NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

UNCERTAINTY IN RISK ASSESSMENT DERIVED FROM SUBJECT MATTER EXPERT ELICITATION

by

Christopher D. Newcomb

June 2023

Thesis Advisor:	Robert A. Koyak
Second Reader:	Louis Chen

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** June 2023	**3. REPORT TYPE AND DATES COVERED** Master's thesis

4. TITLE AND SUBTITLE UNCERTAINTY IN RISK ASSESSMENT DERIVED FROM SUBJECT MATTER EXPERT ELICITATION	**5. FUNDING NUMBERS**
6. AUTHOR(S) Christopher D. Newcomb	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.	**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

This thesis studies uncertainty that is associated with using subject matter expert (SME) elicitation to assess risk. The U.S. Army Engineer Research & Development Center (ERDC) developed the Operational Condition Assessment (OCA) Prioritization Tool to evaluate expected risk to a watershed based upon the condition of components in existing flood control infrastructure. Multiple SME rankings of component criticality are used in conjunction with gate failure consequence data to assign levels of risk to each component and prioritize components associated with the most risk for an OCA. We study three aspects of using the OCA Prioritization Tool. First, we utilize a method of comparing the data from individual SMEs to their peers to assess the consensus in rankings. Next, we utilize the nonparametric bootstrap to quantify the uncertainty in risk mitigation that comes from SME elicitation. Finally, we analyze the effects of having SMEs only ranking a select number of components instead of the entire set.

14. SUBJECT TERMS expert elicitation, rankings, risk assessment, operational condition assessment, OCA	**15. NUMBER OF PAGES** 61
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

UNCERTAINTY IN RISK ASSESSMENT DERIVED FROM SUBJECT MATTER EXPERT ELICITATION

Christopher D. Newcomb
Major, United States Army
BS, Virginia Military Institute, 2011

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
June 2023

Approved by: Robert A. Koyak
Advisor

Louis Chen
Second Reader

W. Matthew Carlyle
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

This thesis studies uncertainty that is associated with using subject matter expert (SME) elicitation to assess risk. The U.S. Army Engineer Research & Development Center (ERDC) developed the Operational Condition Assessment (OCA) Prioritization Tool to evaluate expected risk to a watershed based upon the condition of components in existing flood control infrastructure. Multiple SME rankings of component criticality are used in conjunction with gate failure consequence data to assign levels of risk to each component and prioritize components associated with the most risk for an OCA. We study three aspects of using the OCA Prioritization Tool. First, we utilize a method of comparing the data from individual SMEs to their peers to assess the consensus in rankings. Next, we utilize the nonparametric bootstrap to quantify the uncertainty in risk mitigation that comes from SME elicitation. Finally, we analyze the effects of having SMEs only ranking a select number of components instead of the entire set.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AFB	Air Force Base
DOD	Department of Defense
ERDC	Engineer Research and Development Center
MAE	Mean Absolute Error
OCA	Operational Condition Assessment
SME	Subject Matter Expert
USACE	United States Army Corps of Engineers

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

The Engineer Research and Development Center (ERDC) developed an Operational Condition Assessment (OCA) Prioritization Tool to prioritize which components receive an OCA to maximize the mitigation of risk a watershed faces from infrastructure failures (Alt et al. 2021). The tool relies on input data from a hydrologic consequence model and subject matter expert (SME) elicitation of component rankings in order of criticality to system functionality. This thesis seeks to analyze the uncertainty in the output of the tool for two watersheds: the Trinity River Basin in Texas, and the Willamette River Basin in Oregon.

The purpose of this thesis is to answer three research questions. The first question is: How can we quantify the level of agreement of an individual SME's rankings relative to their peers? To answer this, we determine the absolute difference of each SMEs rankings from the set of their peers, assign weights to differences based on the consensus ranking for each component and sum the values across all components. Evaluation of the two data sets highlighted several SMEs that consistently rank components differently than their peers. While variability in rankings is not necessarily bad, it does create cause for further investigation into the SMEs rankings to determine whether elicitation retraining is needed or if the SME possesses some knowledge their peers lack.

The second research question this thesis seeks to answer is: How can we quantify the variability of mitigated risk estimated by the OCA Prioritization Tool that is attributed to SME component rankings? We use the nonparametric bootstrap to resample which SME rankings are used as the inputs for the tool. We measure the mitigated risk associated with each set of rankings to determine the variability resulting from using different SME rankings. We execute this across six budget levels constraining the number of components the tool selects. We conclude that there is no general rule for the behavior of variability across budget levels in different watersheds. Each watershed is unique, and the variability is reliant upon different SMEs, watershed structure, and consequence of gate failure.

The third research question is: Are there more efficient SME elicitation techniques that can be used? We limit the number of components SMEs rank and assign a value equal to the average of the remaining ranks to the unranked components. We use the nonparametric bootstrap to evaluate the uncertainty and assess the outputs to determine whether truncating rankings greatly reduces the mitigated risk to a watershed. We find that, by limiting the number of components ranked to ten, the OCA Prioritization Tool still mitigates 98 percent of the risk that it would have under a full ranking set. This suggests that the values ranked below ten can be approximated by assigning an average value of the remaining ranks.

Our work shows that SME elicitation injects some uncertainty into the results of the OCA Prioritization Tool. We demonstrate ways to identify SMEs that rank components noticeably different than their peers and suggest the reasons for those differences should be investigated to possibly reduce the variability of SME rankings. We also examine whether limiting the number of components that a SME ranks negatively impacts the effectiveness of a tool. Our findings suggest that at a ranking limit of ten and across all budget levels, the OCA Prioritization Tool mitigated risk value is equal to 98 percent of the full rankings. Further research is required to determine if by limiting the number of components SMEs can rank, the overall ranking quality is improved.

There remain many areas of this topic that offer opportunities for continued research. During exploratory data analysis, we observed that components presented for ranking earlier to SMEs tend to display a higher ranking than components presented later. Future research can inform whether there is a bias in component rankings based upon component presentation order. Additionally, this thesis examines the selection of components over one time period. Investigating the effects of truncation over multiple selection periods may indicate whether there are unforeseen consequences by limiting SME component rankings. Finally, analyzing additional SME component rankings in different watersheds can help validate the conclusions we present in this thesis.

References

Alt JK, Richards JP, Gallarno GE, Brown WH (2021) Risk-informed prioritization of Operational Condition Assessments. *IIE Annual Conference Proceedings* (Norcross, GA), https://www.proquest.com/scholarly-journals/risk-informed-prioritization-operational/docview/2560888656/se-2.

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

Thank you to the team at the Engineer Research and Development Center for providing all the support required to make this research possible. I would also like to thank my advisor, Professor Robert Koyak. Professor Koyak's guidance was immensely helpful through all phases of the thesis research process. Thank you to Professor Louis Chen for his input as well. Finally, this endeavor would not have been possible without my wife, Hedei, and son, Cortland. I will always be grateful for their support during my studies at Naval Postgraduate School.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

The United States Army Corps of Engineers (USACE) is responsible for maintaining over $232 billion worth of water resource infrastructure (Alt et al. 2021). This infrastructure is responsible for mitigating the consequences of flooding, which can have a negative impact on Department of Defense (DOD) installations as well as the lives and property of citizens.

In 2019, a winter storm impacted Nebraska and set the conditions for massive flooding along the Platte and Missouri Rivers. Offutt Air Force Base was inundated with 720 million gallons of flood water that "affected 137 base facilities, 1.2 million square feet of workspace, including $230 million of simulators" (Offutt Air Force Base [AFB], 2019). This demonstrates the risk DOD installations, civilian population centers, and agricultural areas face from flooding rivers.

The floods in Nebraska resulted from a unique combination of heavy rainfall on thick snowpack causing floodwater to flow over levees protecting the airbase (Hasemyer, 2019). While this instance is the result of intense flooding due to rainfall, similar circumstances could occur due to the failure of gates in a dam. Therefore, it is vital to mitigate as much risk as possible from dam failures through implementation of a flood control infrastructure inspection program.

Maintaining the existing flood control infrastructure is a costly endeavor and each individual component of a system is required to periodically undergo an Operational Condition Assessment (OCA) (Alt et al. 2021). This policy places a high priority on ensuring each component is operating correctly while paying a high price in resources spent on conducting component assessments.

A. OCA PRIORITIZATION TOOL

The Engineer Research and Development Center (ERDC) developed a tool to prioritize scheduling of component OCAs based upon the potential risk of component degradation. This would enable more efficient allocation of resources while maximizing the risk mitigated from equipment failures.

Figure 1 illustrates the relationships between flood control components and systems in a watershed. The tool is intended for use at the watershed level and encompasses every facility within that watershed, every gate system within each facility, and every component of each gate system. This tool requires input from district and facility-level subject matter experts (SME) and flood consequence models. This data is subsequently used to execute a mixed integer program to select components for an OCA (Alt et al. 2021).

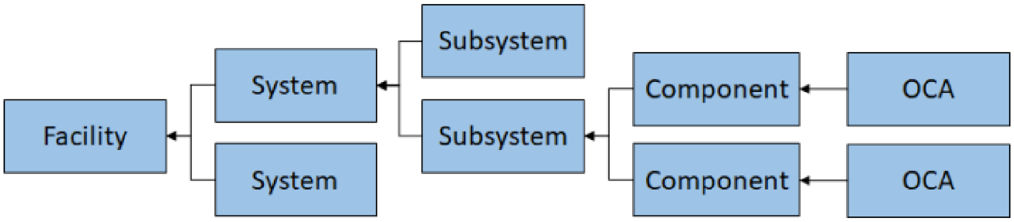


Figure 1. OCA Hierarchy. Source: Alt et al. (2021).

USACE does not possess an abundance of data on the influence of specific gate components on gate failures, so SMEs provide rankings of the influence of each component to gate operability. The rankings are then compiled and normalized to give a relative influence of each component to that system a value between 0 and 1. The relative influence value for each component is multiplied by the consequence of the gate failure to assign a risk level to each component.

The consequence data is derived from simulation models utilizing the Hydrologic Engineering Center Watershed Assessment Tool (Alt et al. 2022). The Watershed Assessment Tool perturbs the number of operational gates and the consequences associated with those circumstances. For smaller watersheds, a full factorial experimental design allows analysts to explore all possible combinations of operational gates. However, for larger watersheds, a more elaborate experimental design must be used, and the outcome translated into a statistical model to gain insight into consequences of specific gate failures.

The tool then inputs the gathered data into a mixed integer program to identify the components for an OCA. The objective is to select the components that maximize the amount of mitigated risk within a given budget of available inspections. An additional stipulation is that selection of one component implies selection of all the same components of a specific gate system, at a specific facility. For example, if the Beaver Creek Dam service discharge gate system consists of three identical gates, a component selected for inspection from that gate system will incur a cost of three inspections, one for each gate.

