foresight 18(1) 59-75.

Carlsen, H., Lempert, R., Wikman-Svahn, P., Schweizer, V., 2016b. Choosing small sets of policy-relevant scenarios by combining vulnerability and diversity approaches. Environmental Modelling & Software 84 155-164.

Carpenter, S.R., Ludwig, D., Brock, W.A., 1999. Management of Eutrophication for Lakes Subject to Potentially Irreversible Change. Ecological Applications 9(3) 751-771.

Coello, C.A.C., Lamont, G.B., Van Veldhuizen, D.A., 2007. Evolutionary algorithms for solving multi-objective problems. Springer.

Comes, T., Wijngaards, N., Van de Walle, B., 2015. Exploring the future: Runtime scenario selection for complex and time-bound decisions. Technological Forecasting and Social Change 97 29-46.

Giuliani, M., Castelletti, A., 2016. Is robustness really robust? How different definitions of robustness impact decision-making under climate change. Climatic Change 135(3-4) 409.

Hadka, D., 2017. Platypus - Multiobjective Optimization in Python [Computer Software]. <http://platypus.readthedocs.io/en/latest/>.

Hadka, D., Herman, J., Reed, P., Keller, K., 2015. An open source framework for many-objective robust decision making. Environmental Modelling and Software 74 114-129.

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.

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.

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.

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.

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

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.

Kwakkel, J.H., Eker, S., Pruyt, E., 2016a. How robust is a robust policy? Comparing alternative robustness metrics for robust decision-making, In: Doumpos, M., Zopounidis, C., Grigoroudis, E. (Eds.), Robustness Analysis in Decision Aiding, Optimization, and Analytics. Springer.

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.

Kwakkel, J.H., Haasnoot, M., Walker, W.E., 2016b. Comparing Robust Decision-Making and Dynamic Adaptive Policy Pathways for model-based decision support under deep uncertainty. Environmental Modelling & Software 86 168-183.

Kwakkel, J.H., Jaxa-Rozen, M., 2016. Improving scenario discovery for handling heterogeneous uncertainties and multinomial classified outcomes. Environmental Modelling & Software 79 311-321.

Kwakkel, J.H., Walker, W.E., Marchau, V.A.W.J., 2010. Classifying and communicating uncertainties in model-based policy analysis. International Journal of Technology, Policy and Management 10(4) 299-315.

Lempert, R., 2013. Scenarios that illuminate vulnerabilities and robust responses. Climatic Change 117(4) 627-646.

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.

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.

Lempert, R.J., Popper, S.W., Bankes, S.C., 2003. Shaping the Next One Hundred Years: New methods for quantitative, long-term policy analysis. RAND, Santa Monica, CA.

McPhail, C., Maier, H.R., Kwakkel, J.H., Giuliani, E., 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.

Peterson, G., Carpenter, S., Brock, W.A., 2003. Uncertainty and the management of multistate ecosystems: an apparently rational route to collapse. Ecology 84(6) 1403-1411.

Quinn, J.D., Reed, P.M., Keller, K., 2017. Direct policy search for robust multi-objective management of deeply uncertain socio-ecological tipping points. Environmental Modelling & Software 92 125-141.

Reed, P.M., Hadka, D., Herman, J.D., Kasprzyk, J.R., Kollat, J.B., 2013. Evolutionary multiobjective optimization in water resources: The past, present, and future. Advances in Water Resources 51 438-456.

Singh, R., Reed, P.M., Keller, K., 2015. Many-objective robust decision making for managing an ecosystem with a deeply uncertain threshold response. Ecology and Society 20(3).

Thissen, W.A.H., 2013. Diagnosing Policy Problem Situations, In: Thissen, W.A.H., Walker, W.E. (Eds.), Public Policy Analysis: New Developments. Springer: New York.

Tietje, O., 2005. Identification of a small reliable and efficient set of consistent scenarios. European Journal of Operational Research 162(2) 418-432.

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.

Trutnevyte, E., 2013. EXPANSE methodology for evaluating the economic potential of renewable energy from an energy mix perspective. Applied Energy 111 593-601.

Trutnevyte, E., Guivarch, C., Lempert, R., Strachan, N., 2016. Reinvigorating the scenario technique to expand uncertainty consideration. Climatic Change 135(3-4) 373-379.

Van der Lei, T., Enserink, B., Thissen, W.A., Bekebrede, G., 2011. How to use a systems diagram to analyse and structure complex problems for policy issue papers. Journal of the Operational Research Society 62(7) 1391-1402.

Van Notten, P.W.F., Rotmans, J., van Asselt, M.B.A., Rothman, D.S., 2003. An updated scenario typology. Futures 35(5) 423-443.

