Developing Implementation Strategies for Risk Based Inspection (RBI)

October 2024
Final Report

Project number TR201910
MoDOT Research Report number cmr 24-020

PREPARED BY:

Glenn A. Washer, Ph.D., P.E., University of Missouri

Henry Brown, P.E., Co-Principal Investigator, University of Missouri

Robert Connor, Ph.D., P.E., Purdue University

Mohammad Hammed, Ph.D., Post Doctoral Associate, University of Missouri

Research Assistant(s): Victor Higgenbotham; Susmit Kute; Blandine Therese Mbianda Kemayour, University of Missouri

PREPARED FOR:

Missouri Department of Transportation

Construction and Materials Division, Research Section

Technical Report Documentation Page

1. Report No.
cmr24-020

2. Government Accession No.

3. Recipient's Catalog No.

4. Title and Subtitle
Developing Implementation Strategies for Risk Based Inspection (RBI)

5. Report Date
June 2024
Published: October 2024

6. Performing Organization Code

7. Author(s)
Washer, G., https://orcid.org/0000-0002-2723-2296
Brown, H., https://orcid.org/0000-0003-1473-901X
Connor, R, https://orcid.org/0000-0002-6964-3317
Hammed, M., Higgenbotham, V., Kute, S., Kemayour, B.

8. Performing Organization Report No.

9. Performing Organization Name and Address
University of Missouri
Civil and Environmental Engineering
E2509 Lafferre Hall
Columbia, Missouri 65211

10. Work Unit No. (TRAIS)

11. Contract or Grant No.
TPF-5(388)
MODOT project # TR201910

12. Sponsoring Agency Name and Address
Missouri Department of Transportation (SPR-B) (lead agency)
Construction and Materials Division
P.O. Box 270
Jefferson City, Missouri 65102

13. Type of Report and Period Covered
Final Report (November 2018-April 2024)

14. Sponsoring Agency Code

15. Supplementary Notes
Conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration. MoDOT research reports are available in the Innovation Library at https://www.modot.org/research-publications.

16. Abstract

This project's goal was to improve asset management through Risk-Based Inspection (RBI) practices. The research was intended to amplify the results of prior research that produced the Proposed Guidelines for Reliability-Based Inspection Practices. Under the new provisions of the National Bridge Inspection Standards (NBIS), bridge owners that implement a risk-based inspection (RBI) analysis can determine risk-based inspection intervals of up to 72 months for certain bridges. This research focused on developing implementation strategies to aid bridge owners in implementing these new NBIS rules. Reliability Assessment (RAP) meetings were held in six states during the research. Risk models were developed and applied to 60 sample bridges. These data were analyzed by comparing results with target values developed during the research. A data-driven methodology for analyzing the risk models based on bridge inventory data was developed. This methodology provides a means of calibrating and verifying the risk models. The method was found to be effective at analyzing the models and communicating their effectiveness. It was also found that there was consistency in the risk models developed by different RAP and these risk models were consistent with the target ranges developed through the research.

17. Key Words
bridge inspection, risk-based inspection, bridge asset management, Monte Carlo Simulation, risk analysis, Bridges; Inspection; Risk assessment; Reliability; Bridge management systems

18. Distribution Statement
No restrictions. This document is available through the National Technical Information Service, Springfield, VA 22161.

19. Security Classification (of this report)
Unclassified

20. Security Classification (of this page)
Unclassified

21. No. of Pages
114

22. Price

DEVELOPING IMPLEMENTATION STRATEGIES FOR RISK BASED INSPECTION (RBI)

By
Glenn A. Washer, Ph.D., P.E., Professor
University of Missouri

Co-Principal Investigator
Henry Brown, P.E., Research Engineer
University of Missouri

Robert Connor, Ph.D., P.E.
Purdue University

Post Doctoral Associate
Mohammad Hammed, Ph.D.
University of Missouri

Research Assistant(s)
Victor Higgenbotham, Susmit Kute, Blandine Therese Mbianda Kemayour
University of Missouri

Prepared for
Missouri Department of Transportation

June 2024

Final Report

Copyright

Disclaimer

Acknowledgments

The research team wishes to thank the participating partners for supporting the research through the pandemic and to completion. In particular the research team would like to thank the members of the RAPs that provided invaluable input and insight, as well as the foundational data for the report.

Table of Contents

List of Tables

List of Figures

List of Equations

List of Abbreviations and Acronyms

ADT	Average Daily Traffic
ADTT	Average Daily Truck Traffic
CF	Consequence Factor
CP	Corrosion Protection
CR	Condition Rating
CS	Condition State
Delam	Delamination
ECR	Epoxy Coated Reinforcing
DOT	Department of Transportation
FHWA	Federal Highway Administration
LRF	Load Rating Factor
NBI	National Bridge Inventory
NBIS	National Bridge Inspection Standards
NDT	Nondestructive Testing
NSTM	Nonredundant Steel Tension Member
OF	Occurrence Factor
POF	Probability of Failure
PSC	Prestressed Concrete
RAP	Reliability Assessment Panel
RBI	Risk Based Inspection
R/C	Reinforced Concrete
SS	Superstructure
SNBI	Specification for the National Bridge Inventory
Stl	Steel
Sub	Substructure

Definitions

Attributes: Characteristics that affect the reliability of a bridge or bridge element.

Condition Attributes: Characteristics that relate to the current condition of a bridge or bridge element. These may include element ratings, component ratings, and specific damage modes or mechanisms that significantly affect an element's reliability.

Consequence Factor: A factor describing the expected outcome or result of a failure.

Damage mode: Typical damage affecting the condition of a bridge element (e.g., spalling of concrete, cracking, etc.).

Delphi process: The Delphi process is a method of expert elicitation that involves consulting a panel of experts through a series of systematic feedback rounds to develop consensus opinions on parameters needed for decision-making. Experts are surveyed anonymously and then consensus is formed.

Design Attributes: Characteristics of a bridge component or element that are part of its design. These attributes typically do not change over time except when renovation, rehabilitation, or preservation activities occur.

Deterioration mechanism: Process or phenomena resulting in damage to a bridge element (e.g., corrosion, fatigue, etc.).

Element: Identifiable portions of a bridge made of the same material, having a similar role in the performance of the bridge, and expected to deteriorate in a similar fashion.

Failure: Termination of the ability of a system, structure or component to perform its intended function (API, 2016). For bridges, the condition at which a given bridge element is no longer performing its intended function to safely, and reliably, carry normal loads and maintain serviceability.

Loading Attributes: Loading characteristics that affect the reliability of a bridge or bridge element such as traffic or environment.

Occurrence Factor: Factor describing the likelihood that an element will fail during a specified time period.

Operational Environment: The operational environment is a combination of the circumstances surrounding and potentially affecting the in-service performance of bridges and bridge elements. These include typical loading patterns, ambient environmental conditions, construction quality and practices, maintenance and management practices, and other factors which may vary between different geographic regions and/or organizational boundaries.

Probability: The extent to which an event is likely to occur during a given time interval (API, 2016). This may be based on the frequency of events, such as in the quantitative probability of failure, or on degree of belief or expectation. Degrees of belief about probability can be chosen using qualitative scales, ranks or categories such as “Remote/Low/Moderate/High” or “Remote/Unlikely/Moderate/Likely/Almost Certain.”

Reliability: Ability of an item, component, or system to operate safely under designated operating conditions for a designated period of time or number of cycles.

Risk: Combination of the probability of an event and its consequence.

Risk Analysis: Systematic use of information to identify sources and estimate risk. Information can include historical data, theoretical analysis, informed opinions, and engineering judgment.

Risk Model: A collection of attributes, criteria, and weights used to assess the level of risk.

Screening Attribute: Characteristics of a bridge or bridge element that:

- Make the likelihood of serious damage unusually high,
- Make the likelihood of serious damage unusually uncertain,
- Identify a bridge with different anticipated deterioration patterns than other bridges in a group or family.

Executive Summary

The goal of this project was to improve asset management through the implementation of risk-based inspection (RBI) practices. The research was intended to amplify the results of the National Cooperative Highway Research Program research that produced the *Proposed Guidelines for Reliability-Based Inspection Practices* (Washer et al., 2014). Under the new provisions of the National Bridge Inspection Standards (NBIS), bridge owners that implement an RBI analysis can determine risk-based inspection intervals of up to 72 months for certain bridges. The research focused on developing implementation strategies to assist bridge owners in implementing these new NBIS rules.

Background

When the NBIS were first introduced following the collapse of the Silver Bridge, the interval for routine inspection of bridges was established as 24 months. This inspection interval was applied uniformly across the bridge inventory. The interval was established based on engineering judgment and did not consider factors such as the age, durability, or condition of a bridge. A more rational inspection planning approach would involve determining the interval and scope of inspections according to the bridge's condition and the likelihood of damage. The primary objective of an RBI analysis is to prioritize bridges in terms of their inspection needs by considering factors that affect the likelihood that a given bridge will deteriorate significantly over a given time interval. The RBI process analyzes the inspection needs for bridges and adjusts the inspection intervals to match the needs. This leads to a more effective allocation of inspection resources by focusing on bridges with the greatest inspection needs. In this way the RBI approach can improve bridge safety by increasing inspection efforts where they are most needed and decreasing the inefficient application of inspection resources applied to bridges where risk is low, i.e., the likelihood and consequence of damage is low.

The process for risk-analysis for RBI has two primary components – first, an estimate of the likelihood of serious damage developing, and second, an assessment of the potential consequences of that damage. The likelihood of serious damage occurring is estimated based on an Occurrence Factor (OF), which is a measure of the relative likelihood of damage based on expert judgement. The OF is estimated based on attributes of a bridge component that affect its reliability. Attributes are identified by a group of experts at the owner level known as a Reliability Assessment Panel (RAP). During the initial stages of this project, RAP meetings were held in six states to identify attributes for different bridge components. Criteria for ranking the attributes within a simple quantitative scoring methodology were also developed through the RAP meeting. The attributes and related criteria form a risk model for assessing the relative likelihood of damage or deterioration affecting the component being analyzed. The risk models developed through the research are documented and analyzed in this report.

The consequence associated with different damage modes in bridge components is estimated based on a Consequence Factor (CF). The CF is estimated based on the effect of the damage on the ability of the component to safely, and reliably, carry normal loads. Typically, the CF is determined based on the redundancy of the structure, traffic volumes, service level, and feature under the bridge, among other parameters.

The research included holding six RAP meetings in different states to develop risk models for steel and prestressed concrete bridges. The results from the RAP meetings were used to form a series of risk models applied to a sample of 60 bridges to assess their effectiveness. The risk models were applied based on inspection records from the time interval of 2004 through 2020 in a process known as "back-casting." The results showed that the risk models were generally consistent with target ranges for bridge components with different condition ratings when minor adjustments were made to the risk models. A new methodology was developed for analyzing risk models that supported calibrating the risk models to meet target ranges, testing the quality of the risk models, and forecasting risk assessment outcomes based on bridge inventory data. This methodology was based on Monte Carlo simulations and provides a tool for effective development and implementation of the risk-based inspections for bridges.

Chapter 2 of the report provides an analysis of NBIS requirements related to the research and develops rational target ranges that can be used to assess the quality of risk models developed by RAPs. Chapter 3 of the report reviews outcomes from RAP meetings held during the research period. The risk models developed by individual RAPs were applied to a sample of 60 bridges, and the results and analysis of these data are reported in Chapter 4. Chapter 5 of the report develops a new methodology for analyzing risk models based on MC simulations that provides a data-driven process for risk analysis of bridge inventories.

The report includes examples of risk models and analysis of the risk models, criteria and commentary for attributes, and methodologies for analyzing and calibrating the risk models. Case studies in the form of RAP models developed by different states, revised through the research, and analyzed against bridge inventory data, are included in the report and the appendices. Commentary supporting the new attributes were combined with existing commentary from the previous research effort (Washer et al., 2014). These elements of the report form guidance for implementing RBI, the research's main objective.

The conclusions from the research were as follows:

- The methodology developed from the research for analyzing risk models provides a tool for calibrating the weights of attributes in risk models, adjusting criteria for attributes, and forecasting the outcome of the risk models when applied to bridges. The methodology based on MC simulations can provide an effective tool for analyzing risk models and communicating their effectiveness.
- It was found that there was consistency in many of the damage modes and attributes identified by RAPs. The research showed that six different RAPs assessed risk factors in a generally consistent manner, with certain differences related to their inventory of bridges and local environment. This conclusion supported the concept of the RAPs as an effective tool for risk analysis.
- It was found that the risk models developed by the RAPs were consistent with target ranges based on NBIS requirements. Given that the RAPs were from six different states and were composed of individuals with diverse backgrounds and experiences, the consistency of the RAP outcomes with the target ranges was a significant finding.
- The original back-casting procedure envisioned for the research had limited effectiveness due to inconsistencies in inspection data format and content and changing inspection requirements. This was complicated by the diversity of inspection practices between different states. A single state analyzing their own inspection data would likely have more success. Also, insight into the quality of the risk models could not be obtained from reviewing inspection results.

The recommendations from the research are as follows:

- RAPs should prepare for their meetings by obtaining certain statistics describing their bridge inventories that would be expected to arise during the consideration of criteria for damage modes. Examples of these statistics include Average Daily Traffic (ADT) and Average Daily Truck Traffic (ADTT) levels, a means of estimating areas of the state where deicing chemical use is relatively high, and some statistics on typical element-level condition state quantities for bridge components with different condition ratings. This would reduce efforts following the RAP meeting to formulate criteria threshold values and improve the process's efficiency.

- The risk matrix proposed in the NCHRP 782 report should be modified to better meet the new requirements of the NBIS.
- Special inspection procedures to collect the necessary data to support attribute criteria are needed if element-level inspection data is unavailable for a bridge.

Chapter 1 Introduction

The goal of this project was to improve asset management through the implementation of risk-based inspection (RBI) practices. The research was intended to amplify the results of the National Cooperative Highway Research Program (NCHRP) research that produced the *Proposed Guidelines for Reliability-Based Inspection Practices* (NCHRP 782) (Washer et al., 2014). The process described in that report has become part of the National Bridge Inspection Standards (NBIS) (FHWA, 2022a). Under the new rules, bridge owners that implement RBI analysis can develop extended inspection interval policies that include RBI intervals of up to 72 months for bridges in good condition. This research project was focused on developing implementation strategies to assist bridge owners in implementing these new NBIS rules. The project was initiated prior to the new rules being implemented, and adjustments were required during the project to address the final rules in the NBIS. This study was the first large study to explore implementing the new methodology introduced in the NBIS. The study included Reliability Assessment Panel (RAP) meetings in six states. Risk models were developed based on RAP inputs, and a data-driven process for analyzing risk models was developed and is described in this report. The data-driven process for analyzing risk models forms the backbone for implementation of the technology by calibrating risk models and demonstrating their effectiveness. The study also included deterioration modeling of bridge inventories to provide the foundation for extended inspection intervals and developing attribute commentary to serve as a resource for future users of the technology.

The research will improve asset management tools available to optimize limited resources and ensure the safety and serviceability of bridges and highways. The research includes the development of a handbook for implementation of RBI practices that will provide a resource to bridge owners by describing the RAP process and presenting an example risk assessment for a bridge.

1.1. Objectives

The research's main objective was to develop a handbook for RBI of highway bridges. A series of studies were conducted in cooperation with partnering states to analyze families of bridges, develop suitable risk-based models for determining appropriate inspection intervals, and verifying those models through analysis of historical data and modeling based on Monte Carlo (MC) simulations.

To meet the project goals, the following additional objectives were part of the research:

- Study the implementation of RBI processes within programs in partnering states.
- Develop a paradigm for conducting risk-based inspection analysis within the states through training and workshops with participating states. This objective was addressed through RAP meetings completed in cooperation with each state.
- Develop strategies for using data-driven risk analysis within the RBI framework.
- Develop methodologies for supporting preservation activities within an RBI framework.
- Study the reliability of inspection technologies and implementation of nondestructive testing (NDT) to support RBI.

1.2. Scope of Work and Report

The scope of the overall project included holding Reliability Assessment Panel (RAP) meetings with six of the nine participating states to develop initial risk models for families of bridges. The study included four states analyzing bridges with steel superstructures and two states analyzing bridges with prestressed concrete (PSC) superstructures. These families of bridges included typical highway bridges that are subject to routine inspection and did not consider nonredundant steel tension members (NSTMs) or bridges that may have special inspection needs, such as signature or large bridges. During the research, deterioration modeling of each state's bridge inventory was conducted to establish the typical service life for the families of bridges considered in the study. These data were based on component-level National Bridge Inventory (NBI) Data and survival analysis to determine estimates of the Time in Condition Rating (TICR) for bridge components. These data from deterioration models provide a foundation for extended inspection intervals based on the long deterioration patterns of typical bridges.

The research analyzed the risk models developed by the separate RAPs and updated them to meet NBIS rules regarding bridges eligible for extended inspection intervals. A group of randomly selected sample bridges were used to analyze the risk models when applied to historical bridge inspection data. This process, termed "back-casting," was intended to analyze the quality of the risk models and identify any significant deficiencies revealed by examining the historical records. The risk models were applied to bridges based on inspection data from the years 2004 to 2020. The initial efforts at back-casting were conducted before the new NBIS rules were issued. The back-casting process was repeated after the new NBIS rules were issued to consider the limitations on the process with the new rules in place.

A data-driven process was developed to allow "users" (i.e., bridge owners implementing RBI analysis) to analyze the risk models developed by a RAP, make rational adjustments to the weight of attributes in the models, and verify the performance of the risk models. This process uses existing bridge inventory data, to the extent possible, to estimate the potential outcomes of the risk models when applied to the intended family of bridges.

This report focuses on the overall outcome of the research. Two interim reports were completed as part of the research. The first interim report described the RAP meetings and documented the inputs and data acquired from the meetings. A second interim report described the data-driven process for analyzing risk models developed through the research. This is the final report that summarizes the overall research including summaries of the key results from the RAP meetings and development of the data-driven data analysis methodology.

Chapter 2 of the report provides an analysis of NBIS requirements related to the research and develops rational target ranges that can be used to assess the quality of risk models developed by RAPs. Chapter 3 of the report reviews outcomes from RAP meetings held during the research. Chapter 4 summarizes the back-casting results for a population of 60 sample bridges. The risk models developed by individual RAPs were applied to a sample of 60 bridges, and the results and analysis of these data are reported in this chapter. Chapter 5 of the report describes an innovative new methodology for analyzing risk models based on MC simulations and provides a data-driven process for risk analysis of bridge inventories. Chapter 6 describes a categorical model for the Consequence Factor (CF), and Chapter 7 includes the conclusions and recommendations from the research.

1.2.1. Overview of Methodology

This section of the report provides a brief overview of the RBI process being studied through the research. The RBI process discussed in this report is based on previous research reported in NCHRP Report 782, "*Proposed Guideline for Reliability-based Bridge Inspection Practices*," and included in the latest revisions to the NBIS. The summary provided here is to help review the report and does not include all process details. A detailed description of the methodology can be found in Appendix A, *Handbook for Implementation of RBI*.

1.2.1.1. Rationale for RBI Intervals

Risk analysis is something engineers do every day in formulating designs and maintaining systems. The fundamental idea behind engineering analysis of risk is to estimate how likely it is for a certain adverse event to occur, and to estimate the potential consequences of that event. This type of analysis occurs every day in an ad-hoc manner as a part of decision-making regarding, for example, identifying urgent repairs or bridges the require an inspection interval of less than 24 months. When the NBIS was first introduced following the collapse of the Silver Bridge, the interval for routine inspection of bridges was established as 24 months. The 24-month interval was based on engineering judgement and no quantitative engineering assessment was performed in establishing the interval. The interval was established based on the rationale that a one-year interval was too short, and a ten-year interval seemed far too long. Hence, 24-months was settled upon under the assumption that this interval was adequate to monitor the condition of a bridge and detect serious deterioration or damage prior to an adverse event.

The time-based NBIS interval was applied uniformly across the bridge inventory, resulting in the same inspection interval for new bridges as for aging and deteriorated bridges. Newer bridges have improved durability and are typically in good condition. Certain older bridges exhibit minimal deterioration, even after many years of service, while other bridges of the same age might be severally deteriorated, depending on attributes of the bridge such as design features, environment, materials, and loading. Bridges located in arid or mild environments may be in service for many years with little deterioration. Bridges in more aggressive environments deteriorate more rapidly but many steps are taken to increase the durability of these bridges. Design features, materials, and corrosion protection strategies are used to minimize deterioration and ensure a long service life. Bridges with attributes like high traffic volume may deteriorate more quickly as compared with bridges that have very low volumes. Regardless of these facts, a uniform inspection interval was established at the outset for the NBIS, and this uniform interval has existed for many years, with some exceptions to allow extended intervals to 48-months introduced in the 1980's but not implemented widely (FHWA, 1988).

A more rational approach to inspection planning would determine the interval and scope of an inspection according to the bridge's condition and the likelihood of damage. Principles of reliability and risk assessment are used in many industries to match inspection requirements to inspection needs. For example, pipelines, offshore structures, components in nuclear power plants, and dams all have risk-based approaches integrated within their inspection programs. These methodologies evaluate the specific characteristics of components (i.e., attributes) such as resistance to damage and deterioration, current condition, and loading history to analyze the reliability of the component and determine appropriate inspection requirements. In this way the safety and operation of the component is maintained over its service life. The RBI procedure for bridge inspection was modeled on these procedures from other industries and customized for the unique needs of bridges.

The primary objective of the risk analysis is to prioritize bridges in terms of their inspection needs by considering the factors that affect the likelihood that a given bridge will deteriorate significantly in the next 72-months. Figure 1.1 illustrates the underlying concept for prioritizing bridges and bridge components (deck, superstructure, and substructure) based on attributes such as corrosion protection, exposure environment, loading, and condition. The figure shows a log-normal plot of the time-in-condition-rating (TICR) for a bridge deck with a condition rating (CR) of 7, good condition. The plot shows that there are some bridge decks that deteriorate rapidly and have a short TICR, shown in area "A" of the plot. These are decks that may have poor durability attributes, emerging damage, or construction defects, and/or may be in an aggressive environment. These decks have a short TICR and transition to CR 6 after only a short interval, many in fewer than six years. On the other hand, there are many bridges that fall in area "B" of the plot. Many of the decks in area "B" stay in CR 7 for 20 years or more. As will be discussed later in the report, the median TICR for bridge decks across the participating states was found to be greater than 11 years based on analysis of the NBI data records. The median value shown in the figure (11 years) represents 50% of the decks having a TICR less than 11 years and 50% of decks having a TICR of 11 years or greater.