B. RESEARCH OBJECTIVES

As the OCA Prioritization Tool is developed, continual updates must be made as the scope of the tool grows from focusing on the selection of components of an individual facility with few subsystems to analyzing the outcomes of major watersheds with numerous facilities and subsystems. SME elicitation is critical in selecting components from a single system. However, the uncertainty that accompanies SME elicitation is relatively unknown within the context of this process. Additionally, when the population of possible components expands to include many systems, the influence of SME rankings may be overshadowed by the consequence of a failure of a particular gate or the cost of conducting an OCA on those components.

The purpose of this thesis is to investigate the characteristics of the current component ranking method and evaluate the amount of uncertainty that accompanies SME rankings. This thesis studies SME ranking data from two separate watersheds: the Trinity River Basin in Texas, and the Willamette River Basin in Oregon. Fourteen SMEs provided rankings for the Trinity River and six SMEs provided rankings for the Willamette.

We utilize nonparametric methods to evaluate the uncertainty stemming from the rankings and to measure the resulting discrepancy in component selection from using an alternate ranking technique in which a SME does not rank the entire set of components. For example, an SME may assign ranks only to the top five components in a system with twenty components. In principle, this would allow a SME to concentrate effort on the most important components.

The following are the research questions that guide this thesis:

1. How can we quantify the level of agreement of an individual SME's rankings relative to their peers?

2. How can we quantify the variability of mitigated risk estimated by the OCA Prioritization Tool that is attributed to SME component rankings?

3. Are there more efficient SME elicitation techniques that can be used?

C. THESIS STRUCTURE

The thesis is organized in the following manner. Chapter II presents a review of the current literature associated with infrastructure resilience, SME elicitation, and construction of the model this thesis analyzes. Of particular interest are the techniques of SME elicitation and ranking of components using Borda Count, a technique of consolidating rank-based voting often used in elections. Chapter III explores the characteristics of the current model based upon the data provided by ERDC concerning two distinct watersheds. We also describe our approach to analyzing the limited data in our possession. First, we focus on the current approach followed by an analysis of utilizing truncated rankings to calculate the influence of specific components on a system. Chapter IV presents the results of our analysis by comparing the model outcomes using the methodology outlined in Chapter III. Finally, Chapter V presents our conclusions and recommendations for further research to assist in the development of the OCA Prioritization Tool.

II. LITERATURE REVIEW

This chapter summarizes some of the literature relevant to the topic of risk-based prioritization and methods used to mitigate risk. While the OCA prioritization model is somewhat unique in operation and does not possess abundance of literature, there are insights to be gleaned from previous research when focusing on specific features of this model.

A. INFRASTRUCTURE NETWORKS

Alderson et al. (2015) discuss how it is possible to evaluate the consequences of a loss of a particular component within a system by modeling infrastructure in a quantitative model. They emphasize that possessing a "validated mathematical model of system operation" provides insights about the consequences of a system component failure to operators (Alderson et al. 2015, sec. 6).

Jacob et al. (2011) discuss methods of estimating the probability of system failure given limited information about subsystems and components. This paper discusses how the probability of system failure can be estimated using the probability of failure of components with known failure values, estimated failure intervals, or expert estimations of failure probabilities. It also offers insights on how to approach the dependency of components when estimating the probability of system failure.

B. SME ELICITATION

Boring (2007) discusses the challenges of using expert opinion when operational data is unavailable. Boring states how expert judgements, or "the assignment of a rank, quantity, or probability to an event," can be skewed based upon the framework experts approach the task (Boring 2007). These differences in approach can lead to differences in scaling or varying expert bias. The paper laid a framework for mitigating these problems through expert training prior to elicitation.

Teter (2014) discusses techniques analysts can use when eliciting information from SMEs. Teter describes in detail several elicitation techniques and discusses how to tailor them to obtain valuable information. Some of the techniques of interest to this problem include the "rank ordered centroid assessment" and "Saaty's method of pairwise comparisons" (Teter 2014, sec. C-43).

Rank ordered centroid assessment requires SMEs to provide ranked data for all alternatives. Suppose n represents the number of alternatives and w_i is the weight for the i^{th} ranked item. Teter (2014) summarizes this method uses the following formula to transform ranked data into weights for each component as

$$w_i = \frac{1}{n} \sum_{j=i}^{n} \frac{1}{j} . \tag{1}$$

While the rank ordered centroid assessment requires just the ranked data, this limited information limits its effectiveness as the differences between ranks are not well defined (Teter 2014).

Saaty's method simplifies the ranking process for SMEs by requiring them to compare two alternatives and select the preferred one. Additionally, SMEs are asked to evaluate how much better the favored alternative is. Upon compilation of all the comparisons, the data is passed into a matrix and the resulting eigenvalues of that matrix represent the weights of the importance of each alternative. Teter assesses Saaty's method as being more mathematically rigorous but also notes it requires much more work from SMEs and may be less useful in the context of this thesis (Teter 2014).

Doğan and Giritligil (2014) attempt to characterize the probability of selecting the correct Borda Count winner when utilizing a truncated scoring rule. Using an Impartial Culture model to sample preferences, they run Monte Carlo simulations while varying numbers of voters, objects being voted on, and truncation level. Doğan and Giritligil suggest truncated ballots are less likely to return the actual Borda outcome. There is marginal gain as voters are asked to rank one additional alternative, but the effect diminishes as the truncation limit grows. This study is focused primarily on a single winner rather than achieving an accurate score vector of all alternatives.

Kilgour et al. (2022) investigate the weighted score elections, of which the Borda Count is most widely known. They use “large scale simulation to measure Condorcet efficiency, resilience to ballot truncation, and ability to maximize the Copeland score using both complete and truncated ballots” (Kilgour et al. 2022, p. 1). Methods described in this work can be used in framing the SME elicitation portion of the OCA tool.

C. OCA IMPLEMENTATION

Alt et al. (2021) develop the model analyzed in this thesis. The model integrates SME elicitation of component criticality and consequence evaluation of subsystems as an input to a mixed integer program which selects the components to inspect to maximize the risk mitigated for an entire watershed.

Alt et al. (2022) continue to develop the model introduced in their earlier publication. Here they evaluate the performance of the model against two separate watersheds facing a variety of weather conditions. They also discuss their efforts to make the model more generalizable and test it against watersheds with an increasing number of facilities and components.

THIS PAGE INTENTIONALLY LEFT BLANK

III. DATA AND METHODOLOGY

A. SME ELICITATION

Systems consisting of many components can be analyzed to determine the relative importance of each component and to measure negative effects of component failure on system operation. However, to effectively model the impact of component failure, some insight into the components relationship to the larger system is needed. Because there is a lack of data for component failures within dam gate systems, ERDC uses SMEs to provide an estimate of the relative importance of each component to create the basis for prioritization of OCAs (Alt et al. 2021).

There are several ranking methods ERDC explored for SME elicitation including having SMEs estimate the individual importance of each component. Eventually, ERDC elected to use the Borda Score for its ease of implementation and a reduced strain on SMEs as they consider the importance of each component relative to each other (Alt et al. 2022).

The Borda Score is a process for ranking a set of alternatives. Each voter assigns a rank 1 to *n*, with 1 being the highest rank, based upon the perceived ranking of a particular alternative. Upon completion of voting, the responses of all SMEs are compiled, and the summed ranks of each component are recorded. The Borda Score is then derived by subtracting summed ranks from a baseline value equal to the number of voters multiplied by the number of alternatives.

Table 1 demonstrates how rankings are consolidated to calculate the Borda Score for each component within a system.

Table 1. Example of a Borda Score Calculation

Rankings				Borda Results			
Component #	SME 1	SME 2	SME3	Baseline Score	Sum of Rankings	Borda Score	Relative Influence
1	3	2	3	9	8	1	0.11
2	1	1	2	9	4	5	0.56
3	2	3	1	9	6	3	0.33

ERDC scales the Borda Score for each component so that it is between 0 and 1, where larger values imply higher importance. This scaled Borda Score is used to represent the relative influence of each component (Alt et al. 2021). The sum of relative influence values for all components within a given system is equal to one. Alt et al. (2021) defines the formula for the relative influence of a component is

$$Component\ i\ Relative\ Influence = \frac{Component\ i\ Borda\ Score}{Sum\ of\ all\ Component\ Borda\ Scores}. \tag{2}$$

To maximize the amount of mitigated risk through selection of most influential components, the tool estimates the risk of each component as a function of the relative influence and the consequences of gate failure. Alt et al. (2021) expresses the component risk as

$$Component\ Risk = Relative\ Influence\ of\ Component \times Consequences\ of\ Gate\ Failure. \tag{3}$$

B. EXPLORATORY DATA ANALYSIS OF WATERSHED DATA

The data for this thesis comes from two separate watersheds: the Trinity River Basin in Texas and the Willamette River Basin in Oregon (Brown 2022). These two watersheds have a greater number of systems compared to those in initial case studies. In this section, we describe the exploratory data analysis of each system and highlight notable features that may influence the OCA Prioritization Tool.

1. Trinity River Basin

The portion of the Trinity River Basin analyzed in this section differs from the section analyzed by ERDC in their latest case study (Alt et al. 2022). The size of this portion is also larger than the Jennings Randolph Lake system and offers insight into tool behavior as the size of the network grows. Table 2 summarizes the major characteristics of the watershed including the number components in the watershed systems.

Table 2. Summary Statistics of Trinity River Basin. Adapted from Brown (2022).

Trinity River Basin	
Number of SME Rankings	14
Number of Unique Systems	8
Median Number of Components	18
Maximum Number of Components	23
Minimum Number of Components	14

Figure 2 provides an illustration of the variability of SME component rankings in the Navarro Mills Dam spillway gate system. Some components have generally consistent rankings while others have a large range of values. An example is Component 141 which ranges from second to least in importance. Additionally, other components ranked high in importance have outliers which rank them near the bottom. Another interesting observation of this set of boxplots is that the rankings generally increase as a function of the component id. This may be due to the assembler of the list annotating the features that seem most important to them first, or it may be due to SMEs giving higher ranks to components seen earlier in the component list.