Walker, W.E., 2000. Policy analysis: a systematic approach to supporting policymaking in the public sector. Journal of Multicriteria Decision Analysis 9(1-3) 11.

Walker, W.E., Marchau, V.A., Kwakkel, J.H., 2013. Uncertainty in the framework of policy analysis, In: Thissen, W.A.H., Walker, W.E. (Eds.), Public Policy Analysis. Springer: New York, pp. 215-261.

Ward, V.L., Singh, R., Reed, P.M., Keller, K., 2015. Confronting tipping points: Can multi-objective evolutionary algorithms discover pollution control tradeoffs given environmental thresholds? Environmental Modelling & Software 73 27-43.

Watson, A.A., Kasprzyk, J.R., 2017. Incorporating deeply uncertain factors into the many objective search process. Environmental Modelling & Software 89 159-171.

1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298

Appendix I: Operator parameters and the convergence of ε-NSGAII

Table 2: Operator parameters of the search algorithm (ε-NSGAII)

Parameter	Value
NFE	10000
Epsilons (for *maximum pollution, utility, inertia and reliability*, respectively)	0.05, 0.01, 0.005, 0.005
Initial population size	100
Simulated binary crossover (SBX) rate	1
SBX Distribution Index	15
Polynomial Mutation (PM) rate	1
PM Distribution Index	20
Injection rate	0.25
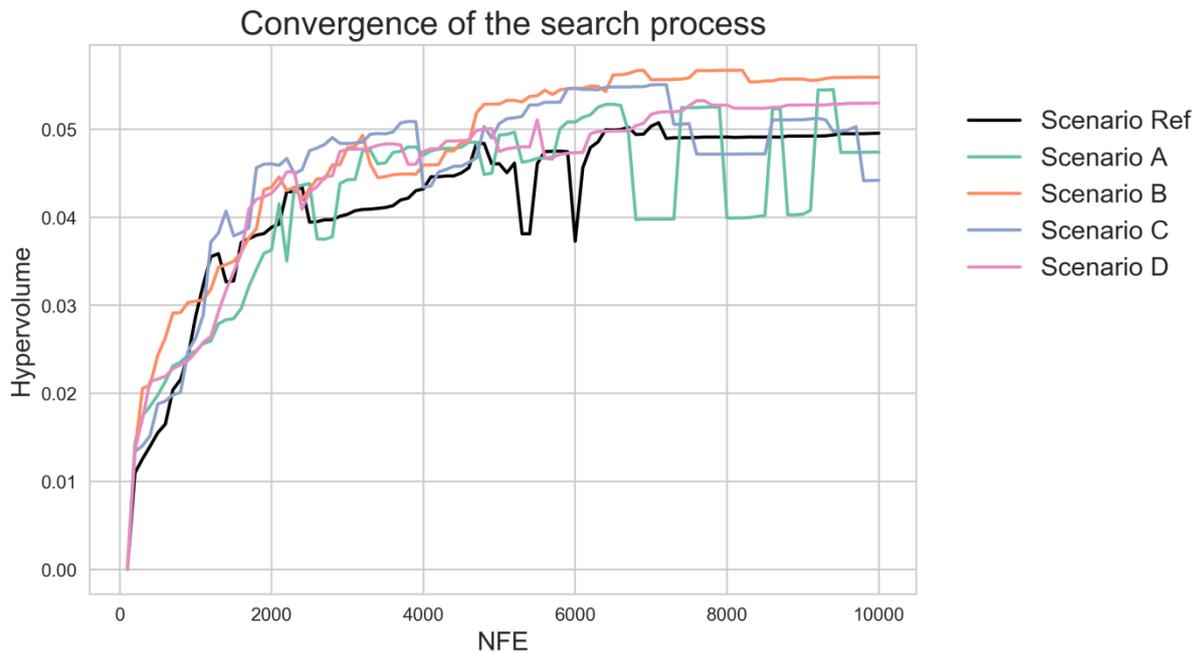
Figure 8: Convergence of the search process under each selected scenario