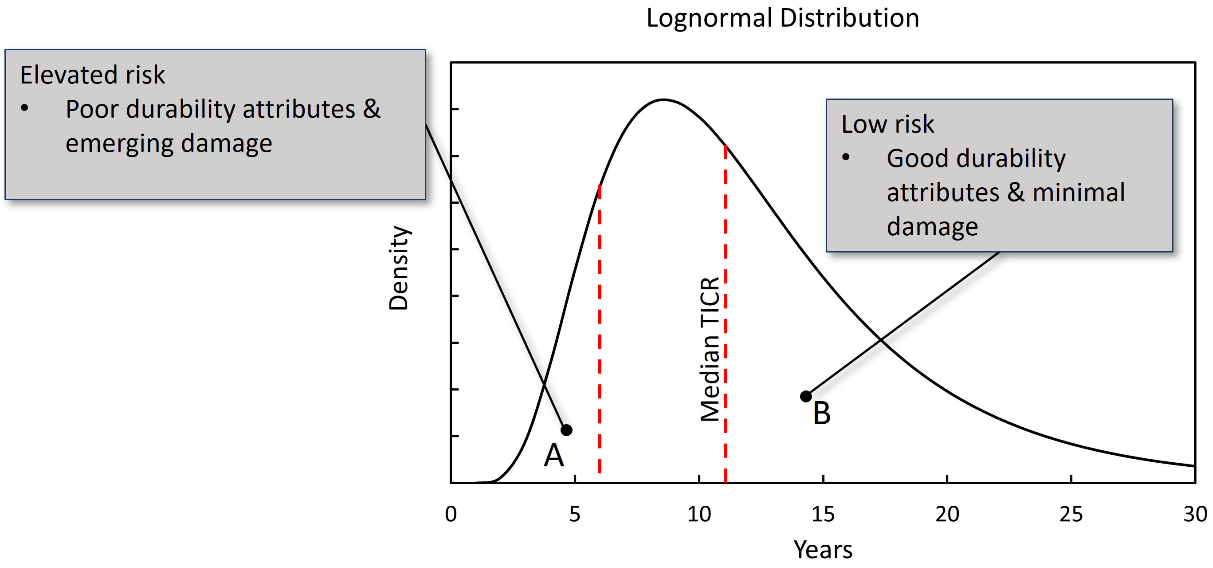


Figure 1.1. Typical deterioration pattern for a bridge deck component in CR 7.

A primary objective of the RBI analysis discussed in this report is to identify those bridge components that are in area "B" of the plot shown in Figure 1.1 by analyzing their attributes combined with expert judgement and experience. In this way a suitable inspection interval can be applied that focuses inspections on components with poor attributes and damage such as those in area "A" of the plot and extended inspection intervals for components in area "B" of the plot. For example, if a certain deck has a TICR of 11 years and is inspected every 24 months, there are five inspections conducted that report no change in CR. Therefore, the cost of the inspection and the risk placed on inspectors having to access the bridge and work in traffic areas has little benefit in terms of assessing the condition of the bridge. This also reduces the perceived value of the inspections since inspectors are required to frequently inspect a bridge even though the condition is not changing. This can lead to complacency that undermines the reliability of the inspections because the inspections may be perceived as mundane tasks of little practical value or importance.

A more effective strategy is to analyze the inspection needs for bridges and adjust the inspection intervals to match the needs. This leads to a more effective allocation of inspection resources, reduces the risks to inspectors deployed to conduct the inspection, and focuses resources on bridges where deterioration and damage are likely to be occurring. In this way the RBI approach can improve bridge safety by increasing inspection efforts where most needed, increasing the perceived importance of the inspection task, and decreasing the inefficient application of inspection resources applied to bridges where risk is low, i.e., the likelihood of damage is remote or low.

1.2.1.2. Methodology

The process for risk-analysis for RBI has two primary components – an estimate of the likelihood of serious damage developing in the next 72 months, and an assessment of the potential consequences. The likelihood of serious damage occurring is estimated based on an Occurrence Factor (OF), which is a measure of the relative likelihood of failure based on expert judgement. The OF is estimated based on *attributes* of bridge components, which are characteristics of a bridge component that affect its reliability. Generally, these attributes are characteristics that affect the durability of the component. For example, epoxy-coated reinforcing (ECR) steel is an attribute of a bridge deck that provides increased durability as compared with uncoated steel reinforcing. Attributes are identified by a group of experts at the owner level known as a RAP.

The consequence associated with different damage modes in bridge components is assessed based on a CF. The CF is estimated based on the effect of the damage on the ability of the component to safely, and reliably, carry normal loads. Input from the RAP is also sought to provide criteria to estimate the CF. Typically, the CF is determined based on the redundancy of the structure, traffic volumes, service level, and feature under the bridge, among other parameters.

The attributes identified for determining the OF are scored using criteria developed through the RAP meeting to estimate the OF; attributes associated with the consequence of each damage mode are used to estimate the CF. These two factors are combined to locate a particular bridge component on an example risk matrix as shown in Figure 1.2. Bridge components that tend toward the lower left corner of the matrix have lower risk and require less frequent inspections; components that tend toward the upper right corner have higher risk and require more frequent inspection. Inspection intervals envisioned by the methodology range from 12 to 72 months, with the lowest risk bridges being assigned a 72-month interval. Analysis of the risk matrix shown in Figure 1.2 is included in this report.

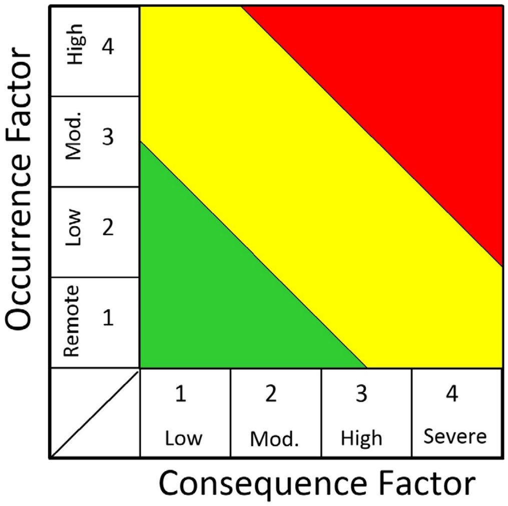


Figure 1.2. Risk matrix for Risk-based Bridge Inspection.

1.2.1.3. Attributes

A key element of the RBI process is to identify attributes that impact the durability of bridge components and the probability of failure (POF). The definition of “failure” adopted for risk analysis of bridge components is as follows: *The condition at which a given bridge element is no longer performing its intended function to safely, and reliably, carry normal loads and maintain serviceability* (Washer et al., 2014). This condition is deemed to be analogous to condition rating (CR) 3, Serious condition, for the purposes of the risk analysis (FHWA, 2022c). The definition of CR 3 is provided in the Specifications for the National Bridge Inventory (SNBI) as “Major defects; strength and/or performance of the component is seriously affected. Condition typically necessitates more frequent monitoring, load restrictions, and/or corrective actions” (FHWA, 2022c).

The attributes are divided into groups for organizational purposes. Attribute groups include screening attributes, design attributes, loading attributes, and condition attributes. Screening attributes are attributes that make the POF unusually high or unusually uncertain such that the bridge should be screened from the RBI process. For example, a concrete bridge with shear cracking may not be suitable for an extended inspection interval, regardless of its other attributes, and should therefore be screened out of the risk analyses process. Design attributes are related to design features such as type of reinforcement used (ECR vs uncoated rebar) and typically do not change during a bridge's service life. Loading attributes commonly describe loads or stressors applied to an in-service bridge, such as Average Daily Truck Traffic (ADTT) or rate of deicing chemical application. Finally, condition attributes are related to the current condition of a bridge component or element. Condition attributes are typically assessed based on inspection data. Examples of condition attributes include the CR of components and condition state (CS) of elements, condition of joints, and other damage indicators. Attributes are identified using an alpha numeric moniker to link a given attribute to commentary that describes the rationale or reason for the attribute.

A complete description of the RBI process for identifying attributes, developing risk models based on input from an RAP, and determining an RBI inspection interval is provided in Appendix A, *Handbook for Implementation of Risk-Based Inspection (RBI)*.

Chapter 2 Analysis of RBI Requirements

This chapter describes and analyzes the recent changes in the NBIS and how these changes affect the implementation of RBI for bridges. The purpose of this analysis was to compare the new NBIS requirements with the risk matrix proposed in the original NCHRP 782 report and establish target ranges for risk models.

2.1. RBI and NBIS Analysis

The update to the NBIS published in 2022 included requirements for implementing RBI intervals for routine, underwater, and NSTM inspections (FHWA, 2022a). The Federal Highway Administration (FHWA) subsequently issued a memorandum with the subject “National Bridge Inspection Standards Inspection Interval Guidance” to provide additional information and assistance for bridge owners implementing the new NBIS requirements (FHWA, 2022b). This memo addressed the two methods identified in the NBIS for determining the inspection interval, named Method 1 and Method 2. Method 1 is a simplified risk assessment approach to determine reduced and extended intervals for routine, underwater, and NSTM inspections. Extended intervals of up to 48 months are allowed for bridges meeting certain criteria defined within the NBIS and clarified with the FHWA guidance memo. In general, Method 1 requires that bridge components have a CR of 6 or higher, have a load rating factor (LRF) of 1.0 or greater, minimum vertical clearance of at least 14 ft, and minimal scour vulnerability. Bridge owners must also consider other factors such as material, ADT, design, etc. in developing a Method 1 policy.

Method 2 is a more rigorous approach that allows for risk assessment by quantified statistical analysis and/or qualitative expert judgement. The maximum routine inspection interval using Method 2 is 72 months, and only bridges in “good” condition are eligible for a 72-month interval. A bridge in “good” condition has the minimum (i.e., lowest) CR of 7, 8, or 9 (FHWA, 2022c). The risk models being formed in this project are the first risk models and processes developed using Method 2 under the new policies. This section of the report summarizes the requirements for the two methodologies to provide context on the needs, criteria, and opportunities within the new risk-based approach to inspection planning. Some of the Method 1 requirements' characteristics are useful for analyzing the risk framework and providing general guidance on requirements for Method 2 analysis, as described in this section of the report. Certain data from the back-casting analysis completed in this research are also presented to illustrate how the new requirements align with research results.

2.1.1. Method 1 Analysis

Method 1 allows for bridges meeting certain criteria to have extended routine inspection intervals of up to 48 months. Table 2.1 summarizes the criteria established in the NBIS and the FHWA memo for an extended 48-month inspection interval. The detailed criteria shown in Table 2.1 refer to items defined in the traditional FHWA Recording and Coding Guide (i.e., the *Coding Guide*) and the new Specifications for the National Bridge Inventory (SNBI) (FHWA, 1995, 2022c).

A key element of Method 1 is bridges that meet the criteria listed in Table 2.1 can be assigned 48-month inspection intervals without FHWA approval when the bridge owner establishes an extended inspection interval policy. The extended interval policy must consider other factors such as structure type, design, materials, etc. determined by the bridge owner. The factors identified by the bridge owner are intended to capture other risks not included in the Method 1 requirements based on expert judgement and knowledge of their bridge inventory. This allows the bridge owner to assign the 48-month interval for any bridge meeting the identified NBIS criteria and additional factors the owner has included in their extended interval policies. The routine inspection interval is reduced to the traditional 24-month interval when one or more of the Method 1 criteria are not met.

The criteria for Method 1 do not include broader risk factors such as durability characteristics of a bridge, the aggressiveness of the environment, or other factors that are expected to be considered in a Method 2 analysis and are included in the risk models developed during this research. Additional factors expected in the Method 2 analysis include Average Daily Traffic (ADT), the feature under the bridge, and the degree of redundancy.

The Method 1 criteria are useful for analyzing Method 2 assessments to determine if the assessment generally meets FHWA requirements, although there can be differences since the Method 2 analysis involves different criteria and a more comprehensive approach to the analysis. A primary difference between Method 1 and Method 2 is that various attributes that affect the likelihood of damage developing in the future are incorporated in Method 2. Method 1 analysis relies entirely on the present condition of the bridge. Because Method 2 includes attributes that look forward in time at the potential for damage, not just the present damage, it provides a more rigorous analysis that may produce criteria that vary from the Method 1 criteria. However, the Method 1 criteria provide a general framework for RBI analysis when implementing intervals determined through Method 2 analysis. For example, the scour vulnerability criteria shown in Table 2.1 would likely be required under most Method 2 risk models. It should be noted that there is not an explicit requirement that the Method 1 criteria be met when implementing Method 2. For example, one of the Method 1 criteria prohibits bridges with E or E' details from having an extended interval. If the bridge had minimal loading such that likelihood of fatigue damage was *remote* or an analysis showed infinite fatigue life, a Method 2 analysis theoretically could be used to establish an extended routine inspection interval. The extended inspection interval policy is subject to FHWA review.

2.1.2. Method 2 Analysis

Method 2 risk assessment allows routine inspection intervals of up to 72 months based on a risk assessment process developed by a RAP. The method requires that a set of screening criteria be used to determine how bridges will be considered in the assessment and to establish maximum inspection intervals. Five different requirements for screening criteria are shown in Table 2.2. The first three screening criteria are to be developed by the RAP and must include flexural and shear cracking in concrete members and fatigue cracking and corrosion in steel members. Criteria for considering details, loadings, conditions, etc. that are likely to affect safety and serviceability of bridges must also be included in the screening criteria. The final two required screening criteria are specified and indicate the maximum allowable inspection intervals based on general CRs. These requirements indicate that the maximum interval for bridges classified as being in "Fair" condition, i.e., bridges with a lowest component rating of CR 5 or 6, is 48 months (FHWA, 2022c). The maximum interval for bridges classified as being in "Poor" condition, i.e., CR less than or equal to 4, is 24 months.

The required screening criteria indicate that only bridges classified as being in "Good" condition, i.e., with CRs of CR 7 or greater, are eligible for a 72-month interval. Bridges in "Fair" condition have a maximum interval of 48 months, indicating that bridges with CR 5 or 6 could be eligible for a 48-month interval even if the bridges do not meet the Method 1 criteria. For example, a bridge that does not meet one or more of the criteria for an extended interval under Method 1 may be eligible for an extended interval if Method 2 analysis is completed.

Requirements for attributes and deterioration modes that should be included in a risk model are summarized in Table 2.3. The table rows are numbered 1-5 for reference. Row 1 of the table provides a list of the attribute types that must be included in each analysis, including material properties, loads, safe load capacity, and condition. Rows 2 and 3 list deterioration modes based on the material that forms the bridge. The deterioration modes for steel members must include section loss, fatigue, and fracture. Models for concrete structures should include damage modes of flexural cracking, shear cracking, and corrosion of reinforcing steel. There are also component-level requirements described for the bridge superstructure and substructure (rows 4 and 5). Superstructure member deterioration modes must include settlement, impact damage, rotation, and overload. Substructure component deterioration modes must include settlement, rotation, and scour.

Table 2.1 Summary of FHWA requirements for Method 1 analysis.

Description	Coding Guide Item	Coding Guide Criteria	SNBI Item	SNBI Criteria
Deck CR	58	≥ 6	B.C.01	≥ 6
Superstructure CR	59	≥ 6	B.C.02	≥ 6
Substructure CR	60	≥ 6	B.C.03	≥ 6
Culvert CR	62	≥ 6	B.C.04	≥ 6
Channel Condition	61	≥ 6	B.C.09	≥ 6
Channel Protection Condition	61	≥ 6	B.C.10	≥ 6
Inventory Load Rating Factor	66	LRF ≥ 1.0	B.LR.05	LRF ≥1.0
Routine Permit Loads	-		B.LR.08	A or N
Fatigue Details	-		B.IR.02	N
Highway Minimum Vertical Clearance	53 and 54B	≥ 4.20 m	B.H.13	≥ 14.0 ft
Span Material	43A and 44A	2, 3, 4, or 5	B.SP.04	C01-C05 or S01-S05
Span Type	43B and 44B	01, 02, or 05	B.SP.06	A01, B02-B03, F01-F02, G01-G08, P01-P02, or S01-S02
Scour Vulnerability	113	5, 8, or N	Item B.AP.03	A or B
Scour CR	-	-	B.C.11	≥6

Table 2.2 Required screening criteria for Method 2 analysis.

No.	Requirement
1	Requirements for flexure and shear cracking in concrete primary load members.
2	Requirements for fatigue cracking and corrosion in steel primary load members.
3	Requirements for other details, loadings, conditions, and inspection findings that are likely to affect the safety or serviceability of the bridge or its members.
4	Bridges classified as in poor condition cannot have an inspection interval greater than 24 months.
5	Bridges classified as in fair condition cannot have an inspection interval greater than 48 months.

Most of the deterioration modes and attributes included in the FHWA guidance documents are addressed by the risk models developed in this project. The FHWA guidance provides a general framework for the analysis of these risk models. Importantly, the FHWA guidance provides some target ranges for the analysis used to update the risk matrix that defines the inspection interval for bridges as discussed in the following section.

Table 2.3. Attributes required for Method 2 analysis.

Row No.	Category	Attributes
1	Attributes for each assessment must include:	Material properties, loads, safe load capacity, and condition
2	Steel members damage modes must include:	Section loss, fatigue, and fracture
3	Concrete members damage modes must include:	Flexural cracking, shear cracking, and reinforcing or prestressing steel corrosion
4	Superstructure members damage modes must include:	Settlement, rotation, overload, and vehicle/vessel impact
5	Substructure members damage modes must include:	Settlement, rotation, and scour

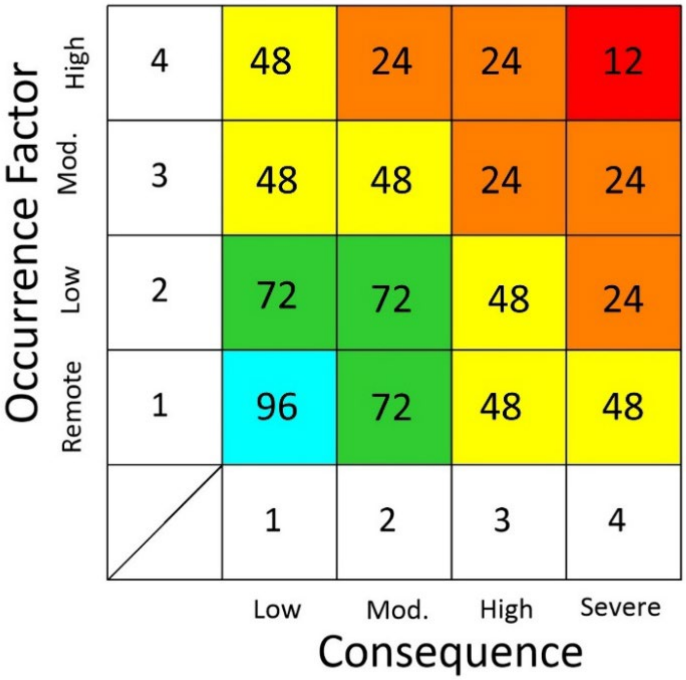
2.1.3. Target Ranges for Risk Models

The general framework provided by the FHWA guidance and the updated NBIS requirements provide some expected outcomes from a risk assessment for determining extended inspection intervals. This framework can be used to make some judgements on the risk matrix used to determine the inspection intervals for bridges based on the OF and CF determined from a Method 2 risk-based analysis. This section of the report discusses proposed changes to the risk matrix based on a combination of the revisions to the NBIS, the associated guidance provided by FHWA, and results from the back-casting.

The project that produced NCHRP Report 782 was the initial effort to develop a reliability-based bridge inspection practice that could be implemented for highway bridges in the US (Washer et al., 2014). The project developed a framework for RBI that was subsequently adopted in the new NBIS. However, the study did not include broad implementation of the framework developed through the research. Further, the 2022 revisions to the NBIS and associated guidance from the FHWA were not available at the time of the study. Therefore, the framework developed in NCHRP 782 needed to be assessed in terms of the new NBIS requirements to enable the implementation of the new policies into practice.

The NCHRP study proposed a risk matrix for typical bridges as shown in Figure 2.1. The 4 x 4 matrix shows the OF on the ordinate (i.e., vertical axis) and the CF on the abscissa (i.e., horizonal axis). The original risk matrix included inspection intervals ranging from 96 to 12 months based on the OF and CF for a given component and damage mode, as shown in the individual elements of the matrix. The specific elements in the matrix are identified based on their location defined using the nomenclature [$R_{row,column}$], referenced from the bottom left corner of the matrix. For a component damage mode rated as OF 4 (*high*) and a CF of 4 (*severe*) the element [$R_{4,4}$] indicates a 12-month inspection interval. For a component damage mode rated as having an OF 1 (*remote*) and a CF of 1 (*low*), the inspection interval would be 96 months based on the original risk matrix.

The new NBIS requirements can be used to analyze the original risk matrix proposed in NCHRP 782 by comparing the ordinate and abscissa values in the original matrix to the new rules. This analysis provides some general guidance on the appropriate inspection intervals for the different elements of the risk matrix.



Occurrence Factor					
High	4	48	24	24	12
Mod.	3	48	48	24	24
Low	2	72	72	48	24
Remote	1	96	72	48	48
		1	2	3	4
		Low	Mod.	High	Severe
		Consequence			

Figure 2.1 Risk Matrix proposed in NCHRP Report 782 showing inspection intervals.

Considering the ordinate, the Method 1 guidance and NBIS requirements designate 48 months as a suitable inspection interval for components in CR 6, provided the component is not a NSTM and is located within a structure meeting the other Method 1 criteria. Potential consequences are not considered explicitly. This provides general guidance on how the elements on the ordinate should be defined because these CR 6 components can have a 48-month interval regardless of the CF. Since Method 1 allows that any CR 6 component is potentially suitable for a 48-month inspection interval, it would follow that a Method 2 analysis should also identify most CR 6 components suitable for at least a 48-month interval. It should be noted that the weighted sum model used to score the risk models produces results ranging from *remote* to *high*, so not all bridge components of any particular CR will lie in a particular element in the matrix. Rather, the risk model for a particular component produces OF results over a range of values based on the attributes and criteria identified by the RAP and the component being assessed.

Based on engineering judgement, most CR 7 components would be expected have a lower OF compared to most CR 6 components, and most CR 6 components should have a lower OF than most CR 5 components. Based on the proposed risk matrix with four categories for the OF, most CR 7 components would be expected to score in the range of *remote* to *low*, most CR 6 components would rank in the *low* to *moderate* range, and most components in CR 5 would lie in the *moderate* to *high* range. Figure 2.2 illustrates a risk matrix showing these ranges on the ordinate. These ranges provide a reasonable and rational ordering of the expected OF values for components based on the CR. The ranking for individual components is refined by the risk models developed by the RAP. For example, certain CR 6 components may score lower than certain CR 7 components when the risk factors (i.e., attributes) in the risk models are assessed.

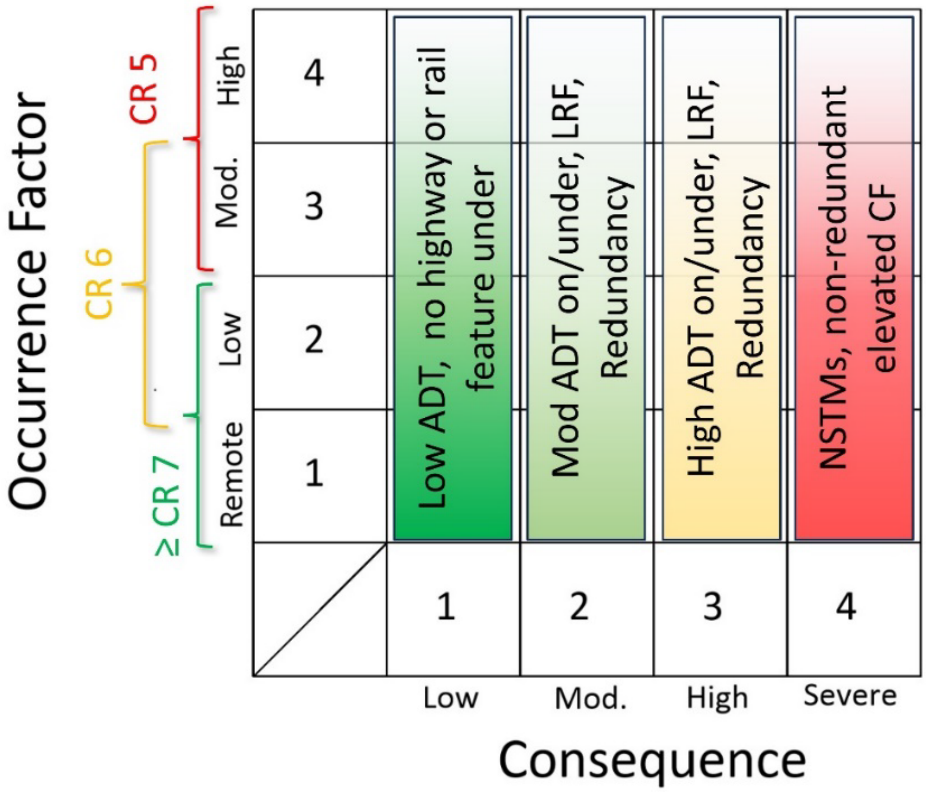


Figure 2.2. Figure showing risk matrix with target ranges for the OF and CF general descriptions.