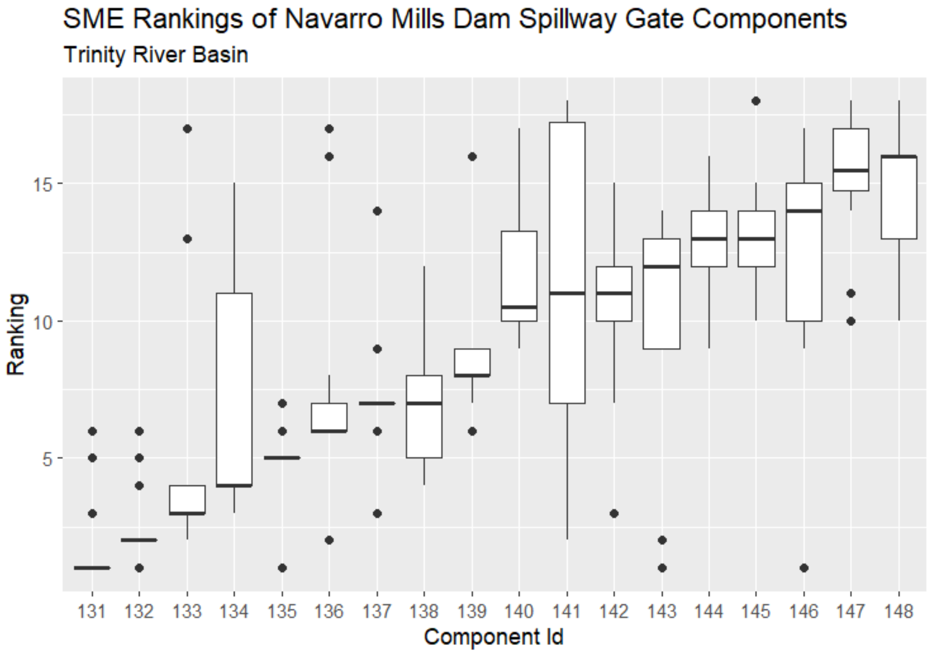


Figure 2. Example of SME Rank Disparity in Trinity River Basin. Adapted from Brown (2022).

Figure 3 depicts the standard deviations of component rankings in the Trinity River watershed. The standard deviation typically is around 3, but some components have large discrepancies in rankings leading to standard deviations as large as 8. Additionally, this distribution appears bimodal which may be the result of the rankings of a subset of components being more variable than the complement of that subset.

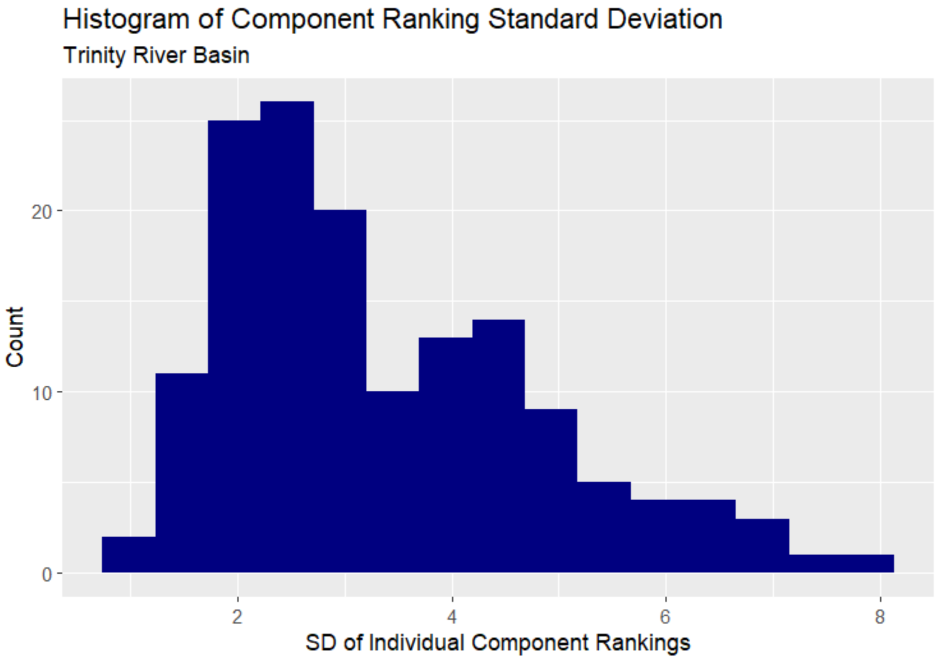


Figure 3. Standard Deviation of Component Rankings in Trinity River Basin. Adapted from Brown (2022).

2. Willamette River Basin

The Willamette River Basin is the most recent watershed ERDC evaluated. The large number of systems and components of this watershed offers valuable insights into tool behavior in a more complex watershed. Table 3 provides summary statistics of the systems in the Willamette Basin. Most noticeable is the large range of number of components in the watershed gate systems. The system with the most components has twice as many components as the system with the fewest.

Table 3. Willamette River Basin Summary Statistics. Adapted from Brown (2022).

Willamette River Basin	
Number of SME Rankings	6
Number of Unique Systems	19
Median Number of Components	17
Maximum Number of Components	22
Minimum Number of Components	11

Figure 4 illustrates the rankings submitted for the components of the Lookout Point Spillway Gates. Most noticeable in this visualization are Components 317, 318, and 319 which received a large range of rankings, demonstrating the uncertainty of their influence. It is also noteworthy that Components 307 and 308 were almost unanimously selected as the most important and Component 327 was generally accepted as least influential. The linear trend in rankings seen in the Trinity River Basin (Figure 2) is seen here as well. Although not addressed in this thesis, further investigation into whether alternatives presented first receive higher rankings may be worthwhile. The rankings for this system seem to be more consistent than the rankings for the Navarro Mills Dam in Section 1. This may be due to the smaller sample of SMEs or additional training implemented into the elicitation phase which may reduce variance.

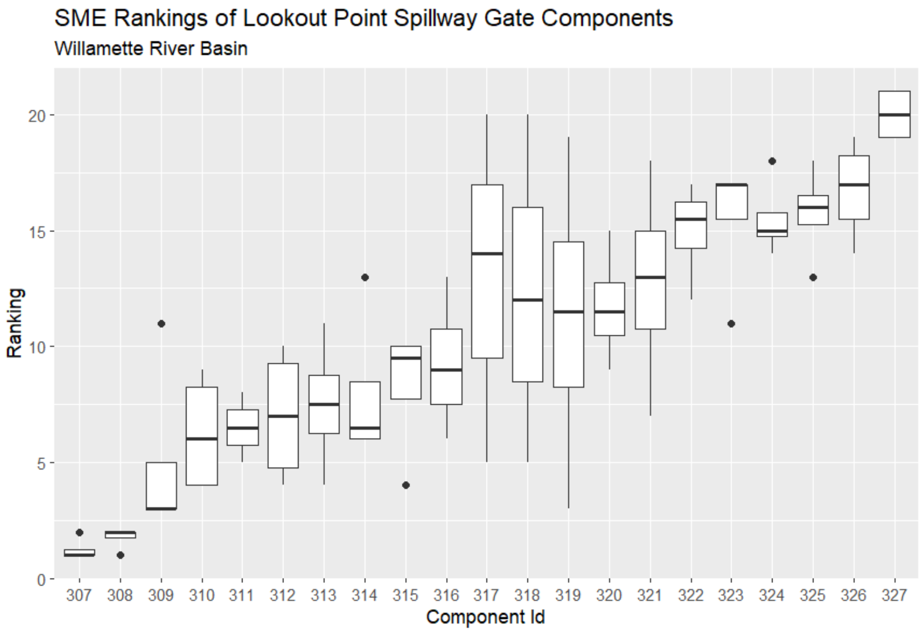


Figure 4. Example of SME Rank Disparity in Willamette River Basin. Adapted from Brown (2022).

Figure 5 depicts the standard deviations of component rankings across the Willamette watershed. The peak standard deviation is approximately 3, however standard deviation values can approach 9 demonstrating potentially large disagreement over the influence of a specific component. It is also important to note that these standard deviations are also likely influenced by the number of components within a gate system. A system with more components will likely have a larger deviation in component influence.

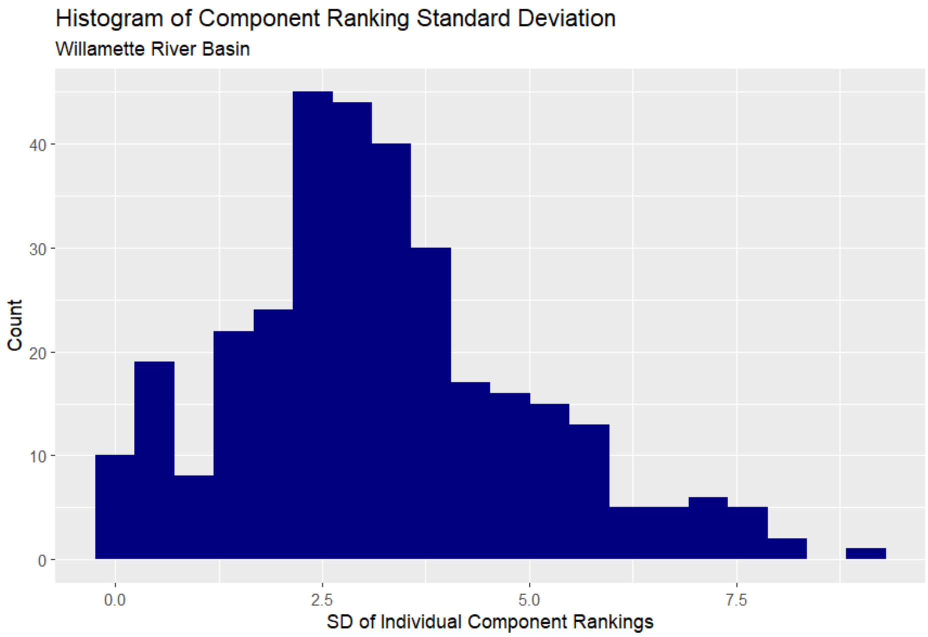


Figure 5. Standard Deviations of Component Rankings in Willamette River Basin. Adapted from Brown (2022).

C. MEASURING SME RANKING PROXIMITY TO PEER AVERAGE

In eliciting data from SMEs, it may be useful to compare each SME's rankings to the others. By doing so, analysts can identify SMEs that stand out as different from the consensus and subsequently seek to identify the reasons for variation in rankings, such as a misunderstanding, or an individual SME having greater knowledge of the site and built-in redundancies that reduce the influence of a specific component. These discrepancies are worth investigating as the insights gained could change the perspective of the group of SMEs or a mistake in rankings can be identified and altered to not degrade the accuracy of the OCA Prioritization Tool.

1. Measuring SME Ranking Distance Methodology

We use a simple method to analyze the distance between SME rankings. Some SMEs may have more knowledge of certain systems than others, therefore, this calculation is completed for each unique system. Additionally, differences in rankings for components at the top of the consensus rankings are likely more important than those at the bottom. To account for this, we calculate a weight for each component based upon its average ranking. The next step is to find the absolute difference between the individual SMEs ranking and the consensus ranking and multiply it by the weight for that component. The error for each component is then summed across the entire system.

Suppose M represents the number of components in a system, c_{ij} represents the rank of component i for SME j, and $\bar{c}_i$ represents the average ranking of component i among all SMEs. We calculate weight of each component i, represented by w_i, and the distance of the rankings of SME j from the average, represented as d_j, with the following two formulas:

$$d_j = \sum_{i=1}^{M} \left| c_{ij} - \bar{c}_i \right| \times w_i , \tag{4}$$

and

$$w_i = \frac{(M + 1 - \bar{c}_i)}{M} . \tag{5}$$

This final ranking provides a measure of distance for how far a SME's rankings are from the consensus. Larger values indicate further distance from the consensus while smaller values indicate the rankings are closer. The values produced vary from system to system and are primarily dependent on the number of components that a system has.