1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357

Appendix II: Density distributions of the 1000 scenarios generated for each candidate solution
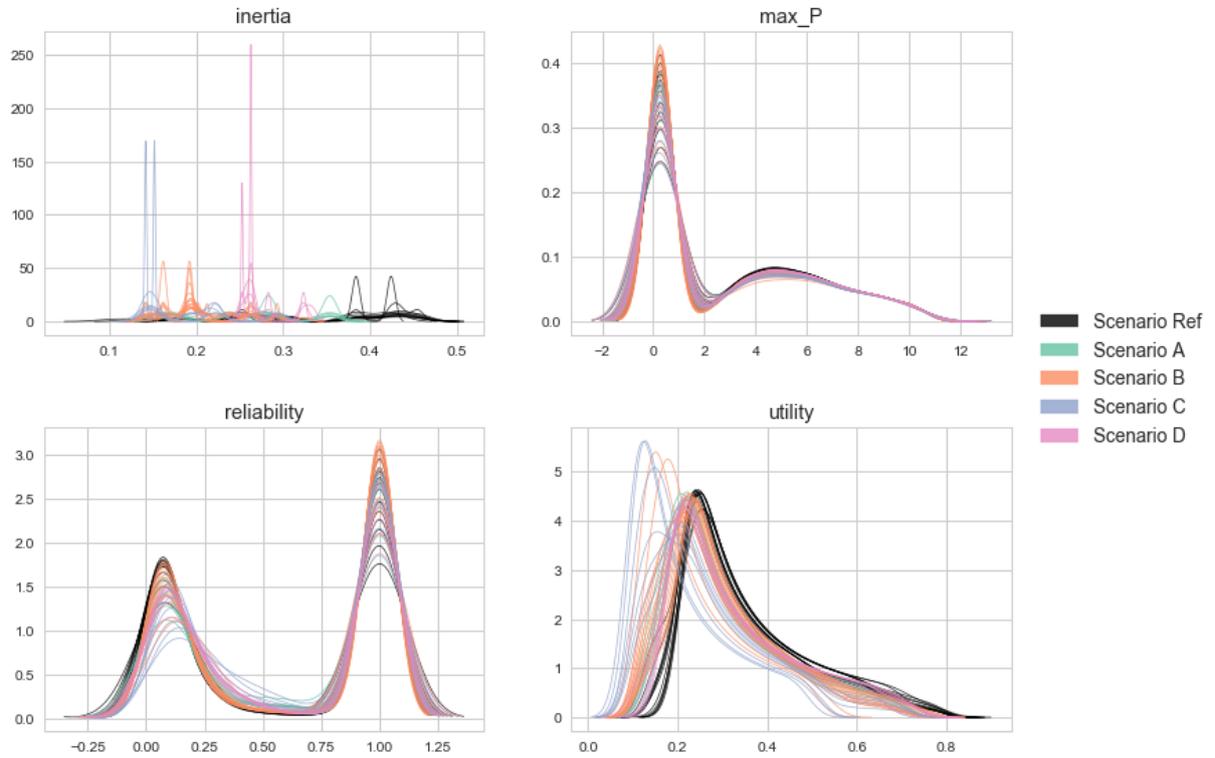
Figure 9: Densities of the scenario ensembles generated with each candidate solution