The horizontal axis can also be analyzed based on the new NBIS rules to infer values in the risk matrix. It is assumed in this analysis that the CF is generally defined in terms of the load capacity of the bridge as expressed by a load rating factor (LRF), the degree of redundancy, feature under the bridge, and traffic volumes (e.g., ADT).

NSTMs require a hands-on inspection at a standard interval of 24 months, which can be extended to 48 months if the criteria for Method 1 for NSTMs are met or if a RAP is used to develop suitable risk models for NSTM inspection. The Method 1 criteria for NSTMs are similar to those for routine inspection but include additional criteria that consider the age of the structure and its fatigue resistance. Historically, the rationale for considering NSTMs differently than redundant steel tension members is an assumed potential for catastrophic collapse resulting from member failure. This is a *severe* consequence, as shown in column 4 in the risk matrix in Figure 2.2. Certain other bridges such as some non-redundant concrete members, structures with only three primary members and wide beam spacing, or other situations where the consequence of serious damage presents substantial risk of life may also have a CF characterized as *severe*. Most bridges that do not have NSTMs would generally be described as having *low*, *moderate*, or *high* consequences, based on characteristics such as the ADT, feature under the bridge, LRF, etc. The Method 1 policy for routine inspections allows components in CR 6 to have an inspection interval of 48 months regardless of the CF being *low*, *moderate*, or *high*. All elements in the risk matrix shown in Figure 2.1 except column 4 could be considered 48-month for CR 6 components under the Method 1 approach.

Most common bridges would typically have a *moderate* or *high* CF. A bridge with a low CF is assumed to be a bridge with uncommonly low ADT and no highway or rail feature under, as shown in Figure 2.2. It is notable that the Method 1 criteria make no mention of ADT levels on or below a bridge in making the assessment that a bridge could have a 48-month interval, although it is among the factors bridge owners may consider in their extended interval policy. It would be reasonable to expect that high ADT alone would not preclude a bridge from a 72-month interval, since ADT is not required in the Method 1 criteria for extended inspection intervals. Further, most CFs would be expected to identify very high ADT as an attribute for components with a high CF. Based on these assumptions and understanding of the NBIS requirements, row 1 of the risk matrix shown in Figure 2.1 should be 72 months for any CF that is not *severe*, such that a CR 7 bridge with *remote* likelihood of failure would qualify for a 72-month interval for CFs of *low*, *moderate*, or *high*.

The distribution of OF values for components in different CRs shown in Figure 2.2 was substantiated when risk models developed by individual RAPs in this project were applied to real bridges. As will be shown in greater detail later in the report, the target values that consider CR 7 components typically having OF in the *remote* or *low* range, and CR 6 components typically having an OF of *low* or *moderate*, etc., were close to those produced from the risk models developed by the individual RAPs and applied to the sample bridges. For example, Figure 2.3 shows a cumulative probability distribution for the OF stemming from the deck component of the 60 sample bridges studied in the back-casting. The abscissa shows the OF category ranging from *remote* to *high* at the bottom of the plot and the numerical values for the risk score at the top of the plot. The ordinate shows the probability of a randomly selected bridge deck having a certain risk score based on the risk models formed by the RAPs in the study. The curve was produced from the original, unweighted risk models developed by the individual RAPs and the risk scores obtained through the back-casting analysis that applied the models to components of the 60 sample bridges. The 60 bridge decks from six states were combined and treated as a single sample population to provide the mean and standard deviation needed to form the cumulative probability distribution plot shown in the figure.

The figure shows that for decks with CR 7, 54% of decks were likely to score in the *remote* range, while 46% would have scores ranging from the *low* to *moderate* range. Those decks scoring in the *remote* range are decks that scored 1.0 or less according to risk models that included attributes such as the CR, CS, rate of deicing chemical application, ADT, corrosion protection level, etc. In other words, these are decks in fair or good condition with good durability characteristics and consequently remote POF (i.e., the likelihood of deteriorating to a CR 3 in the next 72-month interval is *remote*). Decks with CR 6 have increased risk, with only ≈8% of decks expected to score in the remote range, ≈61% scoring in the low range, and the remaining ≈31% scoring in the moderate or high range. Decks in CR 5 are scored with ≈32% in the *low* range, ≈44% scoring in the *moderate* range, and ≈18% scoring in the *high* range. The specific percentage values will obviously vary for different components and different risks models, but these results illustrate the general behavior and trends of the risk models developed by the RAPs and applied to actual bridges. Specifically, the plot shows that the attributes and criteria developed through a Method 2 process identified decks in CR 7 as having relatively lower risk than decks with CR of 6 or 5. Within the group of CR 7 decks, components were rated as having *remote*, *low*, or *moderate* likelihood, which identifies those low-risk components that may be suitable for a 72-month inspection interval and those components with elevated risk.

Methods of calibrating the individual risk models to improve the quality of the results were developed through research based in part on sensitivity studies of the back-casting results. A systematic approach for analyzing and calibrating the risk models using MC simulations was developed and is presented in the report. The values shown in Figure 2.3, which were determined directly from the RAP models developed through the study and applied to real bridges, combined with the FHWA guidance on extended intervals, form the initial expectations for risk categories and risk scores as a function of the CR of a component and provide target ranges for analysis of individual risk models. These target ranges were used to analyze the risk models produced by the RAPs.

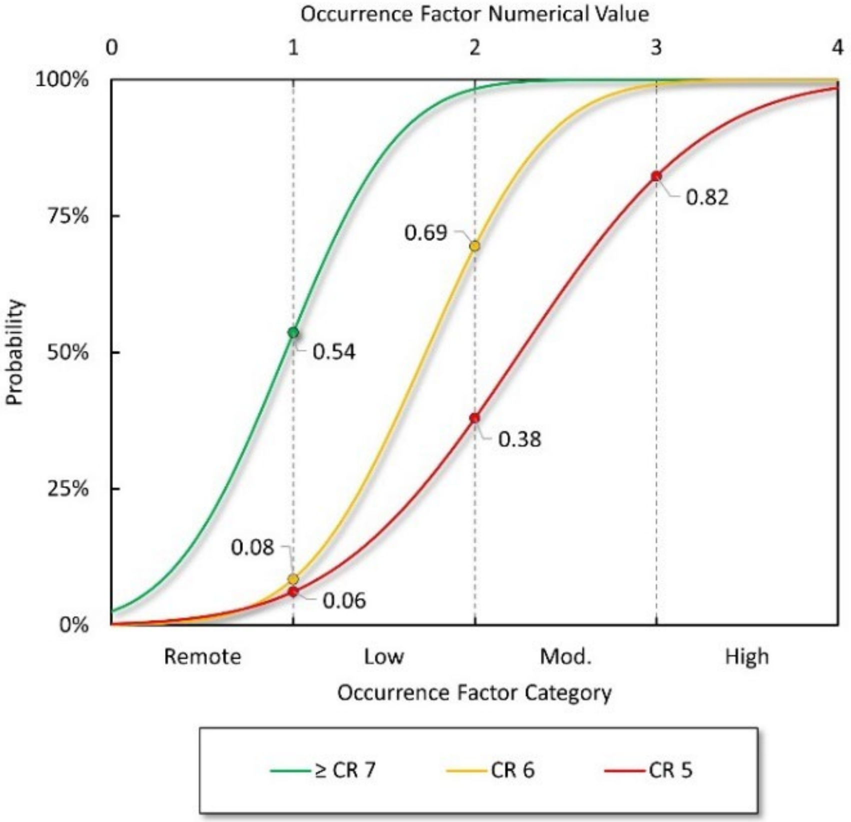
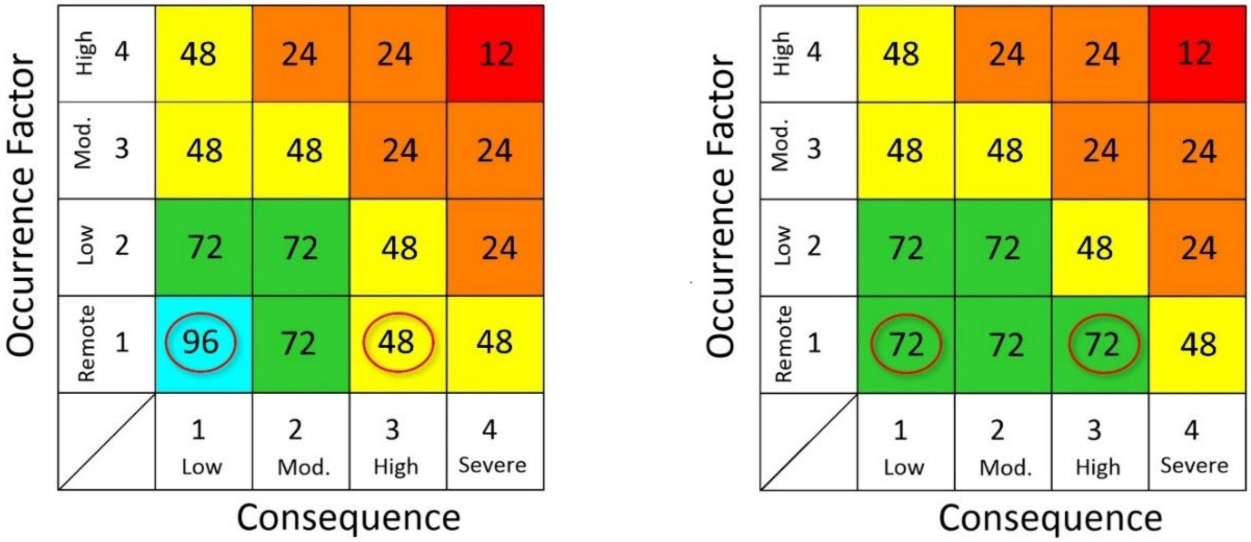


Figure 2.3. Example cumulative probability distribution function for sample bridge decks.

2.1.4. Proposed Changes to the Risk Matrix

Changes to the original risk matrix from the NCHRP 782 report may be justified considering the analysis of the new NBIS requirements and associated target ranges for components with different CRs. Also, a more robust calibration and validation of risk models was completed through this project. Figure 2.4 (A) shows the original risk matrix from NCHRP 782 and Figure 2.4 (B) shows the risk matrix being proposed based on the results of this research. The proposed changes are encircled on each risk matrix. As noted above, most CR 7 components tend to score in the *remote* to *low* range for OF. Since the policy allows a 72-month interval only for bridges with components with CR 7 or greater, it would seem rational that the matrix elements [$R_{1,1}$], [$R_{1,2}$], and [$R_{1,3}$] would be 72 months. This allows CR 7 components with remote likelihood of failure to have a 72-month interval for any CF other than *severe*. The matrix element [$R_{2,3}$] is assigned 48-months and provides granularity in the analysis that aligns with the Method 1 approach that a CR 6 bridge with a high CF can have a 48-month interval. In this way bridges with CFs of *high* are only eligible for 72-month if the OF is *remote*, and the inspection interval is reduced to a 48-month interval when the likelihood is increased from *remote* to *low* (i.e., OF = *low*).

The risk matrix provides a very rational hierarchy shown in Table 2.4 for components with CF of *high,* meaning a bridge has elevated risk based on the CF attributes. A 72-month interval is only possible for CR 7 components with *remote* OF for components with a *high* CF. If the OF for a CR 7 component is *low* the interval is 48 months. Components in CR 6 are expected to have OFs of *low*, resulting in a 48-month interval which aligns with the Method 1 approach. If the CR 6 component has an OF score of *moderate* then the assigned interval would be 24-months, which is more conservative than the Method 1 approach that does not consider the consequence explicitly. If a CR 6 component OF is *remote*, it would seem to qualify for a 72-month interval, although NBIS requirement would not allow a 72-month interval. Regardless, the results from back-casting and MC simulations presented later will demonstrate that there is a relatively low probability of CR 6 components having a risk score of 1.0 or less. For CR 5 components, the interval of 48 months would only apply if the OF was *low*, which is expected to be a relatively small proportion of CR 5 components (see Figure 2.3). Many CR 5 components will score in the *moderate* or *high* range with an assigned interval of 24 months.



(A)

Occurrence Factor	1 Low	2 Mod.	3 High	4 Severe
High 4	48	24	24	12
Mod. 3	48	48	24	24
Low 2	72	72	48	24
Remote 1	96	72	48	48

(B)

Occurrence Factor	1 Low	2 Mod.	3 High	4 Severe
High 4	48	24	24	12
Mod. 3	48	48	24	24
Low 2	72	72	48	24
Remote 1	72	72	72	48

Figure 2.4. Original risk matrix from the NCHRP 782 report (A) and the proposed risk matrix based on the results of the research (B).

This analysis of the new NBIS requirements and their intersection with practical application of risk models provides sound rationale for modifying the risk matrix originally proposed in NCHRP 782. The following changes to the original risk matrix are proposed as on outcome from the research:

- The matrix location [$R_{1,1}$] should be 72 months, since 96-month intervals are not allowable under the NBIS. (It should be noted that should the NBIS be modified in the future, this value could be replaced with 96 months with no negative impacts with respect to the calibration performed herein.)
- The position $R_{1,3}$ was originally indicated as a 48-month interval but is proposed as a 72-month interval to provide appropriate granularity to sort bridges into different "bins" in terms of risk. Using the proposed matrix, when the likelihood of serious damage is *remote* and the CF is *high*, a 72-month interval is allowable.

This revised matrix was used in the research and provided a practical solution when considering the results of back-casting. It should be noted that the risk matrix is subjective and based on engineering judgement. As a result, the weighting of the risk models that describe the OF and the categorization of attributes used to define the CF are interrelated with the definitions of the elements of the matrix. Assuming a bridge with very high ADT may have a CF of *high*, and that ADT alone should not preclude an extended interval, the element $R_{1,3}$ should be a 72-month interval. Otherwise, it may be necessary to assume a bridge with very high ADT fits a *moderate* CF, reducing the granularity of the risk matrix.

Table 2.4. Risk-based inspection intervals for components with CF of *high*.

CR	OF	Interval (Months)
CR 7	Remote	72
CR 7	Low	48
CR 6	Low	48
CR 6	Mod	24
CR 5	Low	48
CR 5	Mod	24
CR 5	High	24

This analysis and proposed risk matrix were used to provide "target ranges" for the analysis of the RAP-developed risk models. Although fixed values were not used explicitly, the target ranges for components were as follows:

- Most components rated in CR 7 have risk scores in the *remote* range for the OF.
- Most components rated in CR 6 have risk scores in the *low* or *moderate* range for the OF, indicating increased risk as compared with CR 7 components and decreased risk as compared with CR 5 components.
- Components rated in CR 5 present increased risk as compared with components rated in CR 6 with many having risk scores in the *moderate* to *high* range for the OF.

Here "most" was considered as being more than 60% and less than the mean values plus one standard deviation (σ), or ≈84%, of CR 7 components should have a remote likelihood of deteriorating to a CR 3 in the next 72 months. These quantitative proportions are subjective, but conservative, and align with the FHWA policy for Method 1. These target ranges are not intended to be defined limits but rather target ranges to provide a means of weighting individual attributes.

Another assumption of the research is that bridge components in CR 4 or lower are screened from the analysis. Components in "poor" condition (i.e., CR ≤ 4) have a maximum interval of 24 months according to the NBIS. Risk analysis could be used to identify bridges in this condition that require inspection intervals of less than 24 months. However, the criteria for attributes in the risk models described in this research are aimed at prioritizing bridges in fair to good condition. Different criteria and different attributes would be needed to prioritize bridges of CR 4 or lower in terms of risk. For example, most deck models may rate a CS attribute as *high* for a deck with more than 5% CS 3 damage. To apply the attribute to bridge components in poor condition the criteria would need to be adjusted. Many CR 4 components may have more than 5% CS 3 damage and rating them all as *high* may not produce any prioritization of the components. The criteria ranges would need to be increased to, for example, CS 3 greater than 25% is *high* relative to other CR 4 components. A separate risk analysis with suitable criteria for components in poor condition would be required to estimate a rational reduced inspection interval.

2.2. Nonredundant Steel Tension Members

The NBIS allows RBI to be applied to NSTMs using Method 2 in a similar manner as for routine inspections. A previous project in 2007 at Purdue University developed a methodology that was very similar to the RBI approach used in the NCHRP 782 report in terms of identifying key attributes and developing a semi-quantitative scoring process to determine the appropriate inspection interval for NSTM members(Parr, Connor, & Bowman, 2010). The updated NBIS allows for an inspection interval of 48 months for NSTMs that meet certain criteria such as being constructed after 1979, have no Category E or E’ details, and have no fatigue details with finite life, history of fatigue cracks, or pin and hanger details. NSTMs meeting these criteria are eligible for a 48-month inspection interval. For NSTMs not meeting one or more of the criteria, the Method 2 approach can be used. To help implement Method 2 for NSTMs, the methodology for analyzing NSTMs has been updated to have a similar scoring approach used in the present research. The updated methodology presents the results of a RAP meeting conducted focusing on NSTMs. The methodology is described in Appendix B, *NSTM Hands-On Inspection Interval Assessment*.

Chapter 3 Reliability Assessment Panel Results

Reliability Assessment Panel (RAP) meetings were conducted with six of the participating states. The RAP meeting's purpose was to develop risk models for RBI planning. The RAP meeting included discussions of the damage modes, attributes, and the consequences for the bridge family being studied. The families of bridges analyzed in each state are shown in Table 3.1. These bridge families were determined based on survey feedback from the participating states and the desire to have both steel and PSC bridges considered in the study.

Table 3.1. Table showing the family of bridges studied in each state.

State	Bridge Family	State	Bridge Family
Connecticut	Steel	Missouri	Steel
Idaho	PSC	Washington	PSC
Illinois	Steel	Wisconsin	Steel

The focus of the RAP meetings was damage modes and attributes needed for analysis to determine the OF as part of determining the inspection interval for routine inspection. The RAP panels commonly included between four and ten participants. A typical composition of panelist included the following:

- Bridge Inspection Expert.
- State NBIS Program Manager.
- Bridge Management Engineer.
- Bridge Maintenance Engineer.
- Materials Engineer.
- Structural Engineer.
- Independent Experts / Academics.

The RAP meetings were 1.5 days in length. The first day of the meeting was focused on identifying damage modes and attributes for the deck, superstructure, and substructure for the subject bridge family. The following day included a summary of the day one activities and discussion of the CF attributes. The following sections describe the preparation for the meeting, description of the activities conducted during the RAP meeting, and the results. RAP meeting results include the damage modes identified for different bridge families and the related attributes.

3.1. RAP Meeting Preparation

A document providing an overview of the RBI process and the RAP meeting objectives was provided to each participating state. Each state was asked to provide a list of RAP meeting participants that included data on the qualifications and experience of the RAP members. The participants were invited to attend a 1.5 hr. training webinar that provided an overview of the RBI process and the RAP meeting procedures for expert elicitation that were to be conducted during the meeting. This webinar set the stage for the in-person meeting by providing participants with a preview of what to expect during the meeting. The topics in the webinar were also repeated at the outset of the RAP meeting itself to reinforce the overall concepts of this new technology of risk assessment by an expert panel.

Workbooks were developed for use in the RAP meeting to support the expert elicitations conducted during the meeting. Each workbook included forms for completing exercises conducted during the meeting to identify the damage modes and attributes for bridge components (deck, superstructure, and substructure) and assessment of consequence scenarios. Damage mode worksheets were used to elicit input on the most likely damage modes for the component under consideration. Attribute worksheets were provided for recording design, loading, and condition attributes, accompanied by ranking of those attributes as *low*, *medium*, *high*, or *screening*. These worksheets were used to collect initial data on key attributes, which were then refined through panel discussion. The elicitations were conducted using a Delphi process as described in the following section.

The consequence scenarios consisted of example bridge images and details on number of spans, span length, material of construction, ADT, and damage mode. Each consequence scenario was supplemented with worksheets to record the expert judgment of the RAP participants regarding attributes to be considered in determining the CF. The workbooks had consequence scenario examples from the NCHRP 782 report to assist RAP participants in completing the consequence scenario worksheets and identifying which attributes should be assessed to determine the CF.

The research team (RT) facilitated the meeting, provided materials, and documented the results. The RT led the expert elicitation, prompted participants for input, and recorded the RAP inputs. Partner states made available key personnel to participate in the RAP meetings. Feedback from RAP participants regarding likelihoods, attributes, criteria, and consequences were recorded on easel pads or dry erase boards. Photographs of the recorded notes were taken to document the results. These notes were subsequently converted to flow charts to organize the damage modes, attributes, and criteria provided by the RAP into a logical framework for review.

The RAP meeting focused largely on the attributes for estimating the likelihood of serious damage developing in a bridge component over the next 72-month interval. Expert elicitation was used to gather data on key damage modes and attributes affecting those damage modes. Limited discussion of the CF was conducted during the RAP meeting due to time limitations.

3.1.1. Delphi Exercises

The Delphi process is a means of aggregating expert opinion through a series of structured questions intended to obtain expert knowledge in areas where available data is limited (Gunaydin, 2006; Kiral, Kural, & Çomu, 2014). The Delphi process consists of anonymous surveys of experts followed by consensus development to form an expert solution to the given problem. The goal of the Delphi process is to elicit expert judgements in an objective manner. The anonymous nature of the initial surveys is intended to avoid bias introduced by certain group dynamics, such as vocal or strongly opinionated participants dominating the discussions.

The process for a Delphi survey includes defining a problem, selecting suitable panel members, developing questions for experts to resolve, providing open-ended questions for the experts to provide anonymous input, and controlled assessment and feedback (Hohmann, Brand, Rossi, & Lubowitz, 2018). The assessment of feedback portion consists of aggregating the anonymous survey results and forming additional rounds of questions (or scenarios) that are presented to the experts in a group setting to develop consensus opinions or judgements from the initial anonymous results.