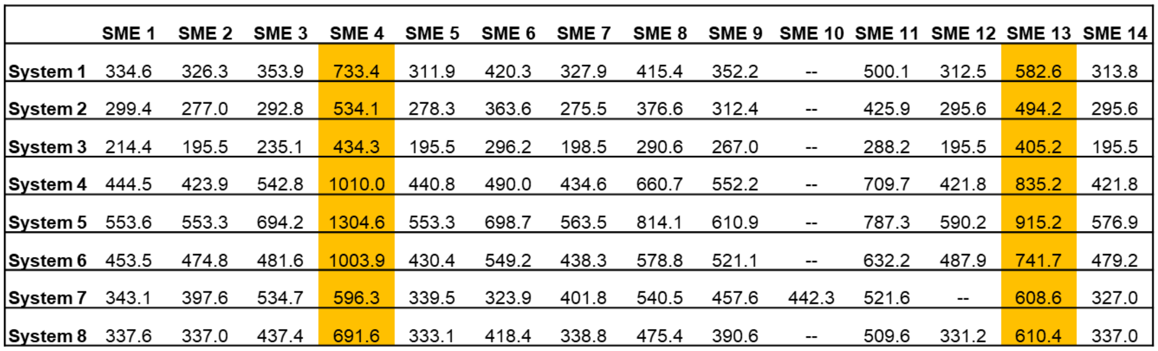
By examining these values at a system level, analysts can identify if a SME has larger values across the entire basin or if the divergence is in a specific system. If the SME has large discrepancies across all systems, the SME may require additional training. However, if there is a large discrepancy in only one or a handful of systems, that SME may possess some knowledge of the system that is not commonly known. Ultimately, analysts can use these distance values to assess SME rankings and investigate the reasoning why larger divergences in rankings occur.

2. SME Ranking Distance Results

a. Trinity River Basin Distance Results

Analysis of the Trinity River SME ranking distance reveals a couple instances of SME rankings that produce large deviance from the consensus rankings. Table 4 shows SME 4 and 13 submitted rankings that are routinely larger than the consensus ranking. Their error value seems to be almost double that of the system average.

Table 4. Trinity River Basin SME Ranking Distance. Adapted from Brown (2022).



	SME 1	SME 2	SME 3	SME 4	SME 5	SME 6	SME 7	SME 8	SME 9	SME 10	SME 11	SME 12	SME 13	SME 14
System 1	334.6	326.3	353.9	733.4	311.9	420.3	327.9	415.4	352.2	--	500.1	312.5	582.6	313.8
System 2	299.4	277.0	292.8	534.1	278.3	363.6	275.5	376.6	312.4	--	425.9	295.6	494.2	295.6
System 3	214.4	195.5	235.1	434.3	195.5	296.2	198.5	290.6	267.0	--	288.2	195.5	405.2	195.5
System 4	444.5	423.9	542.8	1010.0	440.8	490.0	434.6	660.7	552.2	--	709.7	421.8	835.2	421.8
System 5	553.6	553.3	694.2	1304.6	553.3	698.7	563.5	814.1	610.9	--	787.3	590.2	915.2	576.9
System 6	453.5	474.8	481.6	1003.9	430.4	549.2	438.3	578.8	521.1	--	632.2	487.9	741.7	479.2
System 7	343.1	397.6	534.7	596.3	339.5	323.9	401.8	540.5	457.6	442.3	521.6	--	608.6	327.0
System 8	337.6	337.0	437.4	691.6	333.1	418.4	338.8	475.4	390.6	--	509.6	331.2	610.4	337.0

b. Willamette River Basin Distance Results

The Willamette River basin offers some different results when analyzing the SME ranking distance from the consensus. Table 5 displays the resulting SME distance values from the Willamette River Basin. SME 4 rated only three systems and each rating value was higher than the rest of the SMEs. This may indicate that SME 4 requires some retraining or is thinking about the problem in a different manner than the rest of the SMEs.

Additionally, multiple SMEs have one or two systems where they deviate from the consensus rankings. In these cases, the discrepancy is limited to one or two systems so the problem may not correspond with SME elicitation training. It is still worthwhile to investigate the reasons for the difference if possible.

Table 5. Willamette River Basin SME Ranking Distance. Adapted from Brown (2022).

	SME 1	**SME 2**	**SME 3**	**SME 4**	**SME 5**	**SME 6**
System 1	141.8	115.2	124.9	--	143.9	128.9
System 2	125.7	92.3	79.4	--	110.3	95.7
System 3	70.8	90.2	75.2	--	111.9	102.6
System 4	--	164.5	149.6	--	202.2	167.6
System 5	--	63.2	46.2	--	94.7	75.3
System 6	--	33.9	26.5	--	49.6	37.5
System 7	--	47.7	46.1	--	73.2	62.6
System 8	--	88.8	77.0	145.6	88.5	104.1
System 9	149.3	105.8	105.2	--	152.0	217.5
System 10	--	43.3	36.1	--	63.4	52.8
System 11	211.0	158.1	132.2	--	223.5	186.3
System 12	--	88.4	77.0	--	112.1	111.3
System 13	--	119.1	111.8	--	140.1	165.4
System 14	--	60.0	47.8	--	89.2	65.8
System 15	--	41.8	34.8	--	51.0	42.0
System 16	--	140.2	137.2	234.0	166.7	196.5
System 17	--	164.5	159.1	247.5	175.8	186.4
System 18	--	103.0	80.5	--	156.2	126.4
System 19	--	102.3	77.0	--	144.7	117.0

D. STATISTICAL ANALYSIS METHODS

This section describes the statistical analysis methods used in this thesis to gain additional insights into the behavior of the OCA Prioritization Tool. The rankings for each system are manipulated in a couple of ways. The first way is using a nonparametric bootstrap to produce a “new” sample collection of rankings. The second method of data perturbation was instituting a limit on SME rankings. We measure the effects of these perturbations in both mitigated risk and average change in component selection rates.

1. Nonparametric Bootstrap

The nonparametric bootstrap is a resampling technique that allows analysts to observe characteristics of a population from a random sample that represents the population well (Sprent and Smeeton 2007, p. 445). By using the bootstrap method, analysts do not have to make assumptions about the distribution of the data.

For this analysis, we use the bootstrap to resample the entire sets of rankings of the SMEs. By running the bootstrap multiple times, we analyze how different sets of SME rankings can affect the variability of the mitigated risk value produced by the OCA Prioritization Tool.

2. SME Ranking Truncation

ERDC's method of SME elicitation requires SMEs to rank the importance of all components for each system. This method gathers a lot of data in a relatively efficient manner, but still requires intense focus on the part of SMEs as they work through multiple iterations of ranking many components. By truncating the rankings SMEs can focus on ranking the most important components without making possibly arbitrary decisions about components of lesser importance.

We consider the effects of limiting the number of components that a SME is asked to rank. Using this method, SMEs rank components up to a set limit, the remaining components receive a rank value equal to the average of the remaining possible ranks. To apply this rule to SME $j's$ full rankings of components, suppose our data contains m components, r represents the limit to which SMEs rank components, and c_{ij} represents SME j*'s* ranking of component i. We evaluate SME rank data with the formula summarized as

$$\tilde{c}_{ij} = \begin{cases} c_{ij} & if \;\; c_{ij} \leq r \\ \dfrac{r+1+m}{2} & if \;\; c_{ij} > r \end{cases} . \tag{6}$$

E. EVALUATING UNCERTAINTY IN OCA PRIORITIZATION TOOL'S OUTPUT

To evaluate the uncertainty derived from SME rankings, we examine the outputs of the OCA Prioritization Tool in the form of components selected for OCAs and mitigated risk in dollars.

1. Mitigated Risk

The objective function of the OCA Prioritization Tool is to maximize the mitigated risk through the selection of a set of components constrained by a budget. Therefore, mitigated risk is a natural way to measure the uncertainty derived from SME rankings as you receive different values based upon the set of SME rankings that is used. Each iteration of bootstrapped SME rankings produces a new value for mitigated risk. These values can be compiled and analyzed with standard statistical methods.

2. Component Selection Rate

The second output from the OCA Prioritization Tool is the components selected to receive an OCA. Each iteration of the bootstrap can produce a different combination of components selected. We capture the data for these selections and measure the proportion of times each component is selected. After collecting this data for the full ranking bootstrap, we can use the same process for measuring component selection in truncated rankings and subsequently compare the truncated-ranking output with full-ranking output.

We measure the difference in these proportions using the Mean Absolute Error (MAE) formula. Suppose m represents the number of components in a watershed, p_i represents the proportion of time component i is selected under a truncated ranking system, and $\bar{p}_i$ represents the proportion of time component i is selected under a full ranking system. The mean absolute difference of proportion of time a component is selected is represented by

$$MAE\ of\ \ component\ selection\ proportions = \frac{1}{m}\sum_{i=1}^{m}\left|p_i - \bar{p}_i\right|. \tag{7}$$

IV. RESULTS AND ANALYSIS

This chapter analyzes the results of the OCA Prioritization Tool tested with the methods described in the previous chapter. We quantify uncertainty stemming from SME elicitation by bootstrapping SME rankings. We then assess how limiting the number of components SMEs rank affects the tool's estimated mitigated risk and component selection.

Although we use the tool to analyze the outputs based upon perturbations in ranking data, we lack the true consequence values for the failure of each specific gate. To mitigate this issue, we place notional values as placeholders for the consequences to analyze how the model reacts to different rankings. The consequence values we use remain constant throughout our analysis. This means that our observations relate to the behavior of the model and are not representative of the results that the two watersheds would produce if we had the actual consequence data from the Watershed Assessment Tool.

A. VARIABILITY IN FULL RANK OUTPUTS

To quantify the uncertainty originating from SME elicitation in the OCA Prioritization Tool, we ran 1000 iterations of the nonparametric bootstrap of SME rankings. Each iteration provided a new collection of SME rankings that was then used to select components for OCAs, and to estimate the resulting mitigated risk. We also varied the budget, equivalent to the proportion of total number of components within a watershed, to assess if budget levels affect the variability of the outcome as well. The proportions we use as budget levels are 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, and 0.7.

1. Trinity River Basin

The uncertainty in estimated mitigated risk is depicted in Figure 6, which shows a marked dependence on budget levels. These plots depict the distribution of mitigated risk using 1000 bootstrapped iterations. The line in the center of the boxplot represents the median, the edges of the box contain the middle 50 percent of bootstrapped values, and an 80 percent confidence band for the mean of the mitigated cost lies between the two red lines. As budget levels increase, the mitigated risk increases. This is representative of the components with greater risk being selected first and components with lesser risk being selected as the budget allows. The variability of components appears to be narrower at each end of the budget spectrum and larger in the center.