1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416

Appendix III: Scenario selection

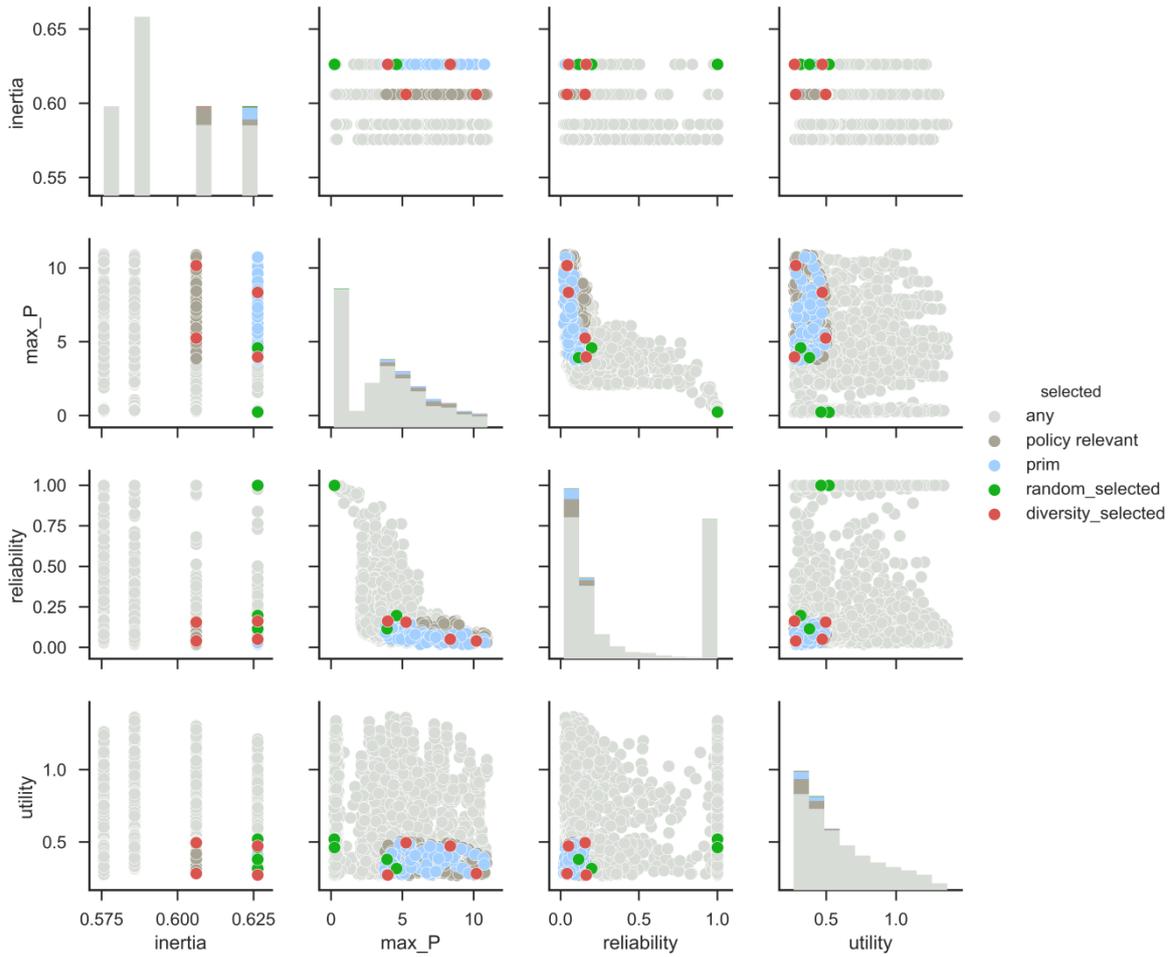Scenario space of the four outcomes of interest, with the 4 systematically and randomly selected scenarios

Figure 10: Scenario selection results with N=2500 scenarios in light grey, M=206 policy relevant scenarios in darker grey and K=4 diverse scenarios in red, and 4 randomly selected scenarios in purple
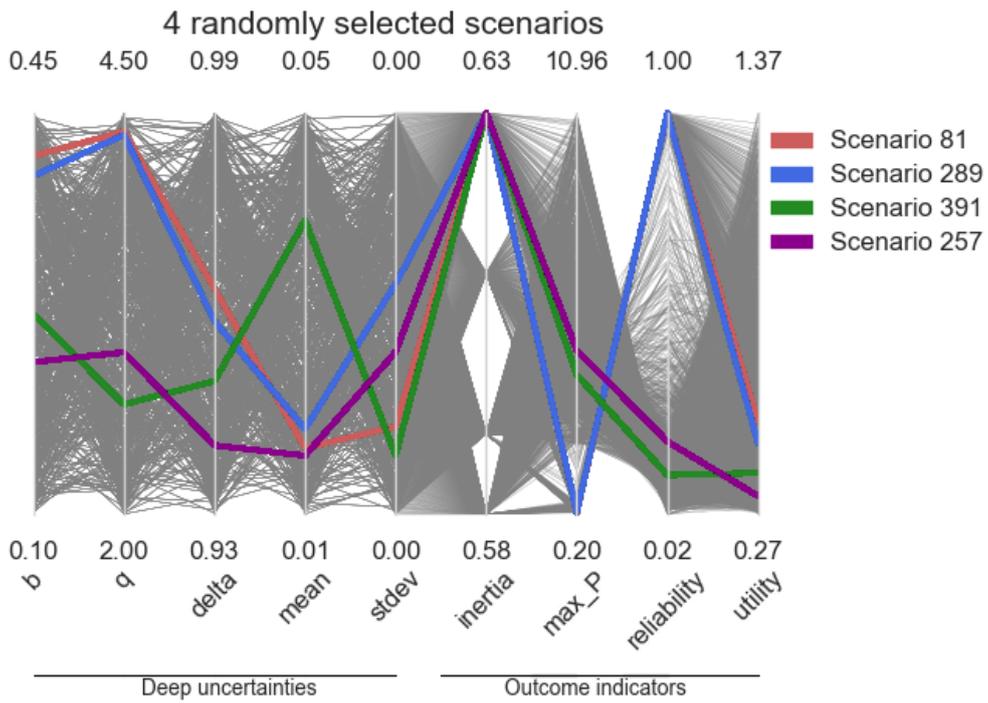
Figure 11: The randomly selected four scenarios among all scenarios shown in gray