The Delphi process was applied to identify the damage modes and attributes for a given component. To identify damage modes, the anonymous survey seeks responses to the question *"If it was reported a [component] is rated as CR 3, based on your experience, what damage would be present in the [component] that has resulted in the low condition rating."* The "component" in the question is either the deck, superstructure, or substructure. This inquiry's objective is to identify the primary damage modes that are likely to lead to a component failure (i.e., CR 3) and identify damage modes which may occur but are unlikely. For the former, the primary damage modes are identified such that risk models can be developed to address the likelihood of that damage mode occurring and causing the component to deteriorate to CR 3. For the latter, the damage modes are considered as potential screening criteria or deemed insignificant and neglected. During the RAP meeting, members assessed this question individually to provide anonymous initial input for the survey. The results from each panelist were summarized by the facilitator and presented to the panel for discussion and consensus-building. The consensus process sought to determine if a risk model was needed for the damage mode, if the damage mode was suitable as a screening criterion, or if the damage mode was too rare to require consideration in the risk analysis.

The individual anonymous assessments for damage modes were conducted using a simple bubble sheet where participants were asked to provide an estimate of the likelihood for each damage mode to occur. An example portion of the bubble sheet is shown in Table 3.2, with only two rows shown to illustrate the appearance of the bubble sheets. The actual sheets provide the user with additional rows to record their input. The bubble sheet includes a space for the RAP member to input damage modes and circles representing different likelihoods or probabilities associated with the damage mode. The members were instructed to list the damage modes they believe are most likely to cause a given component to deteriorate to CR 3, shown as handwritten in Table 3.2. The members were then asked to prioritize the damage based on the likelihood of that damage mode being the cause of the deterioration. The RAP members used the bubble sheets to provide input on the likelihood for each damage mode and were instructed to provide estimates that sum to 100% likelihood in 10% intervals. This provided an independent assessment of the relative priority of the potential damage modes to identify those damage modes that are most likely and those damage modes that are unlikely. This exercise helped to identify common damage modes that are likely to occur and prevent bias toward damage modes that are rare or have occurred recently.

The survey results were then aggregated by listing each of the damage modes identified in the survey on a white board and recording the likelihood estimates provided by the RAP member for each damage mode. These data were then reviewed by the RAP and discussed to form a consensus on the most likely damage modes for the given component.

Table 3.2 Example of table provided to RAP members for assessing likely damage modes.

Damage Mode	**Likelihood (in 10% increments)**									
Delamination and spalling	○	○	○	○	○	●	○	○	○	○
Impact	○	○	○	●	○	○	○	○	○	○
Proportion	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

Once the damage modes were prioritized in this manner, the key attributes that impact the resistance of the component to the specified damage mode were addressed. For each of the key damage modes, the RAP members were presented with a second survey question to independently identify attributes. The questions posed to the panel were as follows: "*For the [component], estimate how long will it be before significant [damage mode, e.g., section loss] develops? What information do you need to know to make that estimate?*" This survey question intended to elicit objective input on the key attributes to be considered in the risk model. The question was formed in a manner that is expected to be familiar to engineers with experience in the condition assessment and maintenance of highway bridges. The question's objective was to identify attributes that significantly impact the development of the subject damage mode.

The RAP members were asked to rank the relative importance of those attributes as *high*, *moderate*, or *low* in terms of the attribute's impact on the damage mode developing and progressing to a serious condition. These data were used as a basis for developing consensus on the priority rank of the attributes through discussion with other RAP members. In this way, the most significant attributes were identified and provided a rank used to develop a scoring scheme in the risk model.

Once the attributes were ranked, group consensus-building was used to discuss and identify criteria for scoring each attribute. The criteria are intended to characterize the attribute for a specific bridge. For example, for the damage mode of section loss in steel girders, the attribute of coating condition may be identified as an attribute with a rank of *high*. If the coating is in good condition, then the coating condition is not increasing the likelihood of the underlying deterioration mechanism of corrosion developing and resulting in damage. If the coating is failing, exposing the steel to the corrosive environment, then it will have a significant impact on corrosion propagating and causing section loss. Criteria were developed to rate the attribute, e.g., if the coating is failing, the attribute will be rated *high*, and if the coating is in good condition, the attribute will be rated *low*. Specific criteria that can be obtained from available records are preferred, for example, the criterion for the coating attribute to be rated *high* is coating assessed to be in CS 4 or significant quantities in CS 3 by element-level inspection.

Following the RAP meeting, the criteria for each attribute and its relative rank were used to develop a simple scoring methodology to be used to rate a particular bridge or family of bridges to determine the OF.

3.2. Documentation and Initial Analysis of RAP Results

The damage modes, attributes, and criteria were captured by taking photos of the whiteboard or the easel pads used during the RAP meetings. These results were transcribed into flowcharts to illustrate the framework of the risk model. The flow charts show the damage modes, attributes and criteria for each damage mode identified by a RAP. The rank of each attribute is also listed in the risk flow charts. The flowcharts organize the RAP results in a systematic fashion and summarize the outcome of the RAP surveys.

A flow chart for bridge deck attributes collected during one of the RAP meetings is shown in Figure 3.1. The figure shows a portion of the flowchart for the damage mode of delamination and spalling in a bridge deck. The figure shows four of the attributes identified by the RAP and the rank assigned to each attribute. The attribute of "Current Condition Rating" and "Spalls and Patches" were each ranked as *high* (H) by the RAP. The attribute of "Rate of Salt Application" was ranked as *moderate* (M), and the attribute of "Corrosion Protection Layers" was ranked *low* (L). Criteria for each of these attributes were also identified as shown in the figure. For the attribute of "Current Condition Rating," the criteria were described in terms of the CR for the deck. For the attribute of "Spalling and Patches," the criteria for the attribute were described in terms of the element CS from element-level inspection. The CS for this element considers the presence of spalling, patches, and cracking in the deck and acts as a surrogate for the "Spalls and Patches" attribute. In this way, the available data from element-level inspection can be used to support a data-driven process. Although the CS is different from the RAP input, it includes the RAP input and was deemed a suitable replacement. In a similar way, the rate of salt application, which is rarely data that is available state-wide, is expressed in terms of available data such as ADT values and functional class of the roadways that commonly receive more (or less) aggressive deicing treatments.

3.2.1. Data Analysis Tools – Excel Spreadsheets

The data from the RAP meeting was used to develop spreadsheets to aid in the scoring of individual bridges using the risk models. The spreadsheet's objective was to provide a simple data entry tool for a user to assess risk for a particular bridge. The spreadsheet was used for calculating the OF for the risk analysis for each bridge component. The data collected during the RAP meeting was scored based on the rankings assigned by the RAP participants for attributes. In this way, relative values for each attribute are assigned based on the RAP input and criteria selected to assign specific values for each attribute. The scoring process could be implemented though asset management software that includes data on many of the attributes and data-driven criteria for assessing the attributes.

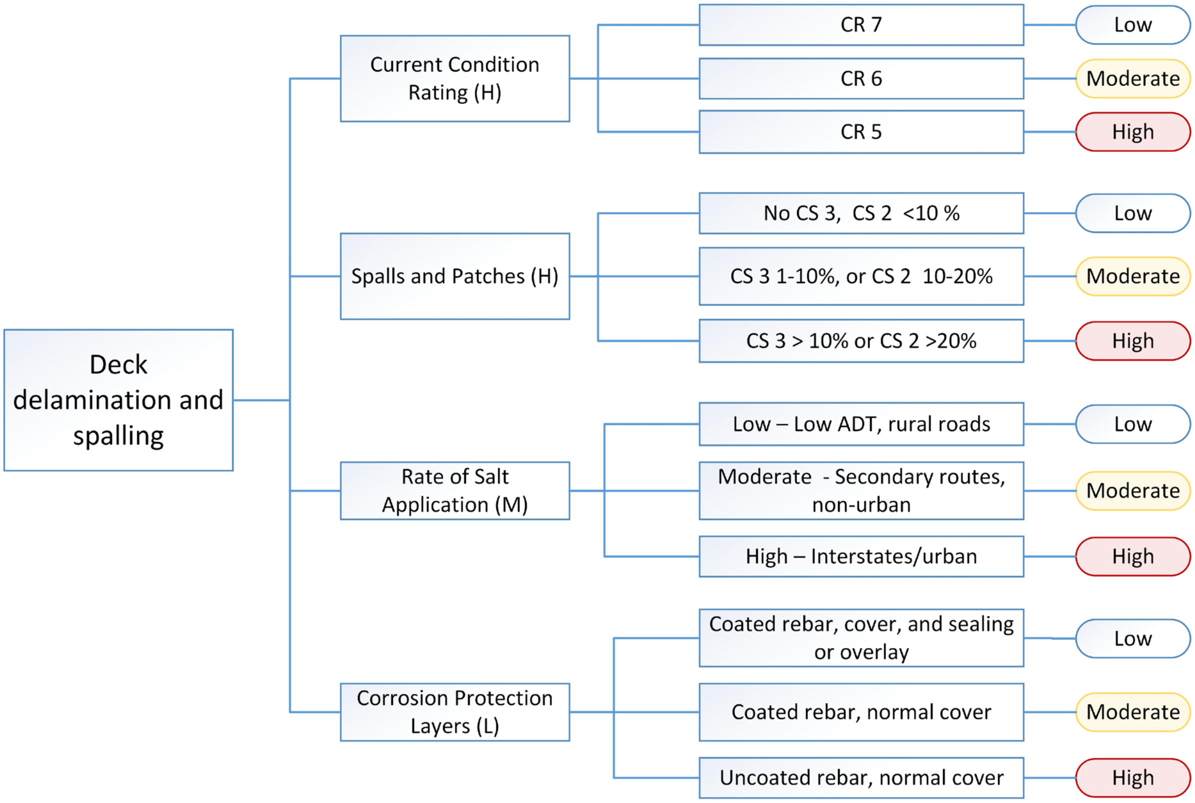


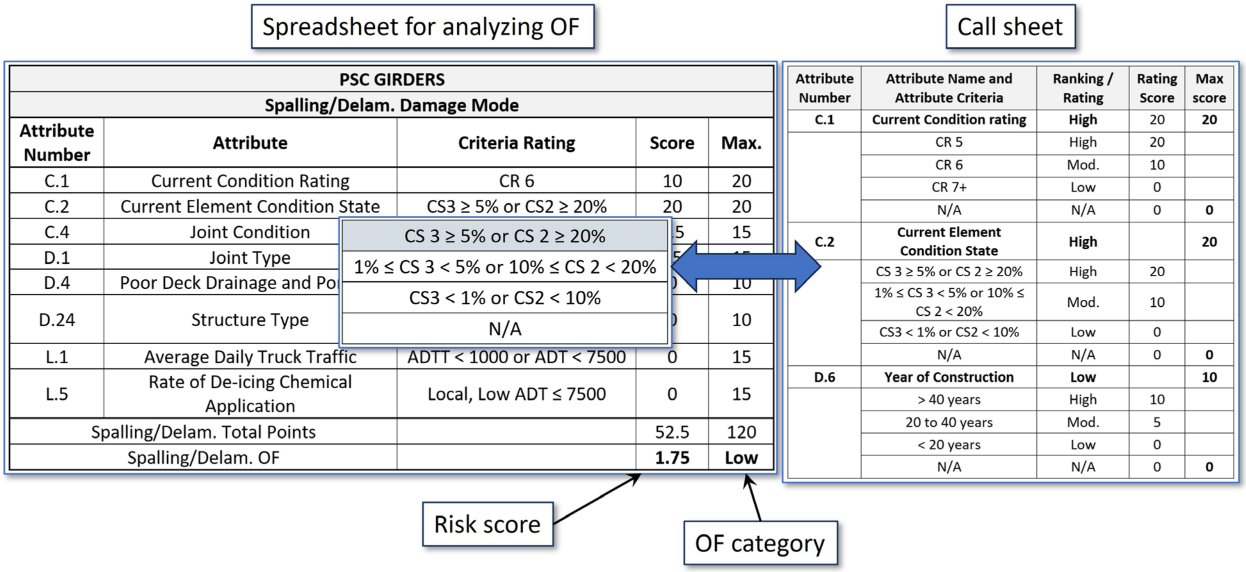
Figure 3.1. Typical flow chart for RBI.

Figure 3.2 illustrates the functionality of the risk analysis spreadsheet. The figure illustrates the damage mode of delamination and spalling for a PSC superstructure. Each attribute identified by the RAP for this damage mode is shown, and the criterion for each attribute is assigned based on the specific bridge being analyzed. A drop-down menu is used to select the specific value for the criteria for the bridge being assessed. The drop-down menu calls values from an associated listing on a call sheet for each of the criteria for that attribute. Once values are selected for each attribute, the risk score for the OF for that component and damage mode is automatically calculated. In this way, individual bridges can be rapidly scored to determine the risk profile for the bridge. The spreadsheet application was developed for each RAP outcome to support risk analysis of the sample bridges used for back-casting.

3.3. Reliability Assessment Panel Results

This section presents the overall results from RAP meetings held as part of the research. The objective of this section of the report is to summarize the primary damage modes and attributes identified by the RAP meetings conducted under the research.

Four RAPs reported damage modes for steel girders and two RAPs reported damage modes for prestressed concrete girders. All six RAPs reported damage modes for deck and substructure components. The following section discusses the damage modes identified by the RAPs and summarizes the common attributes identified for the different damage modes. The damage modes and attributes for bridge decks are presented first, followed by PSC and Steel superstructures, and, finally, reinforced concrete (R/C) substructures.



Spreadsheet for analyzing OF

PSC GIRDERS				
Spalling/Delam. Damage Mode				
Attribute Number	Attribute	Criteria Rating	Score	Max.
C.1	Current Condition Rating	CR 6	10	20
C.2	Current Element Condition State	CS3 ≥ 5% or CS2 ≥ 20%	20	20
C.4	Joint Condition		.5	15
D.1	Joint Type			
D.4	Poor Deck Drainage and Po			10
D.24	Structure Type			10
L.1	Average Daily Truck Traffic	ADTT < 1000 or ADT < 7500	0	15
L.5	Rate of De-icing Chemical Application	Local, Low ADT ≤ 7500	0	15
Spalling/Delam. Total Points			52.5	120
Spalling/Delam. OF			**1.75**	**Low**

CS 3 ≥ 5% or CS 2 ≥ 20%
1% ≤ CS 3 < 5% or 10% ≤ CS 2 < 20%
CS3 < 1% or CS2 < 10%
N/A

Call sheet

Attribute Number	Attribute Name and Attribute Criteria	Ranking / Rating	Rating Score	Max score
C.1	**Current Condition rating**	**High**	20	**20**
	CR 5	High	20	
	CR 6	Mod.	10	
	CR 7+	Low	0	
	N/A	N/A	0	**0**
C.2	**Current Element Condition State**	**High**		**20**
	CS 3 ≥ 5% or CS 2 ≥ 20%	High	20	
	1% ≤ CS 3 < 5% or 10% ≤ CS 2 < 20%	Mod.	10	
	CS3 < 1% or CS2 < 10%	Low	0	
	N/A	N/A	0	**0**
D.6	**Year of Construction**	**Low**		**10**
	> 40 years	High	10	
	20 to 40 years	Mod.	5	
	< 20 years	Low	0	
	N/A	N/A	0	**0**

Risk score

OF category

Figure 3.2. Example of the spreadsheet scoring process for risk analysis.

3.3.1. Concrete Bridge Decks

There were six different damage modes identified for R/C bridge decks as shown in Table 3.3. The damage modes identified included delamination and spalling, cracking, reinforcing steel section loss, and exposed rebar. Wear, abrasion, or rutting was identified by a few RAPs, although the likelihood that this damage mode was the cause of a deck being rated in CR 3 was low. Several RAPs also identified soffit damage such as map cracking, saturation, and efflorescence.

To illustrate the results for the RAP Delphi surveys, the combined results from all six RAPs are shown in Figure 3.3 along with abbreviations used for the different damage modes. These data show composite results as a percentage of all RAP members indicating a particular range of likelihoods for each damage mode. For example, the figure shows that for the damage mode of delamination and spalling (DL / SP), 26% of RAP participants estimated the likelihood for delamination and spalling as between and 10% and 30%, 43% of RAP participants estimated the likelihood between 40% and 50%, and 31% indicated the likelihood as between 60% and 80%. As shown in the figure, only the damage modes of delamination and spalling, cracking (Crk), and soffit map cracking / saturation / efflorescence (Sat.) were assigned a substantial likelihood of resulting in a deck being rated in CR 3. Other damage modes identified by the RAPs included exposed reinforcing (Exp. R), section loss of reinforcement (Sect. L), and wear, abrasion, or rutting (W / A / R). These damage modes were assessed as being relatively unlikely causes of serious damage, less than 30%, and two of these damage modes (exposed reinforcing and section loss) are closely related to corrosion damage and can be addressed as part of a delamination and spalling damage mode.

Table 3.3. Summary of damage modes for R/C Bridge decks.

Damage Mode	Abbreviation
Delamination / spalling	DL / SP
Cracking	Crk
Section loss	Sect. L
Exposed reinforcing	Exp. R
Wear, abrasion, or rutting	W / A / R
Soffit map cracking / saturation / efflorescence	Sat.

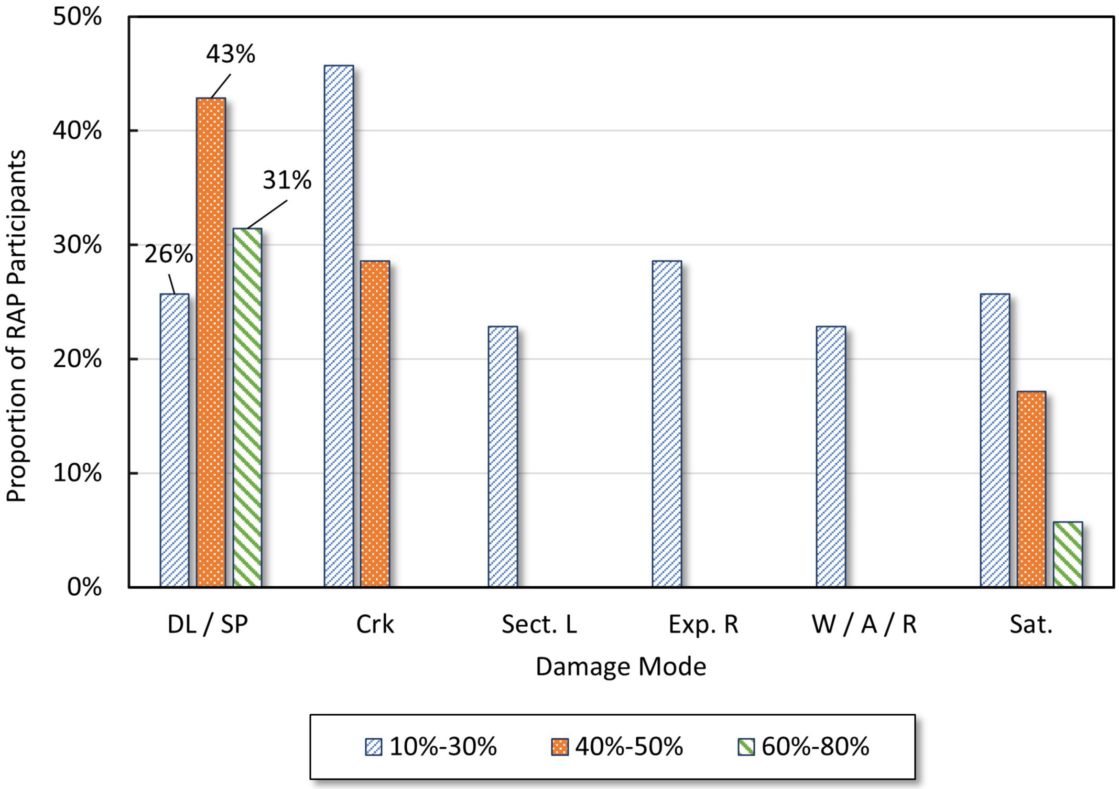


Figure 3.3 Damage modes for R/C bridge decks identified by the RAPs.

In all cases, the damage mode of spalling / delamination was identified as the most significant damage mode and, therefore, was addressed as the primary damage mode for decks. Cracking was another common damage mode identified by each of the RAPs, but this damage mode is commonly associated with corrosion damage that results in delamination and spalling. For instance, a bridge deck with cracking is not usually repaired until associated spalling occurs, although preservation strategies like crack sealing may be employed. Therefore, cracking is an attribute that contributes to the damage mode of delamination and spalling, and not treated as a separate damage mode. Saturation and map cracking of the deck soffit was also identified. Saturation of the deck is a more advanced form of corrosion damage and has the same attributes in terms of deck durability and resistance to corrosion damage as the delamination and spalling damage mode and, therefore, was included in the delamination and spalling damage mode.

3.3.1.1. Delamination and Spalling Damage Mode for R/C Decks

The most common damage mode for bridge decks was delamination and spalling, and there was some consensus among the different RAPs regarding attributes that affected the reliability of bridge decks. Table 3.4 lists the common attributes for R/C decks that were identified by the RAPs, listed in order of how commonly the attribute was identified. Attributes identified by at least 5 of the six RAPs are shown in column 1 on the table, with less commonly identified attributes listed in columns 2 - 4. Columns 2 and 3 both show attributes identified by either two or three RAPs.

Rate of deicing chemical application, CR, and reinforcement type (i.e., ECR, stainless steel, galvanized, or uncoated) were identified by all RAPs. All RAPs identified traffic volume on the deck as an important attribute, with five of the RAPs indicating the attribute as ADTT and one indicating ADT. Five of the six RAPs indicated that overlays or sealers applied to the deck for corrosion protection were an important attribute. Attributes identified by two or three of the RAPs included general cracking, concrete cover, presence of spalled or patched areas, and soffit damage. Frequent exposure to overweight or permit loads was also identified. Concrete mix design and poor deck drainage or ponding were also included.

Several attributes describe specific damage present in a bridge deck such as cracking, soffit damage, and spalling and patches in the deck. These forms of damage would normally be included in an element-level inspection that indicated the quantities of damage in CS 1, CS 2, or CS 3. These attributes were summarized by considering the CS of the deck element (element 12) in the risk models. If element-level inspection is not being used, these attributes would need to be captured in the inspection procedures.

Table 3.4 Listing of common attributes identified for R/C decks.

Attributes Identified by 5 or 6 RAPs	Attributes Identified by 2-3 RAPs	Attributes Identified by 2-3 RAPs	Attributes Identified by 1 RAP
Rate of Deicing Chemical Application	General Cracking	Age	Maintenance Cycle
CR	Concrete Cover	Concrete Mix Design	No. of Freeze-Thaw Cycles
ADT or ADTT	Presence of Spalling or Repaired Areas	Poor Deck Drainage or Ponding	Dynamic Loading
Reinforcement Type	Soffit Damage	-	Rate of Deterioration
Deck Overlay or Sealers	Overweight or Permit Loading	-	Nondestructive Testing

There were several attributes related to the corrosion protection for the deck. These included the rebar type, concrete cover, and overlays and sealers. Each of these attributes improves, in most cases, the corrosion protection level of the deck. An exception is asphalt overlays applied to extend the service life of a deck that has been affected by significant spalling and patching. Such an overlay may not increase the corrosion protection of the deck, but the deck was already in poor condition prior to the asphalt overlay being installed and is not likely to be considered for an extended interval.

There are many types of overlays, sealers, and reinforcing steel coatings. To simplify the analysis, these attributes were used to form a corrosion protection (CP) level attribute. Modeled after the CP levels commonly used for post-tensioned concrete construction, the CP level provides a simple method of describing resistance to corrosion damage based on the reinforcing steel coating type, the depth of cover, and the protection offered by overlays and sealers (PTI/ASBI, 2019).