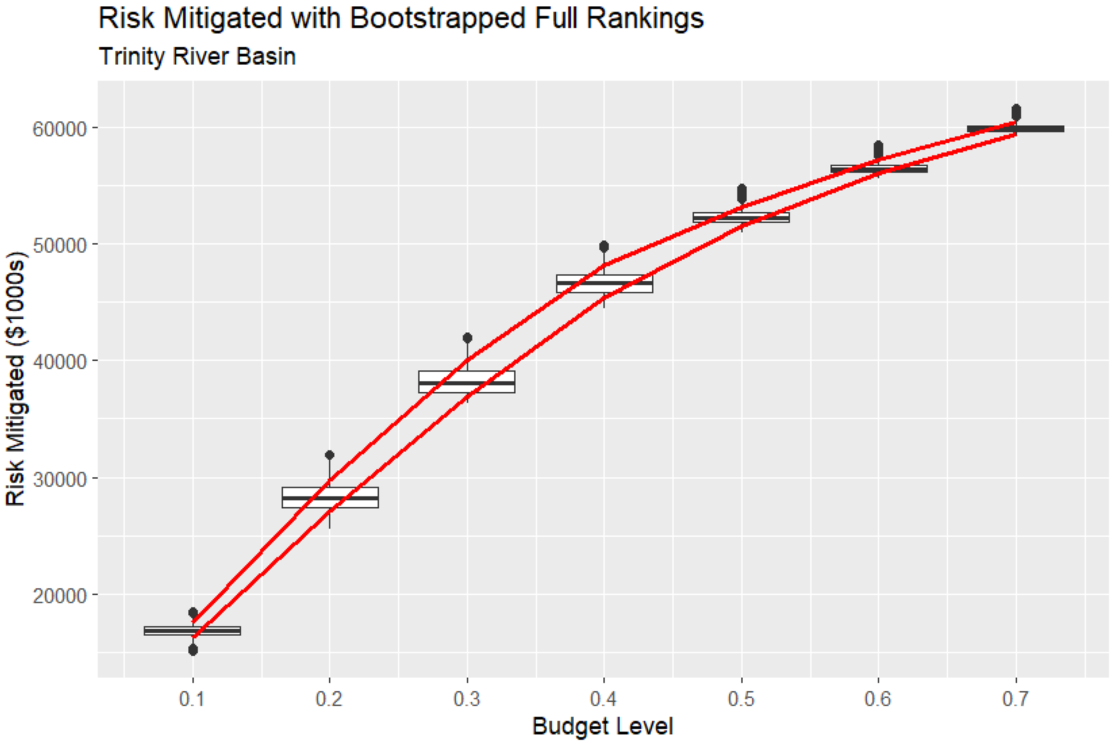


Figure 6. Distribution of Mitigated Risk after 1000 Iterations of a Fully Ranked Bootstrap, Trinity River Basin. Adapted from Brown (2022).

To better display the variability of mitigated risk in Figure 6, we center the data at each budget level on the sample mean and plot the corresponding results. Figure 7 illustrates the results of this transformation. Here we can identify a variability maximum at the 0.20 budget level followed by a clear reduction in variability as more components are selected for an OCA. This may be attributed to general agreement amongst SMEs on the most important components and the diminishing level of component risk among lower ranked components.

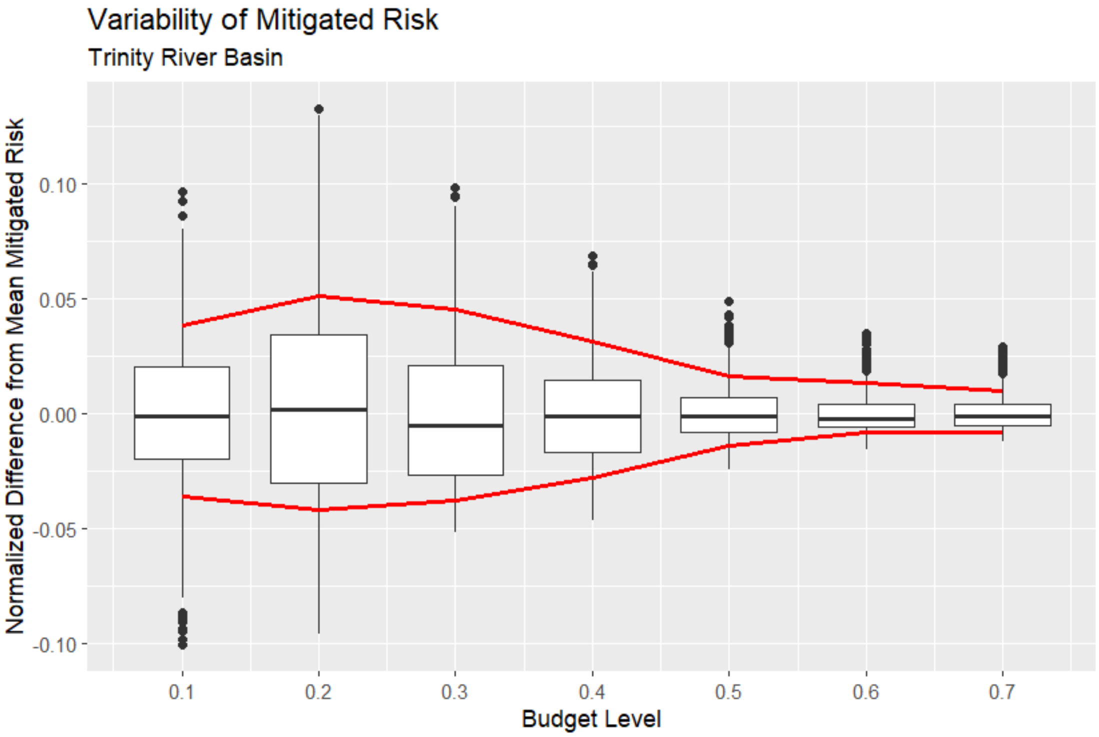


Figure 7. Variability of Mitigated Risk Resulting from 1000 Iterations of a Fully Ranked Bootstrap, Trinity River Basin. Adapted from Brown (2022).

2. Willamette River Basin

In this section, we apply the same analytic techniques used in the Trinity River Basin to the data collected for the Willamette River Basin. The Willamette River Basin contains 327 components, more than twice as many as the number of components in the Trinity River Basin. Additionally, only six SMEs ranked components within this basin.

Figure 8 depicts the distribution of mitigated risk across 1000 bootstrapped samples. Interestingly, there appears to be greater variance at the ends of the budget spectrum and lesser variance in the middle. It is a different shape than the variability we observe in Trinity River Basin's mitigated risk distribution plot (Figure 6). This suggests that the variability of mitigated risks in different watersheds may exhibit different patterns.

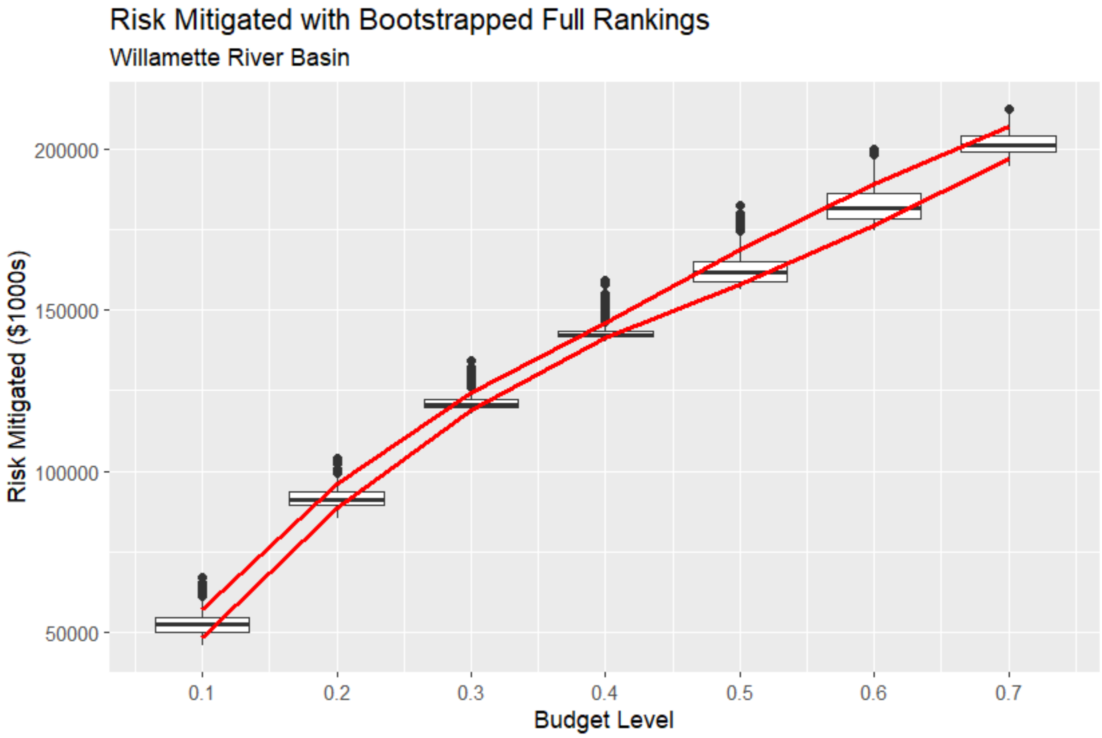


Figure 8. Distribution of Mitigated Risk after 1000 Iterations of a Fully Ranked Bootstrap, Willamette River Basin. Figure created with data from Brown (2022).

Figure 9 centers the variability of each budget level on the sample mean. There is a noticeable reduction in variability at the 0.40 budget level. It is possible that this is the budget level that selects the entire set of components that the SMEs have ranked relatively high and are associated with high levels of risk. Below the set of components estimated to present the most risk, there could be another set of components with moderate risk, but the budget does not allow for selection of these components until the budget becomes much larger.