It should be noted that overlays, in this case, describe overlays that provide additional corrosion protection. An overlay that provides normal concrete cover following a milling or hydro-demolition activity would not typically count as an additional layer of protection. However, if the overlay material is a high-performance material that provides improved corrosion protection as compared with typical concrete mixes, a user may consider the overlay to provide an additional CP level. For asphalt overlays placed on sound concrete decks, the permeability of the asphalt should be considered to determine if the asphalt forms a layer of protection.

To determine the CP level, the user simply counts the number of protective layers between the surface of the element and the embedded reinforcing steel. Epoxy coating, normal concrete cover, overlays placed on existing concrete cover, and sealers all represent one layer of corrosion protection. A membrane under an asphalt overlay would be considered one layer of protection. The use of reinforcing bars that are stainless steel (SS), clad SS, galvanizing coating, or fiber reinforced polymer bars are counted as two levels of corrosion protection based on the increased resistance to corrosion of these materials.

Table 3.5 shows different corrosion protection combinations that fit into the CP levels of 1 through 4. The least amount of corrosion protection is provided by CP 1, which has either 1 or 0 levels of corrosion protection. For example, a bare concrete deck that is unsealed, has uncoated reinforcing bars and normal cover has one layer of corrosion protection (cover) and ranks in the highest category. If the deck with uncoated rebar also has low cover, it would have zero levels of protection, but is scored as CP1. This is rational since a concrete deck with low cover or uncoated rebar would be particularly susceptible to corrosion damage and would likely deteriorate rapidly in an aggressive environment, affecting the condition attribute and increasing the risk score. If the environment is inert and no damage is developing, the effect of corrosion protection is reduced, making the distinction mute. A bare deck with normal cover and EPC reinforcing steel, but without sealing or an overlay, would have two layers of protection, rated as CP2 or *high*.

Additional attributes described by the RAPs included the quality of the drainage (e.g., ponding due to poor drainage). This attribute is intended to represent the increased exposure to moisture and deicing chemicals that results from water not draining properly from the deck.

There was less consensus among other attributes related to decks. Two RAPs indicated the concrete mix design or type has an impact, but for distinct reasons. In one case, the rationale for considering the concrete mix was to identify a geographic location in the state with a history of poor-quality aggregates that resulted in less durable concrete as compared with other locations in the state. The other RAP considered that certain concrete mixes previously used in the state did not have good durability and would increase the likelihood of corrosion damage in the deck.

Other attributes related to the condition of the deck but less frequently cited included the rate of deterioration for the deck, i.e., how rapidly the deck had deteriorated to its current condition, the number of freeze-thaw cycles to which the deck was subjected, and the condition of adjacent deck and header elements. The flexibility of the superstructure was identified based on the experience that decks on bridges with larger deflections under traffic tended to deteriorate more rapidly than decks on bridges with more typical deflection characteristics. An attribute of impact / dynamic loading, referencing the “bump at the end of the bridge” was identified as a driver for more rapid deck deterioration. Two RAPs indicated that overloads on the deck was a consideration in predicting the future evolution of damage because the overload increased the rate of deterioration.

Availability of NDT data for a deck was identified by one of the RAPs. This attribute was identified because NDT data could indicate additional damage not identified using visual inspection or could verify the absence of subsurface deterioration. In either case, applying NDT would reduce the uncertainty of the deck's condition assessment.

Table 3.5. Examples of CP levels for R/C components.

CP level	Rank	Description	Description
CP 1	Very High	The element has either one layer of protection or no layers of protection. One layer of protection may be from the following: • Cover of at least 2 in. • Epoxy coated reinforcing steel. • An overlay applied for corrosion protection. • Effective sealing practice.	Little to no corrosion protection
CP 2	High	The element has two layers of protection: • Typical cover of at least two inches and ECR. • Typical cover, black rebar, and an overlay or sealer. • Low cover with EPC and an overlay or sealer.	Nominal corrosion protection
CP 3	Mod.	The element has three layers of corrosion protection: • Typical cover, ECR, and an overlay <u>or</u> sealer. • Typical cover and galvanized, SS, or FRP rebar.	Improved corrosion protection
CP 4	Low	The element has four layers of protection: • Typical cover, ECR, overlay <u>and</u> sealer applied. • Typical cover, galvanized, SS, or FRP rebar, and an overlay or sealer.	High level of corrosion protection

3.3.2. Steel Bridges

There were four RAPs that assessed the damage modes and attributes for bridges with steel superstructures. The Delphi process was used to elicit expert judgement on the damage modes for steel bridge members. There was a consensus among the four RAPs that the primary damage modes for steel bridges were corrosion damage (i.e., section loss due to corrosion), fatigue cracking, impact damage from over-height vehicles, and collision from debris in a waterway. Three RAPs mentioned overload that results in sagging or member deformation in the listing of damage modes, but this was not deemed sufficiently widespread to be considered. Other damage modes identified by the RAPs included connection damage, fabrication errors, and girder movement or bearing failures. Connection damage and fabrication errors were addressed with the risk models for fatigue cracking. Bearing failures referred to tilting rocker bearings and could be addressed as a screening criterion based on the CR or CS of the bearings.

The following section describes the attributes identified for each of the primary damage modes.

3.3.2.1. Corrosion / Section Loss

Loss of section due to the accumulation of corrosion damage was identified by all four RAPs as the most common damage mode leading to a steel bridge component being rated in CR 3. There was substantial agreement on some of the leading attributes associated with corrosion damage / section loss, as shown in Table 3.6.

Attributes were sometimes represented in slightly different terms during a RAP meeting. For example, three states identified “NaCl” (i.e., salt), two identified “NaCL application,” and one state described “ADT/salt usage” to express the increased likelihood of corrosion damage resulting from exposure to deicing chemicals. Attributes being worded slightly differently by different panels, and even different panel members within the same RAP, was quite common, and these data were converted to uniform language to match the attributes suggested in the RAP meeting to existing or new attribute definitions. For example, the attribute of deicing chemical application is described in the NCHRP 782 report as a Loading Attribute L.5, “Rate of Deicing Chemical Application.” These attributes are presented using standardized language when possible.

Table 3.6. Summary of attributes identified for corrosion damage in steel bridge members.

Attributes identified by 4 RAPs	Attributes identified by 2 or 3 RAPs	Attributes identified by 1 RAP
Rate of Deicing Chemical Application	CR	Embedded Girder Ends
Coating Condition	Local Environment (i.e., near water or exposed to F/T cycles)	Built-up Members
Joint Condition	-	Details Prone to Collect Water
Poor Deck Drainage or Ponding	-	Maintenance Cycle
Subjected to Overspray	-	ADTT

The attributes included the application of deicing chemicals, the condition of the coating intended to protect the member from corrosion, and the leaking of expansion joints. The attribute of deck drainage was also identified by each of the RAPs, sometimes including the type or condition of the deck supported by the steel members. Various descriptions of splash zones were also included by each state, indicating an increase in exposure to moisture and deicing chemicals typical of bridge members with low clearance over highway traffic or a waterway.

The current condition of the steel member was not included by all states in the initial assessment of the attributes. The current condition may simply be overlooked during the RAP as an obvious attribute that represents the culmination of the other attributes. Most of the risk models developed through the research included the CR of the component as an attribute.

There were some attributes that were identified during initial discussions but later disregarded for various reasons. For example, the deck condition was identified by one state based on the concept that a deck in very poor condition may allow additional drainage onto the superstructure members. But a deck in such poor condition is unlikely to be considered for RBI, so this attribute may be unnecessary for a risk model developed under Method 2 for extended intervals. Other attributes were considered to have too small of an influence to be included or were combined with other related attributes.

3.3.2.2. Fatigue Cracking Damage Mode

Fatigue cracking was the second most likely damage mode identified by the RAPs. There was commonality in many of the key attributes that affect the likelihood of fatigue cracking as shown in Table 3.7. It is noted that the likelihood of fatigue cracking was considered much lower than section loss. The top four attributes were common to all RAPs that analyzed steel superstructures.

There was a limited number of attributes identified by the RAPs for fatigue cracking. A damage mode with so few attributes affecting its likelihood of occurrence can be formed into a categorical risk model that determines the OF rating directly from the characteristics of the component rather than through a scoring process. A categorical risk model for fatigue cracking is provided later in the report.

Table 3.7 Listing of attributes for fatigue cracking damage mode.

Attributes identified by 4 RAPs	Attributes identified by 2 or 3 RAPs	Attributes identified by 1 RAP
Worst Fatigue Detail Categories	Frequent Permit Loads	Connection Damage
ADT / ADTT	Impact / Collision/ Fire	Web Gap Details
Age / Year of Construction	Welding Defects or Plug Welds	Field Welding
Section Loss, Previous Cracking, or Current CR	-	-

3.3.2.3. Impact / Collision / Fire

All RAPs identified impact as a damage mode for steel bridge superstructures. Impact damage results from a random event of an over-height vehicle impacting the bridge superstructure. Therefore, there is not a deterioration mechanism associated with impact damage. The likelihood of impact damage is also unrelated to the inspection interval or condition of the bridge. Therefore, there are only a few attributes typically associated with impact damage, as shown in Table 3.8. The attributes are related to the exposure of a bridge to the potential for impact damage. Bridges that are not over roadways or have very high vertical clearance are unlikely to be exposed to impact by vehicles. The traffic volume on the feature under the bridge will rationally affect the likelihood of an over-height vehicle traveling under the bridge and potentially impacting the superstructure. If the feature under the bridge is a low-volume rural road with low ADTT, it is less likely that an over-height truck will be travel under the bridge as compared to a highway with high ADTT. Finally, if a bridge has been impacted by a vehicle it is more likely to be impacted again as compared with a bridge that has never been impacted. Therefore, the attributes associated with impact were identified as vertical clearance, ADT / ADTT on the roadway below the bridge, and previous impacts.

Table 3.8. Listing of attributes for the damage mode of impact damage for steel bridges.

Attributes
Vertical Clearance
ADT / ADTT
Previously Impacted

Collision refers to debris or marine vehicles impacting a bridge superstructure or substructure, and therefore requires that the bridge be over water to be exposed. The attributes associated with this damage mode were similar to the attributes for impact, focused on the feature under the bridge being a waterway, vertical clearance, and previous history of collision damage from debris. Fire is a random event that cannot really be addressed through an inspection program in a general sense, although post – event inspections are sometimes required to address such an event.

3.3.3. Prestressed Concrete Bridges

Two RAPs addressed PSC girder bridges and identified different damage modes for PSC superstructures. Results from an additional RAP meeting in Missouri that was held prior to the research's initiation are included in the results reported here to provide broader results. The primary damage modes identified were delamination and spalling, exposed strands, cracking due to overload, and strand damage resulting from impact, as shown in Table 3.9.

There were other damage modes discussed during the RAP meetings. Strand corrosion was identified as a damage mode, but based on discussion, it was included within the delamination and spalling damage mode. Bearing loss or beam end damage was identified by one RAP and was included in the delamination and spalling damage mode. The rationale for combining these damage modes was that the damage modes have a common deterioration mechanism of corrosion. The attributes that indicate the likelihood of corrosion damage would be like the attributes for strand corrosion, bearing loss, or beam end damage.

However, strand corrosion can also result from exposure of the strand because of impact damage within the span, which may also result in mechanical damage (broken or mechanically damaged strands). For this reason, corrosion of strands exposed by impact damage was addressed as the likelihood of impact damage. It may be appropriate to consider exposed strand as representing an increase in the likelihood of corrosion damage (i.e., an attribute) considered as part of the delamination and spalling damage mode. This damage mode also could be treated as a screening attribute since a strand exposed to the ambient environment would have an unusually high likelihood of deterioration as compared with a strand with appropriate concrete cover.

Table 3.9 Listing of damage modes identified by the RAPs.

Damage Modes Identified by the RAPs
Delamination / Spalling
Exposed Strand
Cracking Due to Overload
Impact

The RAPs also indicated that shear and flexural cracking due to overload were potential damage modes for PSC superstructures. However, this damage mode was identified as a significant but unlikely damage mode for most bridges. Therefore, shear and flexural cracking were utilized as screening criteria.

Based on the primary damage modes indicated during the RAP meetings, attributes for delamination and spalling and impact are described in the following sections.

3.3.3.1. Delamination and Spalling Damage Mode for PSC Superstructures

All RAPs identified delamination and spalling resulting from corrosion as the primary damage mode affecting prestressed concrete bridges. One RAP identified beam end corrosion and this damage mode was combined with delamination and spalling as previously discussed. The attributes associated with corrosion-induced damage were similar in nature to those identified for corrosion damage in steel bridges. The attributes included the application of deicing chemicals, condition of joints, and the current condition of the member. The RAPs identified construction defects or damage as a possible source of delamination or spalling. There were several other attributes that were identified by each RAP as shown in Table 3.10.

Table 3.10. Listing of attributes for PSC member delamination and spalling damage mode.

Attributes Identified by 3 RAPs	Attributes Identified by 2 RAPs	Attributes Identified by 1 RAP
Rate of Deicing Chemical Application	ADT / ADTT	Deck Condition
Current Condition (CR or CS)	-	Overspray
Joint Condition	-	Initial Construction Damage
-	-	Girder Type

3.3.3.2. Impact Damage

All three states identified impact damage as a damage mode for PSC superstructures. The attributes associated with the likelihood of impact damage were limited to vertical clearance, ADT / ADTT on the roadway under the bridge, and if a bridge has been previously impacted, as discussed above, and listed in Table 3.8.

3.3.4. Substructure

Analysis of substructure damage modes was not fully completed by all RAPs due to time constraints. Because the damage modes for substructure have commonality with damage modes for other components, much of the work of identifying attributes for substructures was already completed during the superstructure and deck analysis, so priority was not placed on analysis of substructure components.

For the RAPs that completed analysis of substructure components, delamination and spalling was identified as the primary damage mode that would affect substructure elements and result in a CR of 3. Settlement or substructure movement, section loss for elements such as exposed timber or steel piles, and erosion / undermining of the substructure were also identified as potential damage modes. Cracking due to shear or settlement, and impact damage from either water-born vehicles or traffic, were also suggested by the RAPs.

3.3.4.1. Delamination and Spalling Damage Mode for Substructures

The damage modes and attributes for substructure elements were consistent with other components such as R/C decks or PSC superstructures. Table 3.11 lists the attributes identified for the damage mode of delamination and spalling for substructures. As shown in the table, the attributes for delamination and spalling for a substructure are similar to attributes for corrosion damage in PSC superstructures and R/C decks. The joint condition plays a key role because damaged and leaking joints allow the exposure of the substructure to run-off containing deicing chemicals from the deck. The condition and drainage effectiveness of the deck was also considered because poorly drained decks can result in elevated exposure of the substructure to deicing chemicals that cause corrosion.

Table 3.11 Summary of attributes for delamination and spalling of substructures.

Attributes Identified by 3 or 4 RAPs	Attributes Identified by 2 RAPs	Attributes Identified by 1 RAP
Joint Condition	Reinforcing Bar Type	Low Cover
Salt spray, Overspray, Splash Zone	Rate of Deicing Chemical Application	Coating Condition
Poor Deck Drainage	Poor Construction Quality	Debris Damage
-	Ambient Environment / Soil Corrosivity	Protective Coatings

Exposed timber or steel piles were also identified by the RAPs as an attribute that affects the durability of substructures. Because exposed piles formed from timber or steel may not match the deterioration characteristics of R/C substructure elements, it was decided that exposed timber or steel piles should be treated as separate components from R/C substructures.

3.3.4.2. Cracking

Two RAPs identified cracking of the substructure as a damage mode that should be considered in the risk model. Cracking in substructure components can be caused by differential settlement. One RAP considered that the attributes for cracking due to corrosion damage were similar to those attributes already considered for delamination and spalling. A second RAP indicated the attributes shown in Table 3.12. The attributes identified included settlement of the substructure and cracking resulting from consolidation cracking of mass concrete. Three attributes associated with the likelihood of cracking due to corrosion damage were also identified, including the bar type, if there was a joint above the substructure resulting in deicing chemicals from the deck draining onto the substructure, and chlorides that may be present due to deicing chemicals applied to the roadway on the deck above the substructure.

Table 3.12. Listing of attributes for substructure cracking.

Attribute
Settlement
Mass Concrete Consolidation Cracking
Bar Type
Joint Condition
Chlorides

3.3.4.3. Impact Damage for Substructure Components

Three RAPs that addressed substructure damage modes identified impact damage for substructure components. Impact damage would apply to vehicle impacts on substructures supporting overpass bridges, and barge or marine vehicle impact damage for bridges over waterways. Attributes for impact damage were only identified by one RAP. The attributes identified for impact damage generally considered the feature under the bridge, the horizontal clearance from the roadway under the bridge, and if the substructure element was protected by a roadside safety feature (e.g., guard rail or crash attenuator). Impact damage was also considered when the bridge was over a navigable waterway, such as the Mississippi or Illinois Rivers.

3.3.5. Criteria for Attributes

Criteria are used for rating attributes to assign point values. Attributes are rated on a qualitative scale of *high*, *moderate*, or *low*. For example, for the attribute of joint condition, a leaking joint would be rated as *high*, a joint that is not leaking would be rated *low*, and a joint that was leaking to some extent would be rated as *moderate*. An attribute that is rated as *high* based on its criteria is assigned 100% of its weight. If the attribute is rated as *moderate* the attribute is assigned 50% of its weight, and if the attribute is rated *low,* it is assigned zero points. For an attribute ranked *high* by the RAP, a 20-point scale is used. An attribute rated *high* is assigned 20 points, *moderate* is assigned 10 points, and *low* is assigned zero points. Certain attributes may be rated on a qualitative scale that includes four levels of *very high*, *high*, *moderate*, and *low*, based on the assessment by the RAP. Attributes described by four levels are typically assigned points of 100%, 50%, 25% and 0%, respectively. Different point distributions may be used if needed to express the impact of the attribute based on expert judgement. The specific criteria for each attribute may differ between bridge owners due to differences in environment, materials, construction practices, past experiences, etc.

The six RAPs defined attribute criteria for many of the attributes described in the previous section of this report. The attribute criteria identified were documented following the meeting using flow charts as previously mentioned, and then reduced into spreadsheet form. RAP criteria were modified as needed to link them to data available in existing records such as SNBI data or element-level inspection results. Some attributes were not fully characterized during the RAP meeting and were developed as part of the research.

3.3.6. Example Risk Models

This section of the report presents two example risk models developed by RAPs during the research to illustrate how the RAP input was used to form a risk model. The section includes example models for delamination and spalling of an R/C deck and fatigue cracking for steel superstructures. A full example of applying the risk models to a bridge is provided in Appendix A, *Handbook for Implementation of Risk Based Inspection.* Additional examples of the risk models developed through the RAP process are included in Appendix C, *Risk Models from the Research.*

The attributes identified by the RAPs were converted to standard language provided in the NCHRP 782 report, where possible, when the risk models were finalized during research. Several attributes that were not part of the original NCHRP 782 report were identified. Commentary that describes the rationale for these new attributes and example criteria and scoring were developed and combined with the attribute commentary in the NCHRP 782 report. The combined listing of the attributes and commentary are provided in Appendix D*, Attribute Index and Commentary.*

The risk models developed by the RAPs were applied to 10 sample bridges from each state, as will be discussed in Chapter 4. Methods of calibrating the scoring of the risk models were also developed during the research and are described in Chapter 5.

3.3.6.1. Example Deck Model

Table 3.13 shows an example risk model for an R/C bridge deck. The table includes a code used to describe the attribute, such as C.1, C.2, etc. which identifies the type of attribute (C = condition attribute, L = loading attribute, or D = design attribute), the name of the attribute, and the rank for that attribute identified by the RAP. The attribute's rank defines the total number of points assigned to it, with attributes ranked high assigned 20 points, moderate 15 points, and low 10 points. The criteria for rating an attribute as *very high*, *high*, *moderate*, or *low* are also shown. As shown in the table, most attribute criteria include three levels of *high*, *moderate*, or *low.* The only attribute with four levels of criteria is attribute D.26, Corrosion Protection Level. This attribute was developed through research, summarizing different attributes such as the concrete cover, reinforcing bar coating, overlays, sealers, etc. that provide corrosion protection into a single attribute, as previously described.

As shown in the table, most of the attributes can be rated using available data from routine or element-level inspections, or inventory data (e.g., ADT). For example, the CR and CS attributes are rated using available component and element-level inspection results, respectively. The ADT and ADTT attributes are rated based on vehicles per day (vpd) and trucks per day (tpd) data, respectively, available from bridge inventory data (i.e., Coding Guide or SNBI). Certain attributes, such as the corrosion protection level, may require expert judgment to rate the attribute.

3.3.6.2. Example Steel Superstructure Risk Model

This section presents an example risk model for fatigue cracking in steel superstructure members. As noted previously, there was consensus among the RAPs that the primary damage modes for steel superstructures were corrosion damage / section loss, fatigue cracking, and impact damage. The risk model for the damage mode of corrosion damage / section loss included several attributes such as the CR and CS, rate of deicing chemical application, joint condition, and being subjected to overspray. A risk model for corrosion damage in steel members is presented in detail in Chapter 5.

An example risk model for fatigue cracking is shown in Table 3.14. The model includes the CR and the element CS specifically for the defect element (DE) of cracking (DE 1010, Cracking (AASHTO, 2019)). The traffic volume attribute initially selected by the RAP set the *high* threshold at 10,000 vpd, and *low* at less than 1,000 vpd. The fatigue category of the details in the bridge was assessed with Category E and E' details being screened from the process, and Category D details rated *high*. The bridge's design era was assessed according to changes in the AASHTO fracture control plan, and an attribute was identified for welded attachments. A detailed analysis of this risk model is shown later in Chapter 5 of the report.

Table 3.13. Example risk model for a R/C deck showing the attribute, attribute rank, and criteria for scoring the attribute.

Code	Attribute	Rank	Criteria	Rating
C.1	Current CR	High	CR 5 CR 6 CR ≥ 7	High Moderate Low
C.2	Element Condition State	High	CS 3 ≥ 5% or CS 2 ≥ 20% 1%≤ CS3 < 5% or 10% ≤ CS2 < 20% CS 3 < 1% or CS 2 < 10%	High Moderate Low
C.4	Joint Condition	Moderate	≥ 20% CS 3 / CS 4 1% ≤ CS 3 / CS 4 < 20% CS 1 or CS 2, CS 3 < 1%	High Moderate Low
C.13	Efflorescence/Staining	Low	CS 3 ≥ 20% or CS 2 ≥ 20% 1% ≤ CS 3 < 20% or 5% ≤ CS 2 < 20% CS 3 < 1% or CS 2 < 5%	High Moderate Low
L.1	Average Daily Truck Traffic (ADTT)	Moderate	ADTT ≥ 5,000 tpd or ADT ≥ 16,000 vpd 1,000 tpd ≤ ADTT < 5,000 tpd ADTT < 1,000 tpd	High Moderate Low
L.5	Rate of Deicing Chemical Application	Low	Interstate / NHS or ADT ≥ 16,000 vpd 7,500 vpd < ADT < 16,000 vpd Local, Low ADT ≤ 7,500 vpd	High Moderate Low
D.26	Corrosion Protection Level	Moderate	CP 1 CP 2 CP 3 CP 4	Very High High Moderate Low

Both vehicle impact and fatigue cracking damage modes typically had only a few attributes identified by the RAPs. Initial back-casting results showed that when applied to the sample bridges, results for these damage modes were often inconsistent with engineering judgement. For example, a bridge in good condition with low ADT and constructed before key dates in the development of fatigue and fracture requirements resulted in the OF rating of *high*, even though the bridge may have no indication of deterioration and a low likelihood of fatigue cracking. In such cases, categorical models rather than weighted sum models may be more appropriate.