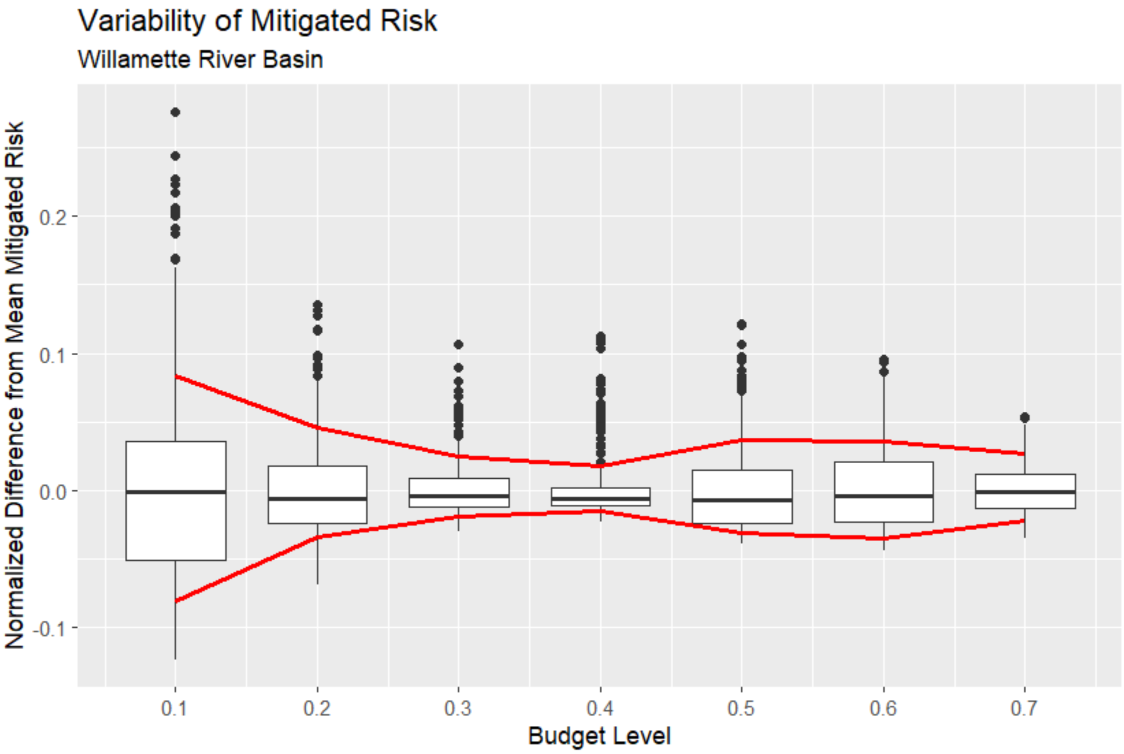


Figure 9. Variability of Mitigated Risk Resulting from 1000 Iterations of a Fully Ranked Bootstrap, Willamette River Basin. Adapted from Brown (2022).

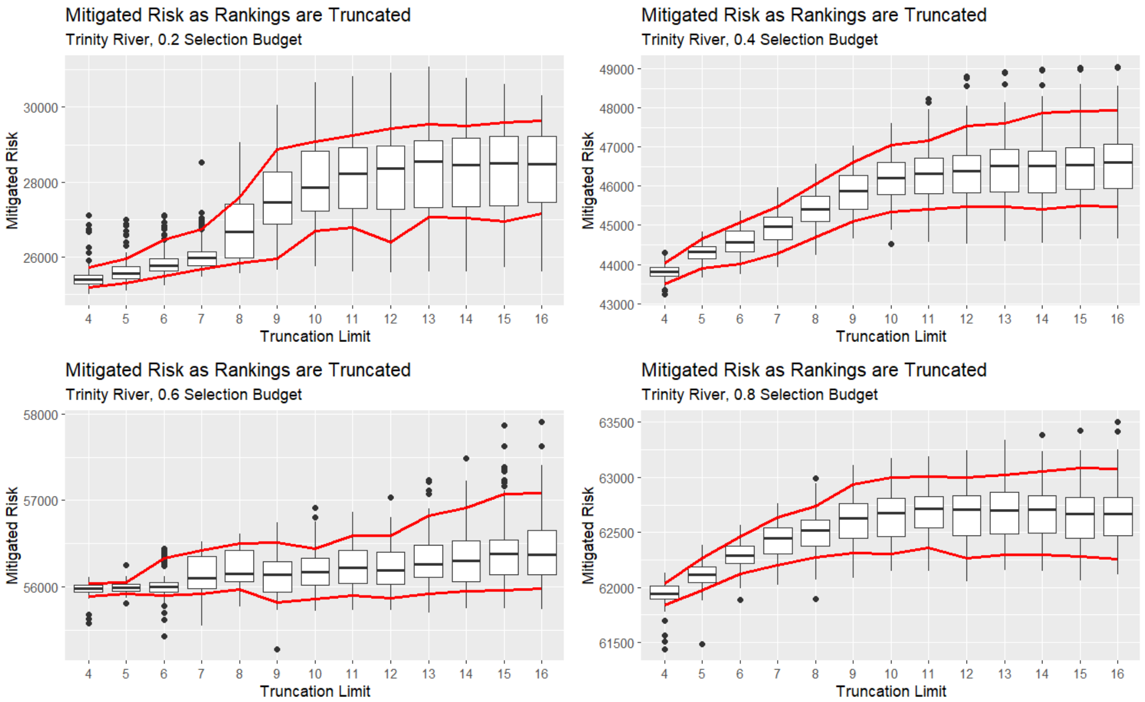
B. EFFECT OF TRUNCATING SME RANKINGS

To measure the effect of limiting the number of components that a SME can rank, we apply Formula 6, described in Chapter III, to simulate SME rankings that rank a limited number of components. We consider this effect for each number of ranked components between four and sixteen. At each level, we use the nonparametric bootstrap with 100 iterations to estimate the variability in mitigated risk. We select four budget levels at which we execute the bootstrap across all truncation levels.

1. Trinity River Basin

Figure 10 depicts risk mitigation across truncation levels for four assessment budgets, relating to inspection proportions of .20, .40, .60, and .80 of the total number of watershed components. At each budget level, the mean estimated mitigated risk increases as the ranking limit increases. In all cases, however, there is a point at which the gain in estimated mitigated risk levels off and further gains are marginal.

The mean mitigated risk and variability both increase as the truncation limit increases. At the 0.60 selection level, however, the mean estimated mitigated risk begins to level off. This may be due to all components with significant weights being selected at the 0.60 budget level and the remainder of the components sharing a diminishing influence.



Mitigated Risk as Rankings are Truncated
Trinity River, 0.2 Selection Budget

Mitigated Risk as Rankings are Truncated
Trinity River, 0.4 Selection Budget

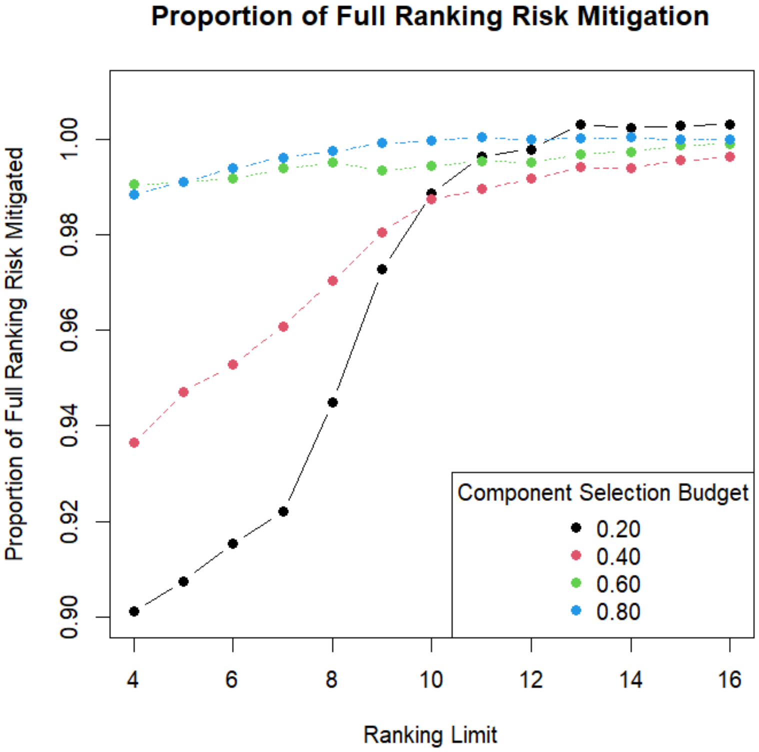
Mitigated Risk as Rankings are Truncated
Trinity River, 0.6 Selection Budget

Mitigated Risk as Rankings are Truncated
Trinity River, 0.8 Selection Budget

The red bounds depicted on these graphs illustrate the 80% confidence band of mitigated risk across all truncation limits.

Figure 10. Estimated Mitigated Risk across Truncation Levels, Trinity River Basin. Adapted from Brown (2022).

It is important to note that the plots shown in Figure 10 have different vertical scales. Figure 11 illustrates how these estimates of mitigated risk compare to the estimates of untruncated rankings. With low selection budgets, there is a substantial gain in mitigated risk as you increase the number of components ranked. This is because, at low selection budgets, the separation at the upper end of the rankings matters more, as that is where the selection cutoff resides. With larger budgets, truncating rankings have a smaller effect as most components are already selected and the components that are not selected account for a smaller share of the risk for a gate. As the truncation limit approaches ten across all budget levels, the estimated mitigated risk is greater than 98 percent of the untruncated mitigated risk and additional rankings provide marginal gain to that value.



This plot compares the mean estimated mitigated risk produced by truncated rankings to the mean estimated mitigated risk produced by full rankings. The x-axis describes the number of components SMEs ranked.

Figure 11. Scaled Estimated Mitigated Risk across Truncation Levels, Trinity River Basin. Adapted from Brown (2022).

We also estimate the amount of mitigated risk from selected components with truncated rankings. Figure 12 depicts the proportion of mitigated risk influenced by truncated rankings. At all budget levels, the proportion of mitigated risk derived from truncated rankings falls as additional rankings are included. The break point between the two categories of mitigated risk appears to be concave downward. This demonstrates that as additional rankings are included, the amount of mitigated risk reliant on truncated rankings diminishes. However, the increase in untruncated mitigated risk gets smaller each step as the focus shifts to less influential components.

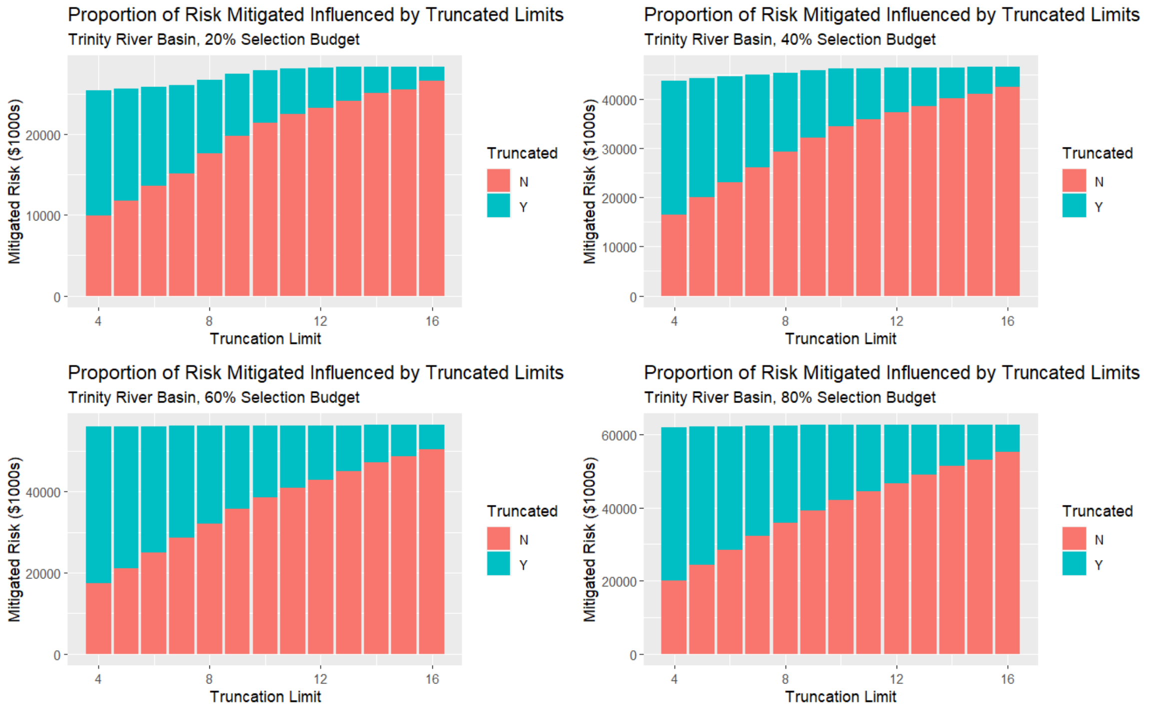
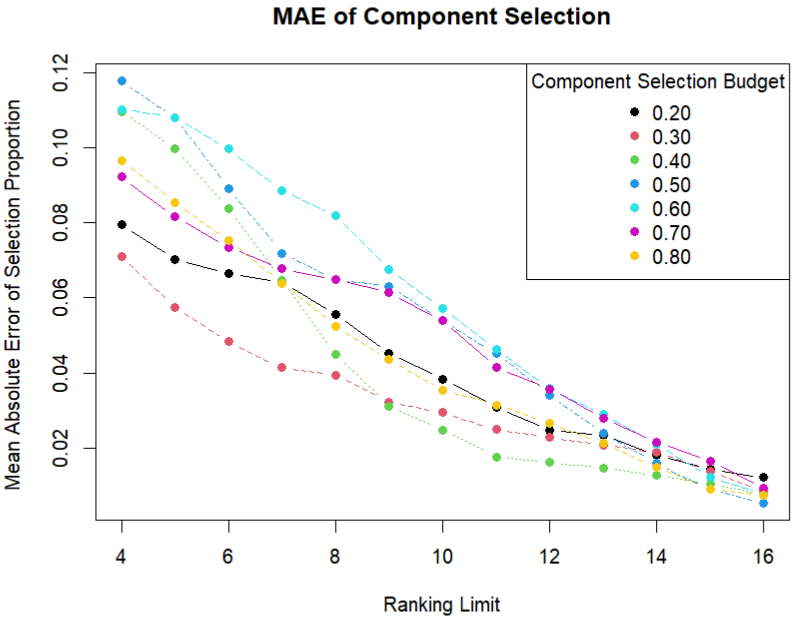


Figure 12. Proportion of Estimated Mitigated Risk Influenced by Truncation, Trinity River Basin. Adapted from Brown (2022).

Truncation of SME rankings also affects the components selected for OCAs. Using the method expressed by Formula 7, presented in Chapter III, we measure the difference in component selection rates under truncation of rankings ranging from four to sixteen relative to sets of rankings where all components are ranked. Figure 13 depicts these values across a range of selection budgets. In all cases, there is a consistent negative trend in the mean difference in selection rate as the number of components ranked increases. For the 0.40 selection budget, there is a large initial difference in selection rates followed by a steep decrease in error between the limits of six to nine before flattening out again. This likely indicates that at this budget level, including additional rankings removes important components from the unranked pool leading to a sharp decline in selection rate error. If the budget is too tight, the mid-tier components likely would not be selected meaning relegation to the unranked subset would not have a large effect. If the budget is sufficiently high, they may be picked regardless of ranking truncation.



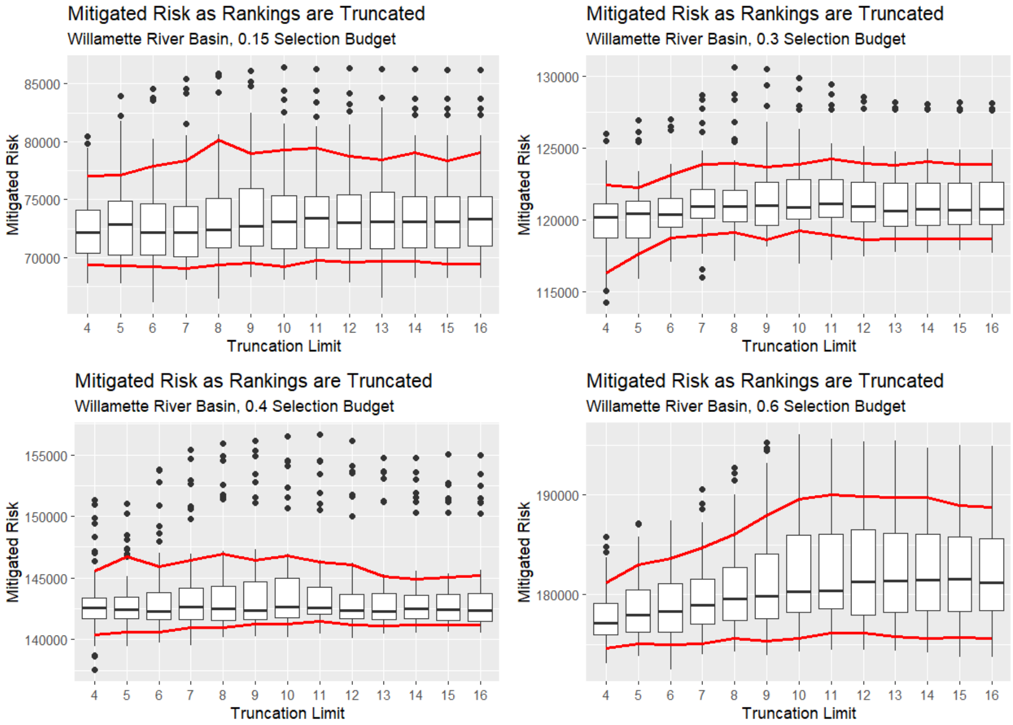
This figure compares the proportion of time a component is selected with truncated rankings to the base case of proportion of time a component is selected with full rankings. The x-axis describes the number of components SMEs ranked.

Figure 13. Mean Absolute Error of Selection Proportion across Ranking Limits, Trinity River Basin. Adapted from Brown (2022).

2. Willamette River Basin

The Willamette River Basin uses a smaller set of six SMEs compared to the fourteen SMEs the Trinity River Basin uses. Two of the six SMEs provide rankings for only a portion of the systems leading to additional sparsity in the data. This difference in available data contributes to some difficulties in conducting analyses on truncated rankings for the Willamette River Basin. During the bootstrap runs, the algorithm failed to find the optimal solution in a timely manner. For example, at the 0.45 budget level and truncation limit of four, the algorithm could not select the optimal set of components in a timely manner. Typical bootstrap iterations took roughly five to ten seconds to run, but at this level, certain iterations took four hours without providing a final selection. This behavior occurs more frequently at budget levels greater than 0.70 when impacted with high truncation. This demonstrates the challenges in implementing the nonparametric bootstrap method of evaluating truncation levels when SME ranking data is sparse.

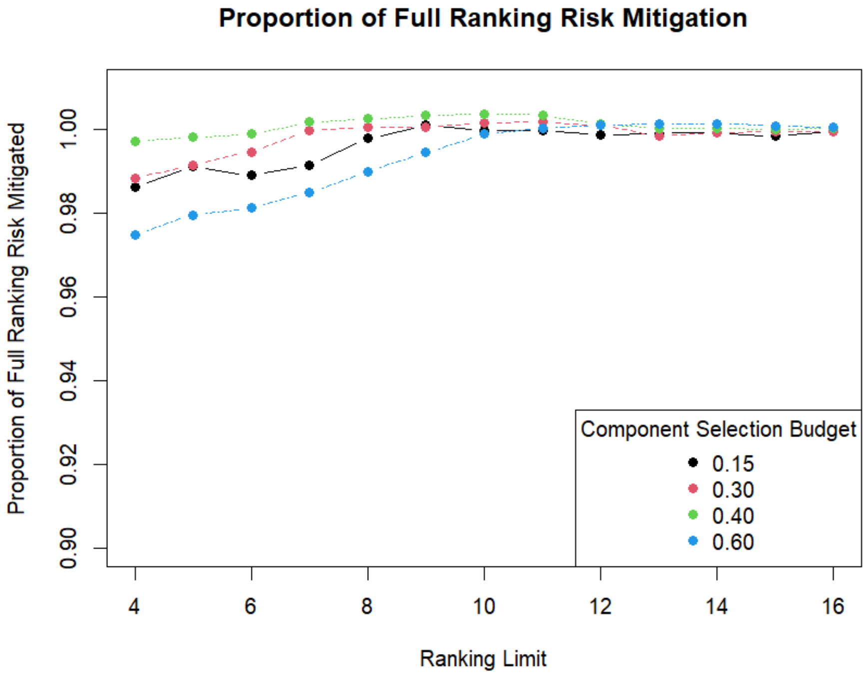
Figure 14 depicts the estimated mitigated risk in the Willamette River Basin across multiple budget levels. A notable observation is that the trend remains relatively flat across truncation levels. The shape of this trend is substantially different than the trend observed in the Trinity River Basin (Figure 10) which presented a more pronounced curve. This suggests that in the Willamette River Basin, truncation of rankings may not have a large effect on mitigated risk.



The red bounds depicted on these graphs illustrate the 80% confidence band of mitigated risk across all truncation limits.

Figure 14. Estimated Risk across Truncation Levels, Willamette River Basin. Adapted from Brown (2022).

Figure 15 supports our observation that the mitigated risk remains relatively constant across all truncation levels. Across all budget levels, the lowest average mitigated risk is roughly 97 percent. As the truncation limit increases to 10, all budget levels mitigate risk at approximately 100 percent of the mitigated risk observed in the full ranking set.



This plot compares the mean estimated mitigated risk produced by truncated rankings to the mean estimated mitigated risk produced by full rankings. The x-axis describes the number of components SMEs ranked.

Figure 15. Scaled Estimated Mitigated Risk across Truncation Levels, Willamette River Basin. Adapted from Brown (2022).