For example, Table 3.15 shows a categorical risk model developed by the RT that could be used to rate the damage mode of fatigue cracking. The criteria shown in the table are presented as examples; specific criteria should be developed by an RAP. In the table, the ADTT for a single lane has been used to be consistent with the commentary for NSTMs (Appendix B). Bridges designed prior to 1985 are assessed differently than those designed after 1985 based on the evolution of fatigue design standards. The AASHTO provisions for resisting out-of-plane distortion cracks were instituted in 1985. As a result, bridges designed after 1985 are less likely to be susceptible to fatigue cracking due to primary and secondary stresses, as discussed in Appendix D. A categorical model of this type can be used to simplify the overall process and ensure the results from the risk models are consistent with expert judgement, which may occur when the number of factors (i.e., attributes) that affect the POF are limited.

Table 3.14. Example risk model for fatigue cracking in steel superstructures.

Code	Attribute	Rank	Criteria	Rating
C.1	Current CR	High	CR 5 CR 6 CR 7+	High Mod. Low
C.2	Current Element CS	High	DE 1010 CS 2 (arrested cracking) No DE 1010 CS 2 (no cracking reported)	High Low
L.1	ADT / ADTT	High	ADT ≥ 10,000 vpd 1,000 ≤ ADT < 10,000 vpd ADT < 1,000 vpd	High Mod. Low
D.17	Worst Fatigue Detail Category	High	E, E' D C A, B	Screen High Mod Low
D.6	Year of Construction	High	Designed before 1975 Designed between 1975 and 1984 Designed between 1985 and 1993 Designed after 1994	High Mod-hi Mod. Low
D.16	Element Connection Type (secondary member connections)	Mod.	Element connected with welds. Element connected with rivets. Element connected with HS bolts.	High Mod. Low

Table 3.15. Example of a categorical model for fatigue cracking.

OF Category	Criteria
High	CR ≥ 6 Designed before 1985, $ADTT_{SL}$ > 1000 tpd Designed after 1985, $ADTT_{SL}$ > 5000 tpd
Moderate	CR ≥ 6 Designed before 1985, 100 tpd < $ADTT_{SL}$ ≤ 1000 tpd Designed after 1985, 1000 tpd < $ADTT_{SL}$ ≤ 5000 tpd
Low	CR ≥ 6 Designed before 1985, $ADTT_{SL}$ ≤ 100 tpd Designed after 1985, 100 tpd < $ADTT_{SL}$ ≤ 1000 tpd
Remote	CR ≥ 6 Designed after 1985, $ADTT_{SL}$ ≤ 100 tpd

Chapter 4 Back-Casting Results

This section of the report describes the back-casting process used to analyze the risk models developed by the RAPs. Back-casting involves the application of the developed risk models to historical bridge inspection records to evaluate if the risk models would be effective if they had been applied in the past. The back-casting process was developed through prior research to address the fact that risk models in general are forward-looking to anticipate the future performance of a system or component (Washer et al., 2014). As a result, the success of the risk model can only be known based on its performance over time by monitoring if emerging risks are detected and managed appropriately by the risk model. Since it is not practical to implement the risk models on bridges and then monitor the performance over time, back-casting was developed to examine the risk models backward in time. In other words, the risk models developed through the research are used to look back at previous inspection data to analyze the performance of the risk models had they been applied at some point in the past.

Back-casting was used to verify if the use of the risk models provided a suitable inspection interval that did not compromise the safety and serviceability of bridges. In the back-casting procedure, the risk models developed by the RAP were applied to individual bridges based on historical inspection records. For example, the risk models were applied to bridges based on 2004 inspection records for the bridge, resulting in an RBI interval that would have been determined in the year 2004 if RBI practices were used at that time. These results were then compared with the actual performance of the bridge, based on the inspection records for 2004, 2006, 2008, etc. to determine if the RBI inspection interval would have adequately addressed the bridge's inspection needs. Criteria for determining the effectiveness of the data model were listed in the NCHRP 782 report (Washer et al., 2014). The original criteria included assessing if the CR for components changed significantly or if major repairs were required in a way that was inconsistent with the risk scores determined from the risk models. The back-casting was also intended to determine if there were any significant risk factors or criteria not identified through the RAP analysis needed in the risk models to provide suitable results.

The back-casting procedure was further developed in this research to examine the overall performance of risk models applied to a collection of randomly selected sample bridges to analyze the quality of the models. This included analysis of risk models for bridge components to determine the appropriate weights for attributes. The target ranges for components in CR 5, 6, and 7 were used to assess the quality of the models and determine the appropriate weights for attributes. Based on the results of these analyses, a new method for calibrating risk models for a given bridge inventory using MC simulations was developed to support the implementation of RBI. The MC simulation process is described in Chapter 5.

This chapter presents the results of back-casting for bridges in six different states. The population of randomly selected sample bridges is described, and data analysis is shown on both a component-level and bridge-level. Deterioration modeling to characterize the general behavior of bridge components based on NBI data analysis is shown. The deterioration modeling provides an overview of bridge behavior that supports the use of extended inspection intervals for bridges based on their typically slow deterioration patterns.

4.1. Back-Casting Bridge Population

The bridges used for back-casting were selected randomly from each state’s bridge population according to the material of focus for that state. The RAPs considered PSC and steel superstructures, with four states focused on steel bridges and two states analyzing PSC bridges. Ten bridges were selected from each state. The bridges were selected at random with two provisions. First, the bridges selected for a particular state were of a certain family of bridges, meaning the bridges had superstructures of a certain material type, either PSC or steel. The bridge family selected for each state matched the risk models developed by the RAPs for that state. Second, bridges were selected to provide geographic distribution across a state. For example, Figure 4.1 shows the geographic distribution of the PSC bridges assessed in the state of Washington.

The rationale for randomly selected bridges was to implement the risk models across a cross-section of the bridge population as compared to, for example, only selecting bridges in good condition. It was hoped that such a distribution of bridges would provide insight into the effectiveness of the model for representing the bridge inventory overall and gain insight into how to weight individual attributes. Additionally, many of the condition-related attributes such as the CR and CS would be rated *low* for bridges in good condition so there would be little opportunity to analyze if the risk models were effective for identifying bridges with increased risk. Finally, the randomly selected bridges would assess if the risk models were durable across the typical bridge inventory in terms of being applicable to all bridges, regardless of the CR for the bridge. Since the risk models assess relative risk, not absolute or quantitative risk, analyzing the effectiveness of the model and developing a methodology for weighting the attributes requires bridges of different conditions, ages, and loading.

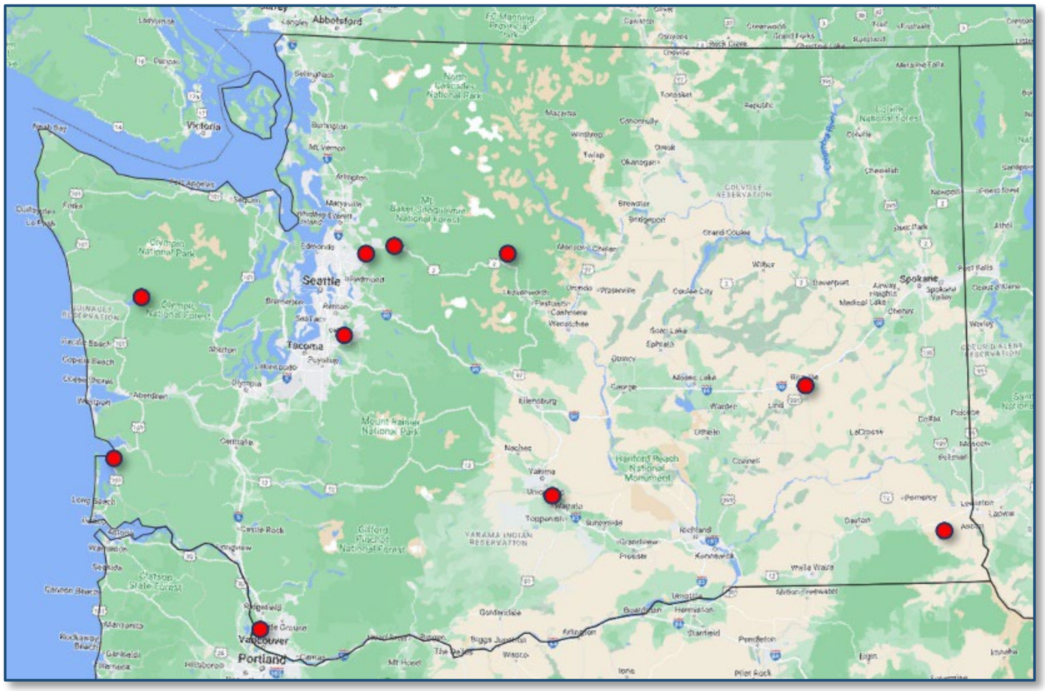


Figure 4.1. Geographic distribution of sample bridges in Washington.

The randomly selected bridges provided a “sample” population of bridges to assess the RAP models and their effectiveness in providing a suitable risk profile of bridges that could be used for inspection planning.

The bridge material types were primarily steel and PSC with 39 steel bridges, 20 PSC bridges, and one bridge that had PSC approach spans and a steel main span. The distribution of different material types is shown in Figure 4.2 using the Coding Guide designations for bridge main spans of steel and steel continuous (steel cont.), and PSC and PSC continuous (PSC cont.). As shown in the figure, steel continuous superstructures formed the largest group among the sample bridges (37%), and PSC continuous formed the smallest group (7%).

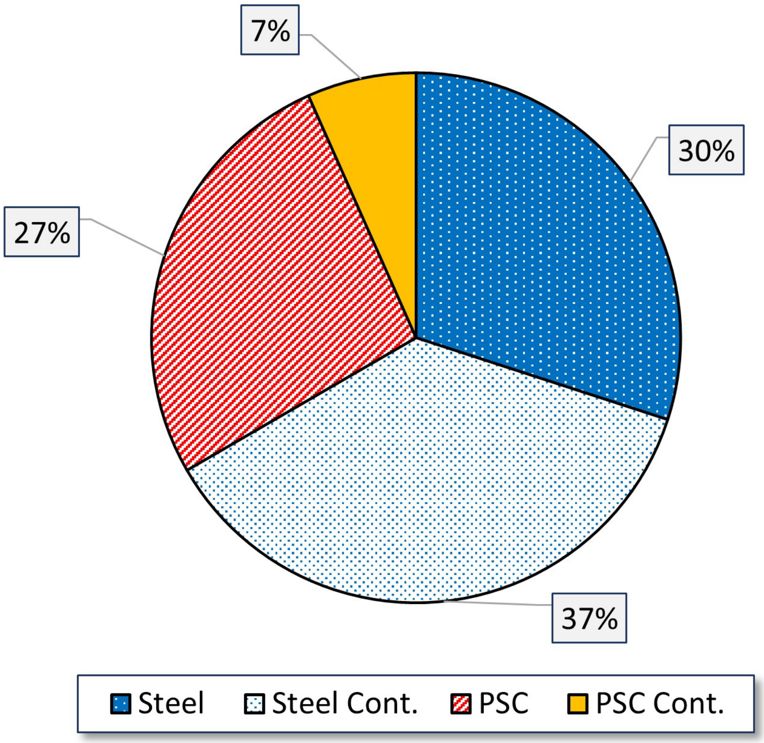


Figure 4.2. Distribution of sample bridge superstructure materials.

Figure 4.3 shows the distribution of the ages of the sample bridges in Figure 4.3 (A) and the ADT characteristics in Figure 4.3 (B). There was a broad distribution of ADT and age of the randomly selected sample bridges. The average age of the sample bridges in 2020 was 48 years (Standard deviation (σ) = 17 years), with a minimum of 14 years and a maximum of 88 years. The average ADT was 11,537 vpd (σ = 21,945 vpd) with a minimum of 12 vpd and a maximum of 136,800 vpd.

The CRs of the sample bridge components of deck, superstructure, and substructure varied from a low of CR 2 to a high of CR 9 during the back-casting interval. The average CR for the sample bridge population was CR 6 (σ = 1). Figure 4.4 shows the frequency plot of the CRs for the sample bridge population for the deck, superstructure, and substructure bridge components based on 2020 NBI data. It should be noted that most of the sample bridges had CRs that changed over the course of the back-casting period, including bridges that had renovations and repairs. This included, for example, overlays installed on decks, and repairs of a fractured bearing area of a substructure.

There were 13 bridges that had a CR of 7 or higher for all three components in 2020. Eight of the bridges included in the random selection of 60 bridges had a scour CR of 5 or less, indicating that these bridges would not qualify for an extended inspection interval. However, the scour ratings were not considered in the analysis because the RAP models do not address scour. The sample population also included five bridges that had a component rated CR 4 or lower during the back-casting period. Overall, the sample bridges reflected the diversity of the bridge population in the US. This provided data that allowed the risk models to be analyzed for bridges with different CRs to assess the performance of the risk models across the existing bridge inventory.

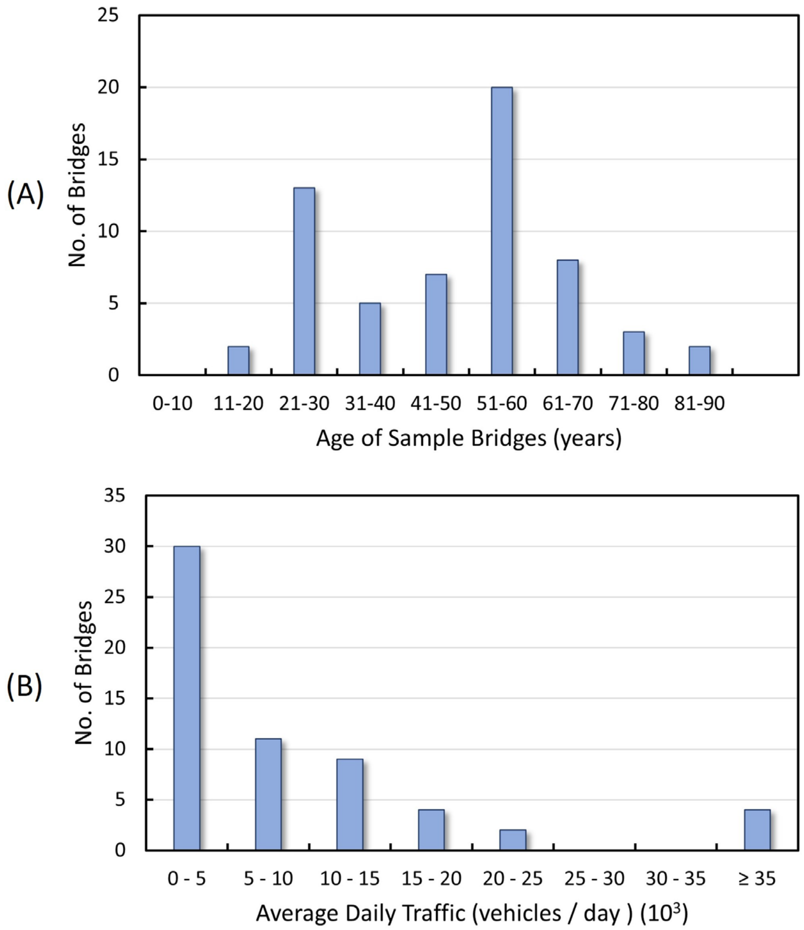


Figure 4.3. Plots showing age of sample bridges (A) and ADT (B).

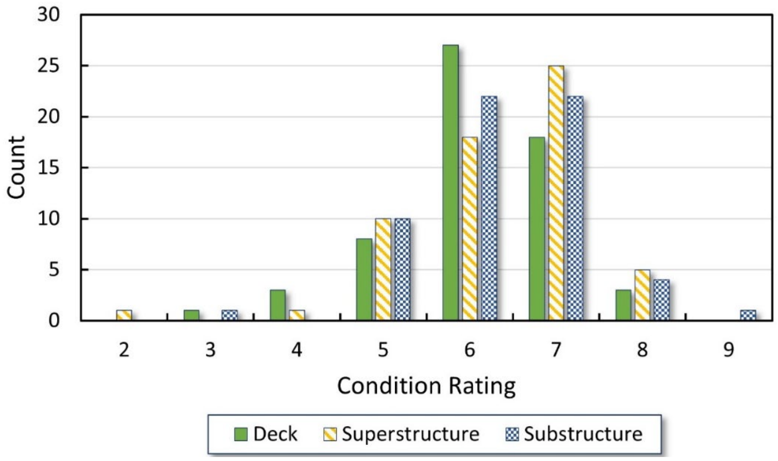


Figure 4.4. CRs for the deck, superstructure, and substructure components of the sample bridges.

4.2. Deterioration Models

The extended inspection intervals realized through RBI can be supported through the analysis of NBI data to determine the time intervals typically required for bridges to deteriorate to poor condition. The TICR values determined through deterioration modeling provide guidance on the length of time for bridge components to deteriorate to poor condition. This data provides critical quantitative evidence for extended inspection intervals by characterizing the relatively slow deterioration patterns of bridges. Typical median TICR values for components in CR 5, 6, and 7 are about 10 years. As a result, many inspections conducted at 24-month intervals will not reflect any change in CR from the previous inspection.

Analysis of NBI data for all nine states that participated in the study was completed. This section of the report describes the deterioration model applied in the research and summarizes the overall results.

4.2.1. NBI Data Analysis

Data from the NBI were analyzed for each of the participating states to analyze the TICR for the bridge families of PSC and steel bridges, R/C decks, and substructures. The analysis's objective was to assess the typical time intervals that components remain in each CR and their service life. Kaplan-Meier survival analysis was used to assess the TICR for bridges. Data was downloaded from the FHWA Infobridge web site for the years of 1992 – 2020 and preprocessed for analysis. The pre-processing included trimming the data and identifying errors in the data. An initial analysis was completed using data covering 1992 – 2017. However, systematic errors in the Infobridge data that was available for downloading from the web resource were found. The RT worked with FHWA to correct the error; corrected data was subsequently downloaded from 1992 – 2020 and were analyzed. Details of the data trimming process undertaken is described in the first interim report and previous research (Nasrollahi & Washer, 2015).

4.2.1.1. Kaplan-Meier Method

Kaplan-Meier (K-M) is a common method for treating discontinuous reliability data such as CRs for bridge components (Allison, 2010; Kaplan & Meier, 1958). The K-M method is a nonparametric maximum likelihood estimator of time-to-event data such as data describing bridge components transitioning to a different CR. The time-to-event data used for the analysis was the transition time for a bridge component to drop from one CR to the next lower CR.

One way to describe a random variable's reliability distribution is using the cumulative distribution function (CDF) graph of the cumulative probability of failure up to each point. In survival analysis, the CDF gives the probability (P) that the survival time, *T*, is less than or equal to a specific time, *t* (Allison, 2010). The CDF for a randomly selected bridge component TICR is the probability that the bridge component stays in a given CR less than or equal to a selected time *t* can be written as:

$$F(t) = P(T \leq t)$$

Equation 1. Cumulative probability of failure function.

When the reliability data are uncensored or only right-censored, the reliability can be calculated using the Kaplan-Meier estimator by the following equation:

$$\hat{S}(t) = \prod_{j:t_j \leq t} \left(1 - \frac{d_j}{n_j}\right) \text{ } for \text{ } t_1 \leq t \leq t_k$$

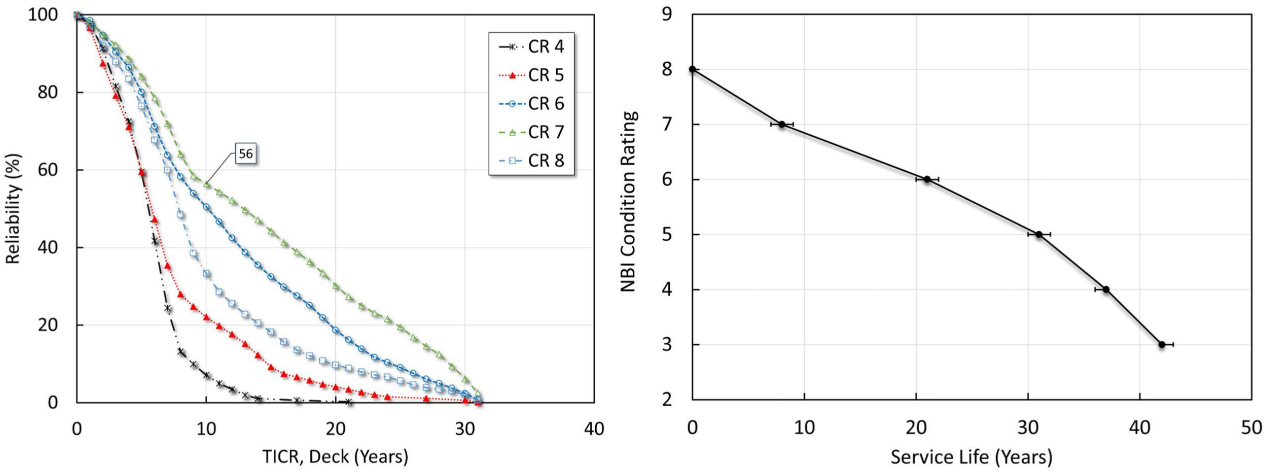
In the above equation, $\hat{S}(t)$ is the Kaplan-Meier estimator, d_j is the number of bridge components for which the event occurred (transitioned to the lower CR) at time t_j, n_j is the number of bridge components at risk of event at time t_j, and t_1 and t_k are the boundaries of the range for *k* distinct event times.

Equation 2. Kaplan-Meier estimator function.

The P in Equation 1 can be interpreted as estimating the conditional probability of surviving to time t_{j+1} given that a bridge component has survived to time t_j (Allison, 2010). For times less than t_1 (before the first event), $\hat{S}(t)$ is equal to 1 (all bridge components are staying in a given CR) and $\hat{S}(t)$ is equal to 0 for the case of no censored data for $t > t_k$ (all bridge components transitioned to lower CR) (Allison, 2010). The K-M estimator is accompanied by statistics such as the mean, median, confidence interval for the median, standard error of the mean, and hazard rate that can be used to analyze results. In the presence of censoring, the mean is not a good measure of the central tendency because the data are "skewed to the right" and therefore the median provides superior statistic to illustrate the central tendency for the data (Hosmer Jr, Lemeshow, & May, 2008).

An example K-M reliability curve is shown in Figure 4.5 (A) for bridge decks in the state of Washington. The ordinate of the figure shows the reliability of R/C decks. The values plotted show the probability that a component has not transitioned to the next lower CR at any given time interval, shown on the abscissa (horizontal axis). For example, after 10 years there is a 56% chance that a CR 7 deck has not yet transitioned to CR 6. The median transition time is 13 years, which corresponds to 50% on the ordinate of Figure 4.5 (A). The median value was utilized to characterize the TICR for components rather than the mean (average) values, which are typically larger values than the median because some components remain in a single CR for many years. Components with good attributes for durability would be expected to have TICR values greater than the median, while those with poor attributes would be expected to have TICR values less than the median.

The median TICR data were used to form service life estimates for the components of decks, PSC superstructure, steel superstructures, and substructures. Figure 4.5 (B) shows an example of a service life estimate for a deck, again based on NBI data from Washington shown in Figure 4.5 (A). The data in the plot were formed from the median K-M values for each CR from 8 to 4.



(A) (B)

Figure 4.5. K-M results showing reliability of bridge decks (A) and service life of decks based on median results (B).

These data provide quantitative assessment, based on historical performance of bridge components in a particular jurisdiction, of the suitability of longer inspection intervals for certain bridges. For example, considering a R/C deck with generally good attributes that has just transitioned from CR 8 to CR 7, and has a median TICR values of 13 years for CR 7, 10 years for CR 6, 6 years for CR 5, and 5 years for CR 4, a median service life estimate for that component to transition to CR 3 is approximately 34 years (13 + 10 + 6 + 5). This simple example illustrates that the service life of a bridge deck before it transitions to serious condition (i.e., CR = 3) is much greater than the current inspection interval of 24 months. In fact, the median time interval before the component would transition to CR 3 is more than five times the maximum RBI interval of 72 months.

Of course, when using RBI to determine the inspection interval, if inspections reveal damage developing on the deck, the condition-related attributes in the risk model would be affected; thereby, the inspection interval may be reduced. Additionally, the intention of the design and loading-related attributes in the risk models, combined with the condition attributes, is to identify those bridges that could be expected to deteriorate more rapidly than the median values of the TICR. As such, the implementation of RBI intervals relies on much more than deterioration models. Regardless, this simple example illustrates how the K-M analysis and TICR data for bridge inventories can be used to provide quantitative analysis to support RBI processes and provide a foundation for extended intervals for low-risk bridges.

For all components and state NBI inventories modeled, the TICR values indicate that components in good condition have a lower bound TICR of at least 8 years across the nine states for components of R/C decks, steel superstructure, PSC superstructure, and substructure. The results of the analysis are summarized in Figure 4.6 that shows the average service life for the key components that were part of the study. The bar graph shows the average service life based on the K-M median values from CR 8 through CR 3 and average service life from CR 8 to CR 5. This data illustrates that median lives of these components are between 45 and 50 years for all components studied when considering CS 3 as failure in terms of risk analysis. The figure also shows that the average service life for a component to deteriorate from good condition (CR ≥ 7) to fair (CR 5), which is the focus of extended inspection interval analysis, is about 30 years. The graph also shows error bars that represent the standard deviation of the service life calculation between the nine state bridge inventories studied.

The individual results from the K-M analysis from nine states that participated in the study are available as supplemental data from the research. Results showing the service life graph, reliability and deterioration graphs, and the cumulative hazard functions are provided in Data Supplement A, *Data from Kaplan-Meier Deterioration Analysis*. Results in the supplement include K-M analysis for R/C decks, steel superstructures, PSC superstructures, and substructures for each of the nine states.

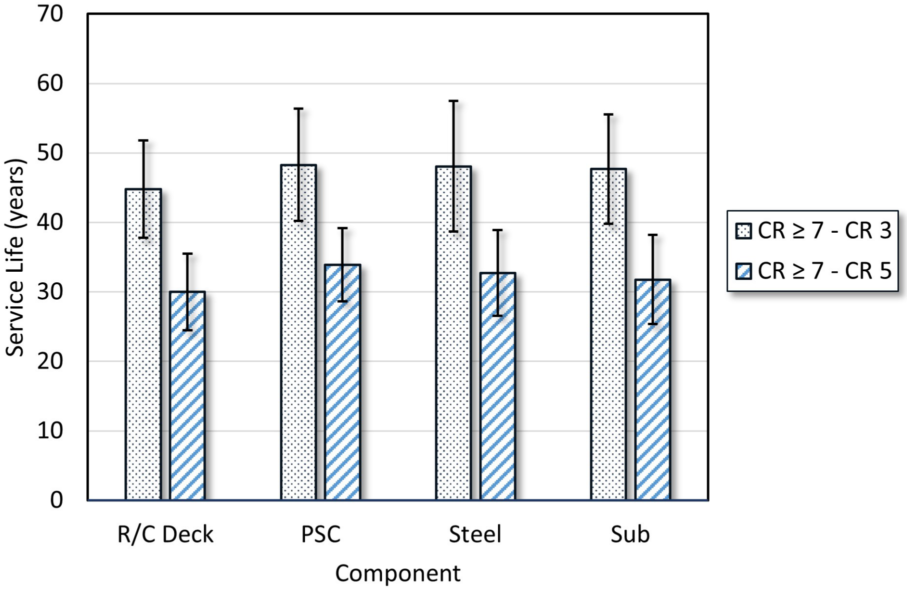


Figure 4.6 Average service life for bridge components of R/C Decks, PSC and steel superstructures, and substructures.

4.3. Back-Casting Results

4.3.1. Approach

The initial risk models developed by the RAP were used to perform back-casting on ten bridges from each state. Inspection records for the sample bridges were collected over time intervals ranging from 1996. Inspection records were reviewed with an emphasis on utilizing inspection notes to determine or infer how the criteria of an individual attribute should be rated for a particular bridge.

A spreadsheet program was used to document the results of inspection report reviews in a systematic manner and allow for the risk model and attributes to be analyzed for a particular group of sample bridges. The spreadsheet program stored data on the attributes and criteria for each risk model and sample bridge. The results from review of the inspection records for each inspection year were stored in look-up tables in the program. The results were stored for each of the attributes in each risk model based on the criteria for the attribute. The program calculated the OF value from the risk model based on the attributes for each component. The resulting inspection interval was determined by considering the CF category of *moderate* and *high*. In this way, the inspection record analysis could be stored and used for sensitivity studies of the effect of weighting attributes to improve the risk models' quality.

An example of back-casting results is shown in Figure 4.7 to illustrate how the back-casting was conducted and provide an example of a bridge with components in CR 7 but elevated risk factors (i.e., attributes that were rated *high*). The CRs for the components of the deck, superstructure, and substructure from the NBI data are shown in Figure 4.7 (A). The inspection interval based on a CF of *moderate* is superimposed on the figure to show how it changed over time. Figure 4.7 (B) shows the results from applying the risk models for the superstructure, substructure, and deck damage modes in each inspection year to determine the OF value. The OF value is shown on the ordinate. For this bridge, the steel superstructure had coating damage, high ADT, leaking joints, deck drainage issues, and the superstructure was subjected to overspray from a roadway below. The bridge's deck deteriorated over the back-casting period, resulting in increased OF values as shown in the figure. The OF also increased for the damage mode of impact because the reported vertical clearance changed from 15.25 ft to 14.78 ft. The criterion for the attribute of vertical clearance was rated *high* when the vertical clearance was less than 15 ft, and *moderate* when the vertical clearance was 15 to 17 ft. As a result, the small change in vertical clearance changed the OF value. Because many of the attributes for the bridge were rated *high* due to the aggressive environment (i.e., deck drainage issues, leaking joints, high ADT, and overspray), the risk scores were relatively high even when the bridge was assessed as CR 7. For other CR 7 components in the study, the risk scores were typically much lower and sometimes remained flat throughout the back-casting period.

The results from back-casting for each of the sample bridges using unweighted risk models are documented in Data Supplement B, *Back-Casting Graphs*.

4.3.1.1. Challenges with Back-Casting

The back-casting proved to be challenging for several reasons. The two primary issues experienced were that the notes and features of older bridge reports did not provide data consistently over time. For example, damage reported in one bridge inspection may not be present in the next inspection report, so it was difficult to track the damage's progression. This was particularly problematic for bridges that did not have element-level reports. In some cases, inspection policies were evolving over the back-casting time interval, resulting in inconsistent inspection data. For example, element-level inspection data became available or reported elements changed because a state policy changed regarding data requirements for inspection reporting. In addition, some attribute qualities identified by the RAPs were not present in the bridge inspection records, and therefore, required some assumptions or inferences to rate the attributes. The inspection reports sometimes included little information regarding why a CR changed from one inspection to another. Older bridges with element-level data frequently included inspection notes that did not align with the CS assigned by the inspector. For example, a bridge deck would be described in notes as having widespread cracking or spalling, but 100% of the deck was recorded as in CS 1. Overall, notes describing damage were very inconsistent in terms of quantity, frequency, and level of detail. Additionally, because the review of the inspection reports often involved assumptions and inferences regarding the appropriate ratings for attribute criteria, the reviews were not repeatable between different reviewers.

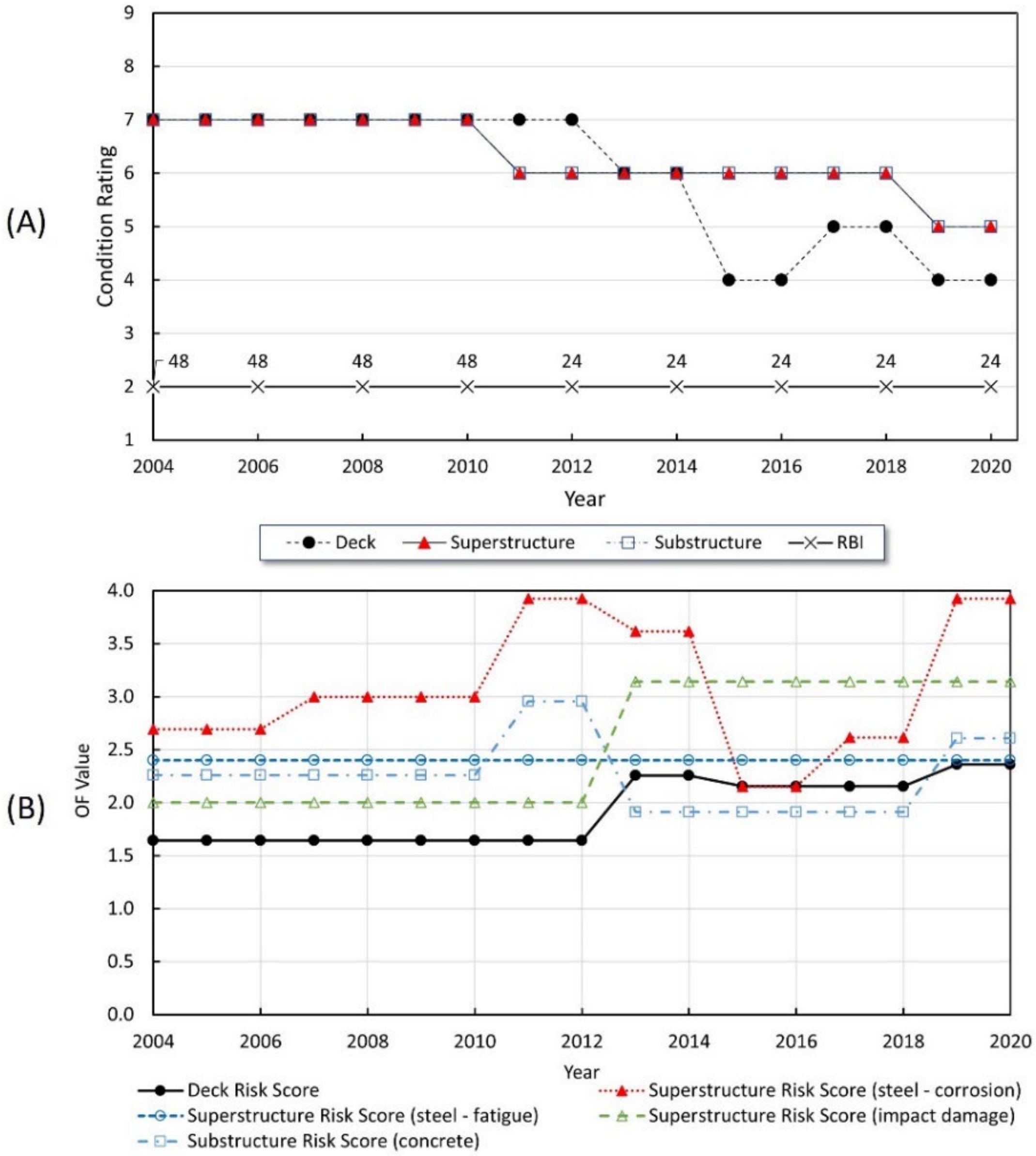


Figure 4.7. Back-casting results for a sample bridge showing the NBI CR and inspection interval (A) and the OF values for each year.

4.3.1.2. Weighted Sum Model

The OF is calculated as a weighted sum model where the initial weights for the model are developed through an expert elicitation process with a RAP. The initial value weights for a given attribute were set by simply ranking a given attribute as *high*, *moderate*, or *low* in terms of its impact on the POF. Attributes ranked as *high* would be expected to have a significant impact or influence of the likelihood and an attribute ranked as *low* would be expected to have a minor impact. The rankings are subjective but provide a starting point for the weighted sum model, which can be adjusted as necessary to better reflect the actual performance of a family of bridges of similar materials and design. A method for weighting the attributes was developed through research and will be described later in Chapter 5.

The weighted sum model used for scoring individual damage modes in the research initially is shown by the equation:

$$OF = {\sum A_i} \Big/ {\sum A_{i,\,max}} \times 4$$

In the equation, A_i is the score for an individual attribute based on its rating criteria and $A_{i,\,max}$ is the maximum score for an individual attribute based on its rank.

Equation 3. Unweighted OF equation.

This equation uses the weights for each attribute according to the rank provided by the RAP and the result of assessing that attribute's criteria. The scores for each individual attribute are summed to produce the numerator and the maximum scores for each attribute are summed to form the denominator.

Most attributes identified by the RAPs were ranked as *high* indicating a significant impact on the POF. It was found through RAPs in this study that very few attributes were ranked as *low*. The attributes identified by the RAPs are those that the individual members of the RAPs consider most important so it would be normal that many of the attributes would be ranked as *moderate* or *high*. Attributes ranked *high* are assigned a maximum value for $A_{i,max}$ of 20 points, attributes ranked *moderate* are assigned a maximum value of 15 points, and attributes ranked *low* are assigned a maximum value of 10 points, as previously discussed.

Once an attribute has been ranked to determine the maximum score for the attribute ($A_{i,max}$), criteria for each attribute are used to rate the attribute. The rating of the attribute assigns its actual score (A_i) when applied to a bridge component. Three criteria are typically developed to determine if the attribute should be rated *high*, *moderate*, and *low*. The points assigned (A_i) are distributed as described in section 3.3.5, with a criterion rated as *high* being assigned 100% of the rank value ($A_{i,max}$), *moderate* 50% of ($A_{i,max}$), and *low* assigned zero points.

The weighted sum model is intended to be a simple and rapid process to apply using engineering judgement. Different approaches of providing additional weighting to attributes were studied to better match the outcome of the risk models with the target values when applied to actual bridges and bridge records. The individual attributes were weighted using the equation:

$$OF = {\sum w_i A_i} \Big/ {\sum w_i A_{i,\,max}} \times 4$$

In this equation, w_i is a weighting factor assigned for a given attribute, A_i.

Equation 4. Weighted OF equation.

This equation allows for the attributes initially weighted by the RAP to have their overall weight in the model increased (or decreased). For example, the score of the CR attribute of a deck was typically 20 points based on its rank of *high*. A multiplier of 1.50 would increase the rank value ($A_{i,\,max}$) of the attribute to 30 points. When the attribute is rated according to the criteria developed by the RAP, the rating of *high* is increased to 30 points, *moderate* increased to 15 points, and *low* rating remains zero. The maximum score for the model is also increased according to Equation 4. In this way increasing the weight of an individual attribute reduces the relative weight of all other attributes in the model, since the denominator is increased.

The research did not find suitable existing procedures for adjusting the weights of individual attributes in a weighted sum model. Several methods were explored and found to be impractical or not related to engineering decision-making. For example, a method for determining the weights of individual attributes based on its statistical properties provided weights that primarily showed which attributes were most likely to vary over the course of time rather than any engineering rationale. Most of the approaches described in the academic literature for weighting attributes in a weighted sum model did not adequately represent engineering decision-making when implemented on the risk models for bridge components. For this reason, new methods of analyzing risk models for bridges were developed and tested to find suitable weights for attributes.

4.3.2. Results from Original Risk Model

This section of the report describes the preliminary results from the original back-casting using the RAP models. Much of the data is analyzed on a component basis to assess the effectiveness of the risk models and provide general results that show how the RAP models performed when applied to actual bridge components. An analysis of the risk scores for all the bridge components in the study is reported.

Some of the original risk models developed from RAP meetings did not include the CR of a component explicitly as an attribute. The RAP meetings focused on attributes and damage modes that indicated an increased relative risk such as existing corrosion damage, rate of deicing chemical application, joint condition, etc. The damage modes that were identified by the RAP would affect the CR even if the CR were not explicitly mentioned in all cases. When the sample bridges were scored with risk models that did not include the CR explicitly as a separate attribute, it was found from initial back-casting that the risk scores often did not align with the target ranges for individual sample bridges. For example, a CR 4 component would have a lower risk score than a CR 7 component because bridge condition was not adequately represented in the risk model. These initial results were not very informative and are not included here. Additionally, the risk models without CR attributes did not reflect the rational assessment that most CR 5 components would be more likely to deteriorate to a CR 3 in the next 72 months than any CR 7 component based simply on the fact that the component is already in CR 5. While it may be possible for a CR 7 bridge to deteriorate more rapidly, it would not be common. The CR was implemented as an attribute for damage modes that would affect CR, such as corrosion-related damage modes. Damage modes for which the risk is unrelated to the component condition such as the impact damage did not have the CR attribute included because the likelihood of a vehicle impacting a bridge is unrelated to its condition.

Figure 4.8 shows the raw risk scores for corrosion-related damage modes for the deck, superstructure, and substructure based on risk models that include the CR for the subject component and the CS for the element under consideration. The ordinate on the left shows the risk score calculated using Equation 3 and the abscissa shows CR for each component. The OF categories are shown next to the ordinate on the right. The data are shown for damage modes of deck delamination and spalling (R/C deck delam. and spalling), steel superstructure corrosion damage (Stl. ss. corrosion), R/C substructure delamination and spalling (R/C sub. delam. and spalling), and delamination and spalling of a PSC superstructure member

(PSC ss. delam. and spalling). The data points are slightly offset from the associated CR for clarity, and a trend line in the figure shows the linear regression for all the data points combined. These data illustrate that the risk models produced risk scores that trended toward larger values for components with lower CR. The average value for components with CR 5 was 2.54, in the *moderate* range. The average value for components with CR ≥ 7 was found to be 1.11, in the *low* range for the OF. However, Figure 4.8 shows that, in some cases, components with CR ≥ 7 have risk scores greater than components with CR ≤ 5. Most of the components with CR ≥ 7 were rated in the *low* or *moderate* range. These data indicate that the design and loading attributes in the models have weights that are too high as compared with condition attributes. As a result, the models did not produce results that aligned with the target ranges and provided suitable contrast between the calculated OF for components in CR 7 and components in CR 5 or lower. The assigned values in the risk model needed to be adjusted to produce results consistent with the target values and engineering judgement. Several different approaches were pursued to properly weight the attributes in the models to better align results with the target ranges.

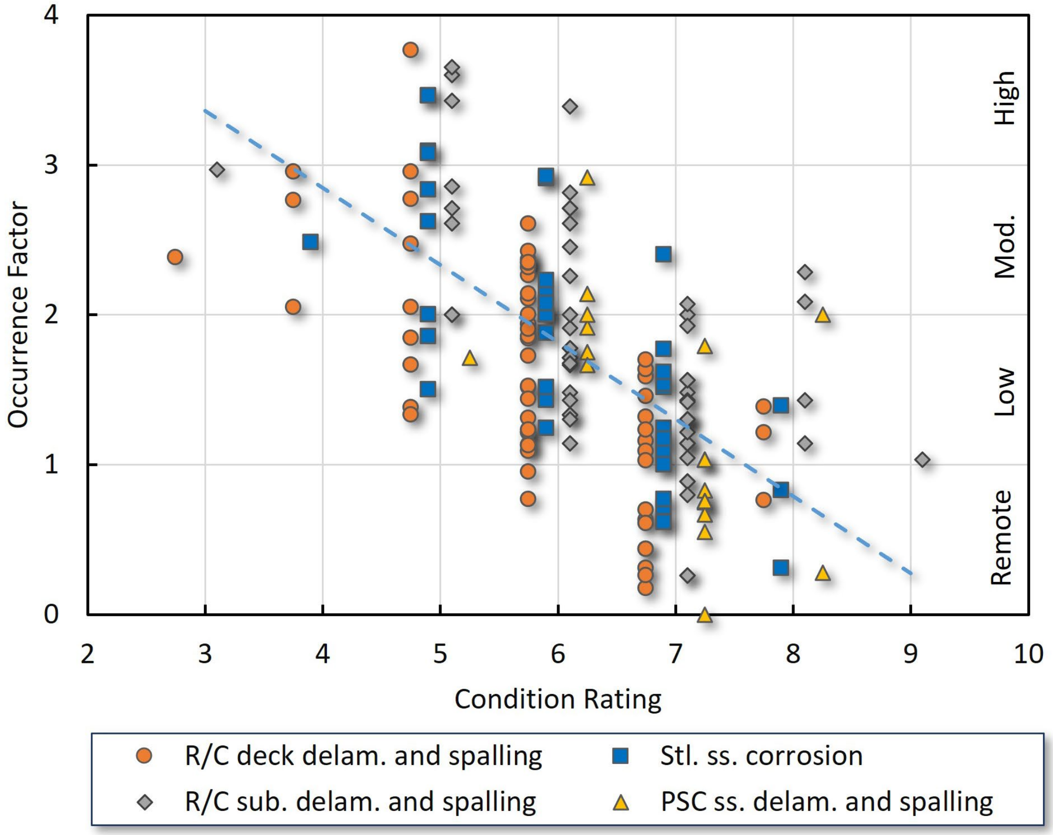


Figure 4.8. Raw OF scores for corrosion-related damage modes for sample bridges.

4.3.3. Assessment of Model Weighting

Based on the results from the initial back-casting process, it was clear that the original risk models were not effectively representing the expected increase in risk as the CR decreases. Sensitivity studies were conducted to assess the impact of different approaches to weighting the attributes. The objective of the studies was to determine if increasing the weights of certain attributes would improve the quality of the risk model when compared with the target ranges and provide insight into how to calibrate the models for implementation.

First, a study was conducted based on a procedure that ranked individual attributes based on statistical analysis to assign individual weights. This approach was ineffective and is not reported here. A second study was completed in which attribute weights were adjusted by two methods. The weights of groups of attributes were increased using Equation 4, in which the attribute value is increased and the total number of points in the model was also increased. As part of that study, the individual attributes for the current component CR and the element CS were weighted without increasing the total value of the model. The study's objective was to determine if groups of attributes or individual condition attributes should be weighted to improve the model's quality when compared with the target ranges.

The primary conclusion from these sensitivity studies was that increasing the weights of condition-related attributes improved the quality of the risk models when compared with the target ranges. While not true for every case, the trend indicated that the most likely approach to improve the quality of the models was to weight the CR and CS in the models.