Figure 16 shows how much of the mitigated risk is influenced by truncated rankings across the Willamette watershed. When compared to the Trinity River Basin, the amount of mitigated risk derived from truncation appears smaller across all budget levels. This may be due to the increase in the number of gate systems in the Willamette. By having the option to select the most influential components in additional systems, the model does not have to select among less influential components that are more likely to have truncated rankings.

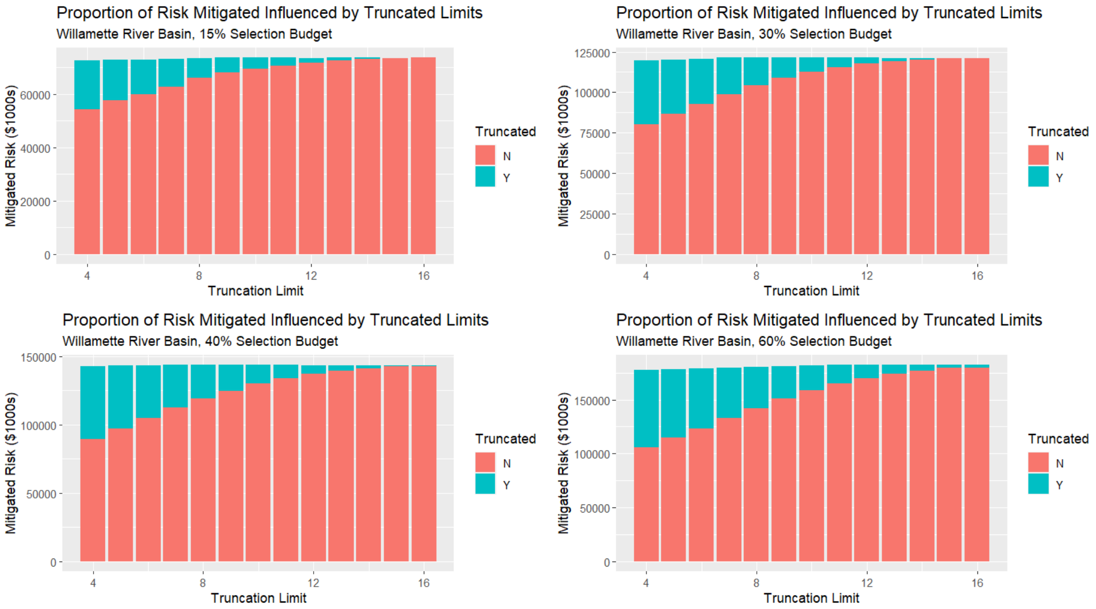
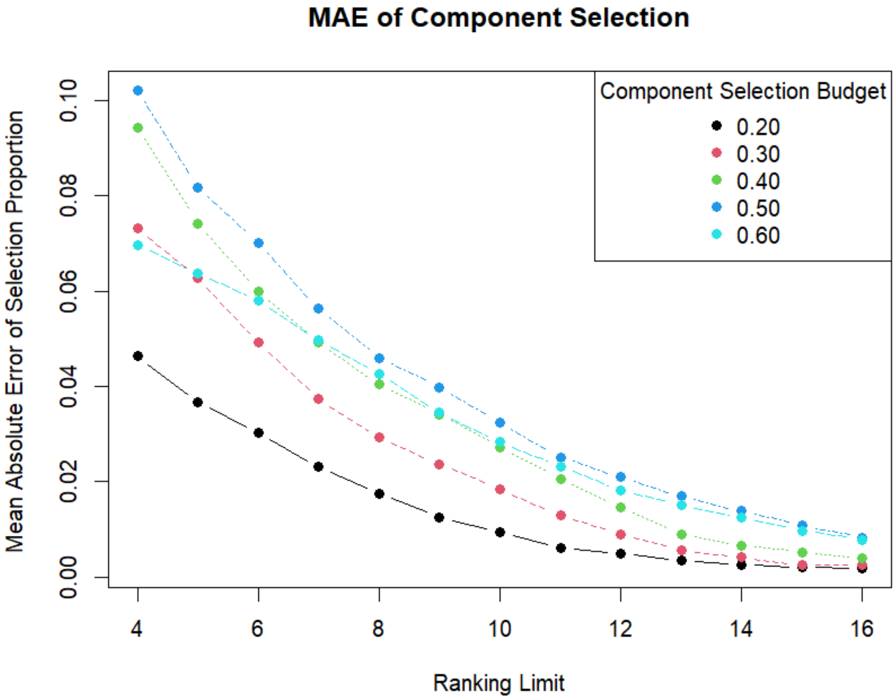


Figure 16. Proportion of Estimated Mitigated Risk Influenced by Truncation, Willamette River Basin. Adapted from Brown (2022).

Figure 17 illustrates how much the set of selected components resulting from truncated rankings differs from the set of selected components derived from the full rankings. Almost all budget levels display a similar curve showing the improvement in selection error becomes marginal as the number of components ranked by SMEs increases.

The 0.60 budget level is interesting as it initially performs better than all budget levels except for 0.20. As the number of ranked components is increased, the reduction in difference in selection rate between truncated and full ranks is a much shallower slope than the curve displayed by the other budget levels. It is possible that if all components share the same risk value, the algorithm selects components from the truncated component pile in an order that, by happenstance, matches the order they are actually ranked in the base case leading to similar results as the base case.



This figure compares the proportion of time a component is selected with truncated rankings to the base case of proportion of time a component is selected with full rankings. The x-axis describes the number of components SMEs ranked.

Figure 17. Mean Absolute Error of Selection Proportion across Ranking Limits, Willamette River Basin. Adapted from Brown (2022).

THIS PAGE INTENTIONALLY LEFT BLANK

V. CONCLUSION

In this thesis, we examine the amount of uncertainty existing in risk estimates derived from SME elicitation. By using the nonparametric bootstrap to resample a set of SME rankings, we assess how much variation exists in the estimated mitigated risk and component selection rates produced by the OCA Prioritization Tool.

A. RESEARCH QUESTIONS

We state the objectives of this thesis in Chapter I in the form of three research questions. These research questions guide our analysis of the SME ranking data of two distinct watersheds provided by ERDC.

(1) How can we quantify the level of agreement of an individual SME's rankings relative to their peers?

An important aspect of analyzing SME rankings is identifying SMEs that produce rankings substantially different than their peers. These differences are not necessarily good or bad, but they allow analysts to identify if a SME needs retraining on the elicitation system or if they possess some knowledge the rest of the group does not have. The data from both the Trinity and Willamette watersheds illustrate some differences in the SME ranking data that deserve some investigation.

(2) How can we quantify the variability of mitigated risk estimated by the OCA Prioritization Tool that is attributed to SME component rankings?

We demonstrate the variability of mitigated risk in the Trinity and Willamette watersheds can be quantified with the use of the nonparametric bootstrap. Variability is not generalizable across all watersheds, but is likely dependent on the SME rankings, structure of the watershed, and consequence data associated with gate systems in the watershed. As an example, in the Trinity River Basin, the maximum variability is observed at the 0.20 budget level but steadily decreases with a higher budget level. In the Willamette River Basin, the variability starts large, decreases through the intermediate budget levels, and increases again at higher budget levels.

(3) Are there more efficient SME elicitation techniques that can be used?

We consider the effects of truncating the number of components in each system that SMEs are asked to rank. In both the Trinity and Willamette watersheds, rankings truncated at ten components mitigate about 98 percent of the risk compared to using full rankings. This shows a tradeoff that may be worth investigating in other watersheds.

B. RECOMMENDATIONS

This thesis examines the uncertainty derived from SME elicitation under one set of environmental conditions, namely, the 100-year flood. By obtaining true consequence data from the Watershed Assessment Tool that show the consequences under a variety of environmental conditions, more realistic insights can be gained from the behavior of the OCA Prioritization Tool in the Trinity and Willamette River Basins.

One noteworthy observation from the ranking data is that the ranks of components within a system appear to be correlated with the order in which they occur on the list. Further investigation should be conducted on potential bias introduced by SMEs being presented the same list of components.

Another potential area of research is the behavior of the model over multiple time periods. This thesis analyzes the results of one time period but does not investigate the impact of ranking truncation in subsequent time periods where the model will likely be presented with components not selected in the first time period.

LIST OF REFERENCES

Alderson DL, Brown GG, Carlyle WM (2015) Operational models of infrastructure resilience. *Risk Analysis* 35(4), http://dx.doi.org/10.1111/risa.12333.

Alt JK, Brown WH, Richards JP, Gallarno GE, Olszewski JM, Rice TL (2022) Risk-informed prioritization of Operational Condition Assessments. Report ERDC/ITL SR-2 2-5, The US Army Engineer Research and Development Center, Vicksburg, MS, https://apps.dtic.mil/sti/pdfs/AD1186404.pdf.

Alt JK, Richards JP, Gallarno GE, Brown WH (2021) Risk-informed prioritization of Operational Condition Assessments. *IIE Annual Conference. Proceedings* (Norcross, GA), https://www.proquest.com/scholarly-journals/risk-informed-prioritization-operational/docview/2560888656/se-2.

Boring, RL (2007) A review of expertise and judgment processes for risk estimation. *Proceedings of the European Safety and Reliability Conference (ESREL 2007)*, INL/CON-07-12774, https://inldigitallibrary.inl.gov/sites/sti/sti/3735689.pdf.

Brown W (2022) Data provided to the author via email, October 14.

Doğan, O, Giritligil AE (2014) Implementing the Borda outcome via truncated scoring rules: a computational study. *Public Choice* 15(1), https://doi.org/10.1007/s11127-012-0019-9.

Hasemyer D (2019) U.S. military knew the flood risks at Nebraska's Offutt Air Force Base, but didn't act in time. NBC News. Accessed May 01, 2023, https://www.nbcnews.com/news/us-news/u-s-military-knew-flood-risks-nebraska-s-offutt-air-n985926.

Jacob C, Dubois D, Cardoso J, Ceberio M, Kreinovich V (2011) Estimating probability of failure of a complex system based on partial information about subsystems and components, with potential applications to aircraft maintenance. Technical Report UTEP-CS-11-23, The University of Texas at El Paso, El Paso, TX, https://scholarworks.utep.edu/cs_techrep/615/.

Kilgour, DM, Grégoire J, Foley AM (2022) Weighted scoring elections: is Borda best? *Social Choice and Welfare* 58, no 2, https://doi.org/10.1007/s00355-021-01347-6.

Offutt Air Force Base (2019) Flood rebuild. Accessed May 01, 2023, https://www.offutt.af.mil/Home/Flood-Rebuild/.

Sprent P, Smeeton N (2007) *Applied Nonparametric Statistical Methods*, 4th ed. (CRC Press, Boca Raton, FL).

Teter, M (2014) Applying subject matter expert (SME) elicitation techniques to TRAC studies. Technical Report TRAC-M-TR-14-036, TRADOC Analysis Center, Monterey, CA, https://apps.dtic.mil/sti/citations/ADA616463.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California