It was also concluded from this sensitivity study that a more systematic method of weighting the attributes was needed to effectively calibrate the risk model to meet the target ranges. The population of sample bridges produced different results for the different weighting scenarios studied. The number of condition, loading, and design attributes varies for different models. The sample of 60 bridges each had unique characteristics and deterioration patterns. Additionally, the historical data obtained from inspection reports was cumbersome to work with and difficult to repeat. While this parametric sensitivity study produced some insight into the behavior of the risk models as compared with actual bridges, it would not be practical to apply this method to calibrate the models. A more effective methodology was needed, and the MC simulations described in Chapter 5 provided a more durable and implementable approach.

Additional studies and analysis of the back-casting were conducted with the CR and CS weighted by a factor of 2 as described below.

4.3.4. CR and CS Weighting

The risk models were implemented with increased weights for the primary condition attributes based on the results of the initial analysis of the back-casting results and the sensitivity studies. The models were weighted by increasing the value of the CR and CS attributes (C.1 and C.2) by a factor of 2 ($w_i = 2$). Increasing the weights of these attributes decreases the weights of all other attributes. This produced results for weighted models that could be compared with results for the unweighted models.

The results for the primary corrosion-related unweighted risk models were previously shown in Figure 4.8. As shown in that figure, there is a general trend that lower CR components have increased values of OF. However, there are cases where CR 7 components have OF values that are equal to or greater than CR 5 components. Most CR 7 components exceed the value of 1 for the OF and would be rated as *low*. If the CF is *high* and the OF is *low*, the inspection interval would be 48 months, based on the proposed risk matrix. Therefore, it was desirable to increase the contrast between components in CR 7, which practically would be expected to have remote likelihood in most cases based on engineering judgement, and components in CR 5, which would not qualify for extended intervals using Method 1 according to the current NBIS requirements. Additionally, the sensitivity study described in the previous section showed that weighting the risk models by increasing the relative value of the condition-related attributes improved the quality of the risk models when compared with the target ranges. To provide additional contrast in the risk values that would better align with expected values, the results from the back-casting were modified by multiplying the CR and CS attributes by 2.

Weighting the models in this way has the effect of reducing the risk values for components with CR ≥ 7, when attribute C.2, Current Element Condition State, is also rated as *low*. The risk values are reduced because the attribute C.1, Current Condition Rating, is rated as *low* for a component with CR ≥ 7, and therefore, scores 0 points, regardless of what weighting factors are applied to the models. The weighted attribute's maximum score is added to the denominator, resulting in a reduced risk score overall. For bridges with CR 6 and lower, points are added to both the numerator and the denominator, resulting in an increased risk score. Additionally, it is much more likely for a CR 6 or CR 5 component to have element CSs that rate as either *moderate* or *low* for attribute C.2 as compared with a component in CR ≥ 7. As a result, the risk scores are reduced for components in good condition and increased for components in fair or poor condition.

The overall results of using the weighted risks models are shown Figure 4.9. The figure includes a linear regression line based on all the data shown. It can be observed that the slope of the regression line is increased as compared with the regression line shown in Figure 4.8. Notable in the figure is that all the components in CR ≥ 7 now score in the *remote* or *low* range. Components in CR 6 are primarily in the *low* or *moderate* range. Components in CR ≤ 5 typically score in the *moderate* to *high* range. These results illustrate greater contrast in the risk scores for components with different CRs, with components in good condition having lower risk scores and components in fair and poor condition having increased risk scores as compared with the unweighted risk models. The results of the weighted models align with the target ranges described earlier.

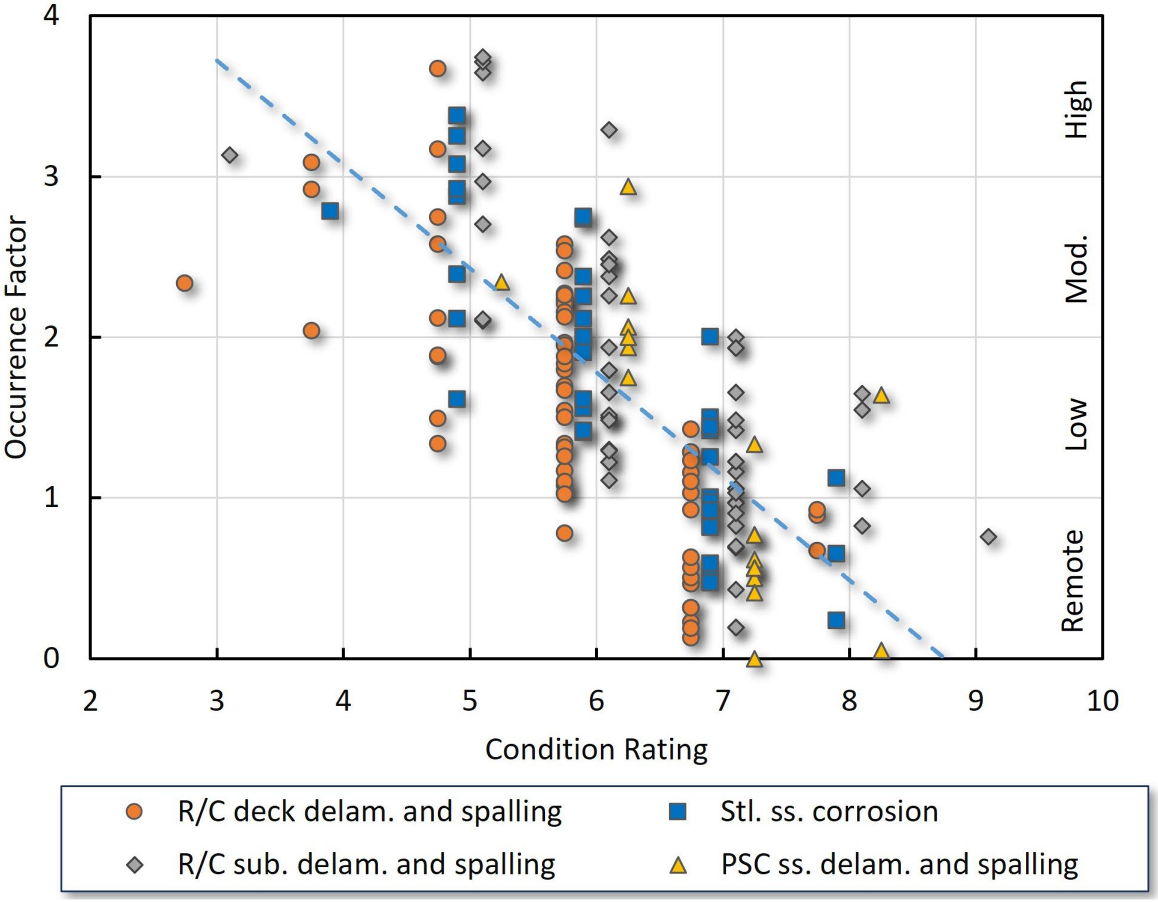


Figure 4.9. OF results for components of 60 samples bridges with weighted CR and CS.

The weighting of the CR and CS attributes has a relatively small overall effect of the average value of the risk score for components in different CR as shown in Figure 4.10. This figure depicts the average of the risk scores calculated from the risk models. The average risk scores for the unweighted and weighted models are shown. Error bars show the standard deviation of the results. The figure illustrates that the average OF value does not change by a large amount, but the change is significant because the average for CR ≥ 7 components drops from 1.11, in the *low* range, to 0.9, in the *remote* range. When compared with the proposed risk matrix shown in Figure 2.4, components in CR ≥ 7 with a *remote* OF would qualify for extended intervals even if the CF were *high*. As discussed earlier in the report, the target range for CR 7 is in the *low* to *remote* range*,* based on the rationale that the bridges in good condition rarely, if ever, deteriorate to a CR of 3 in and 72-month interval.

Based on these data from the sensitivity studies, the overall back-casting results were analyzed for scenarios where the CR and CS are weighted by a factor of 2 and compared with the original, unweighted risk models. This data was used to estimate the resulting inspection interval that would apply based on the risk scores.

4.3.5. Risk-Based Intervals

This section discusses the overall trends in the data formed from the back-casting process. For the back-casting, the inspection intervals were determined for two different CF scenarios, CF = 2, *moderate*, and CF = 3, *high*. The CF = 4, *severe*, was not included in the analysis because this CF applies to bridges that lack redundancy such as NSTMs. The CF factor of *low* was not included because it has the same intervals as the CF of *moderate* except for bridge components with a *high* OF according to the risk model shown in Figure 2.4. As shown in the back-casting data, a *high* OF typically occurs for bridge components with a CR of 5 or less. As a result, there was little relevant information contained in an analysis of a *low* CF.

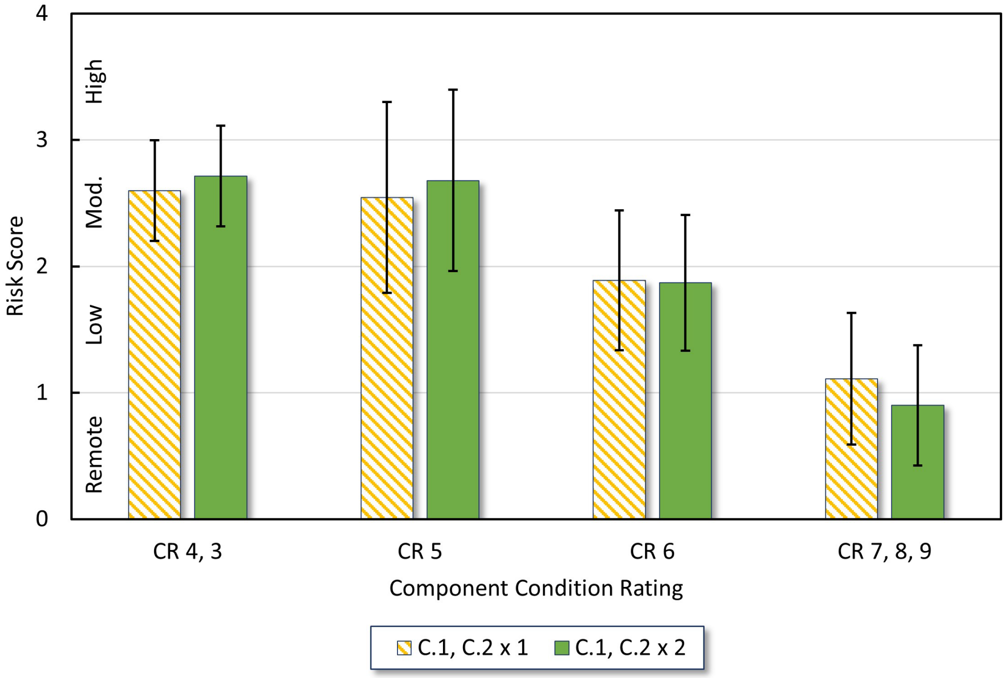


Figure 4.10. Bar chart showing the average OF for sample bridge components for weighted and unweighted models.

The inspection interval for a bridge is controlled by the highest risk score for any component. As such, the controlling component with the highest risk score was used to assess the sample bridges and determine the risk-based inspection interval.

4.3.6. Inspection Intervals Based on Component Ratings

This section shows results for applicable inspection intervals considering the risk scores for the controlling component and damage mode for each of the sample bridges. Results are presented with different CFs to illustrate the potential outcome of the analysis in a general way considering that different owners may have different parameters for the CF. The CFs of *low* (CF = 2) and *high (*CF = 3) were used to determine the inspection intervals based on the unweighted and weighted risk scores using the original risk matrix included in the NCHRP 782 report. In addition, the inspection intervals were determined with the proposed risk matrix, discussed earlier in the report, in which a component with a remote OF and the high CF could be assigned an interval of 72 months. The latter scenario is listed as "CF 3P." The results presented in this section consider the CR for components in the year 2020 as compared with the results from the risk analysis. Results for weighted and unweighted models are presented. The data are analyzed based on the controlling component risk score for each bridge. The NBIS requirement that only bridges in good condition are eligible for extended intervals of 72 months was not considered in the analysis. An analysis of the bridges in the sample set that had all three components with CR ≥ 7, i.e., good condition, is provided later in the report.

The results showed a distribution of inspection intervals that were slightly different if the weighted models were used as compared with the unweighted model. Table 4.1 shows the overall results for the unweighted and weighted models. It was found that for the CF = 2, 42% of the sample bridges could be assigned an inspection interval of 72 months. For CF = 3, there would be no bridges in the sample population that would qualify for a 72-month interval using the risk matrix from NCHRP 782. If the proposed risk matrix were used, 5% of the sample bridges would have an interval of 72 months.

When the weighted models were used, there was a small difference in the number of bridges with CF = 3 that would have a 72-month interval, increasing from 5% (3 bridges) to 8% (5 bridges) as shown in Table 4.1. It is notable that the percentage of bridges with a 24-month interval increases when the condition factors (CR and CS) are weighted as compared with the unweighted model. The slight increase in the number of components that have a 72-month interval does not seem that significant; however, the number of bridges in the sample population with all components of CR ≥ 7 was relatively small, only 13 of the 60 bridges. It is also notable that for CF = 2, the percentage of bridges eligible for a 72-month interval goes down when the weighted model is used. This occurs because some of the components eligible for a 72-month interval in the weighted model are controlled by components in CR 6. As a result, the risk score for these components is increased when the model is weighted.

Analyzing these results according to the CR of the bridge components provides some insight into how the weighted and unweighted models compare for the sample bridges. The bridges considered in this analysis were those that did not have an impact damage mode controlling the inspection interval, since this damage mode is unrelated to the CR of the component. There were 11 sample bridges that had the controlling damage mode of impact for either the superstructure or substructure. Components from the remaining 49 bridges were analyzed to assess the effect of weighting. The proposed risk matrix was used to determine the inspection interval based on the risk score and the resulting OF category. Figure 4.11 presents the results of the analysis showing the calculated inspection interval for the 49 sample bridges considered, based on the controlling component and damage mode for each bridge.

Table 4.1. Inspection intervals determined from the controlling damage mode for unweighted and weighted models.

Consequence Factor	24 Months (%)	48 Months (%)	72 Months (%)
CF 2, Unweighted	18	40	42
CF 3, Unweighted	58	42	0
CF 3P, Unweighted	58	37	5
CF 2, Weighted	23	42	35
CF 3, Weighted	65	35	0
CF 3P Weighted	65	27	8

There was an increase in the number of bridges in good condition that would have a 72-month inspection interval when the weighted models were used as shown in Figure 4.11. Mainly, bridges assigned a 48-month interval changed to 72-month interval. The increased weight of the damage attributes of CR and CS reduces the risk score for these bridges, resulting in a change in the assigned interval. There was also a decrease in the number of bridges with CR 5 components that would be assigned an interval of 48 months, with those components typically changing from a 48-month interval to a 24-month interval. Additional analysis of those bridges with CR 7 is shown in Section 4.3.7.1.

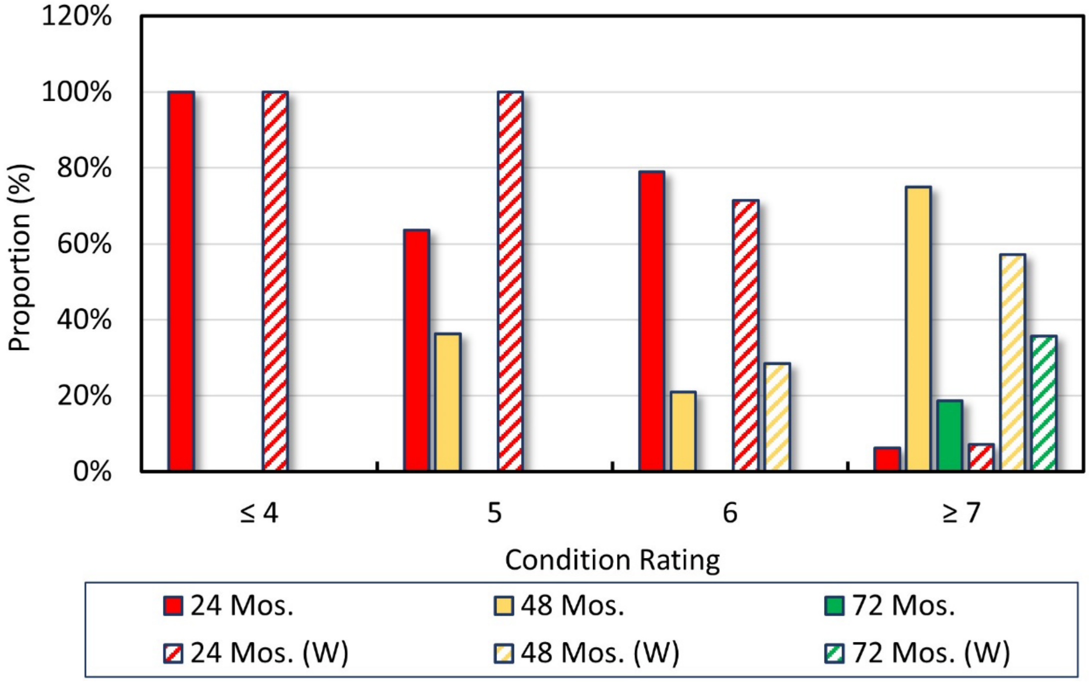


Figure 4.11. Inspection intervals determined from weighted and unweighted risk models.

4.3.7. Damage Modes

Data from the 60 sample bridges were analyzed to determine the prominent damage modes that controlled the inspection interval for a given bridge. The data presented here is for the weighted model, with CR and CS multiplied by a factor of 2, and other attributes remaining the same weight as defined by the RAP. The results are shown in Figure 4.12 which presents two pie charts showing the proportion of bridges with their inspection interval controlled by each of the different damage modes analyzed. The predominant damage modes for steel bridges (Figure 4.12 (A)) were deck delamination and spalling and substructure delamination and spalling. A significant portion of the bridges (18%) were controlled by the likelihood of impact damage due to low vertical clearance of the bridges from the roadways below. It is notable that 13% of the bridges were controlled by the fatigue cracking damage mode, while only 10% of the bridges were controlled by superstructure corrosion damage, i.e., likelihood of section loss. For PSC bridges (Figure 4.12 (B)), about 1/3 of the bridges were controlled by superstructure and substructure delamination and spalling. The deck delamination and spalling modes controlled another 25% of the bridges.

It was notable that the analysis showed that there was no dominant damage mode for the randomly selected population of 60 bridges. In fact, the damage modes were evenly distributed among the superstructure, substructure, and deck. There was a significant proportion of the bridges that had their inspection intervals based on the likelihood of impact damage due to either low clearance, in the case of superstructure impact, or location close to the roadway, for substructure impact damage. Overall, almost 20% of the bridges were controlled by either superstructure or substructure impact.

The NBIS and associated FHWA guidance allows only bridges in good condition to be considered for intervals of up to 72 months for inspection. The 13 sample bridges that were in good condition in 2020 were analyzed separately to assess those bridges that could be eligible for extended inspection intervals.

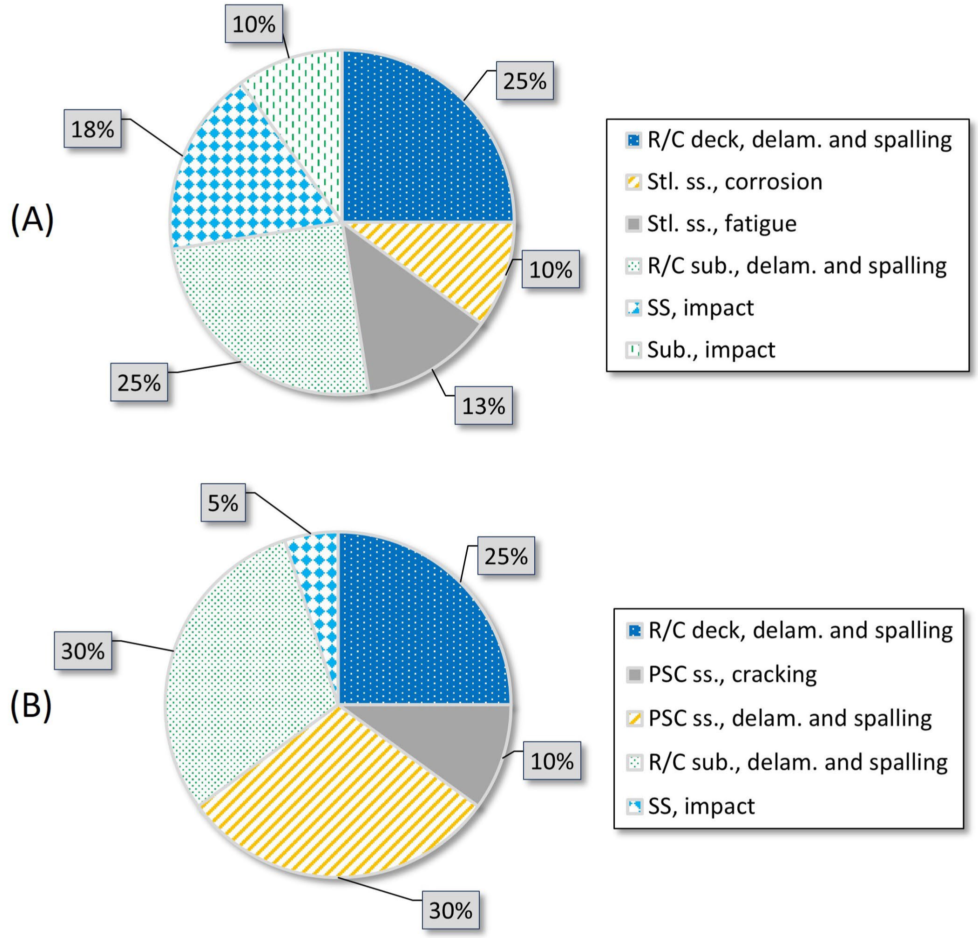


Figure 4.12. Distribution of damage modes for sample bridges showing steel (A) and PSC (B) bridges.

4.3.7.1. Bridges in Good Condition

There were 13 sample bridges that had a CR ≥ 7 for all three components of superstructure, substructure, and deck. There were seven steel bridges and six PSC bridges in this group. Three of the bridges had controlling damage modes of superstructure or substructure impact. The most common controlling damage mode was PSC delamination and spalling controlling for three out of the six bridges with PSC superstructures. Overall, there was an even and broad distribution of damage modes, with eight different damage modes controlling for sample bridges in good condition. This included R/C deck delamination and spalling, steel superstructure fatigue cracking, and superstructure impact with two bridges each, and steel superstructure corrosion, PSC superstructure cracking, R/C substructure delamination and spalling, and substructure impact with one bridge each.

Bridges in good condition with a controlling risk score of impact damage were re-analyzed without considering the impact damage mode. For each bridge, the damage mode with the highest weighted risk score other than impact damage was used in the analysis. This resulted in the controlling damage mode being one of the condition-related damage modes such as delamination and spalling. Specifically, it was assumed that one bridge was controlled by deck delamination and spalling, the second bridge was controlled by fatigue cracking, and the third bridge was controlled by substructure delamination and spalling, rather than impact damage.

The resulting damage modes were proportioned as shown in Figure 4.13. The figure shows the proportion of bridges in good condition controlled by each damage mode. The data showed that the predominant damage modes were deck delamination and spalling, PSC superstructure delamination and spalling, and steel superstructure fatigue. Delamination and spalling of the substructure also played a significant role. Overall, the results demonstrated that among randomly selected bridges in good condition, there was a distribution of the controlling damage modes divided somewhat equally between the deck, superstructure, and substructure components.

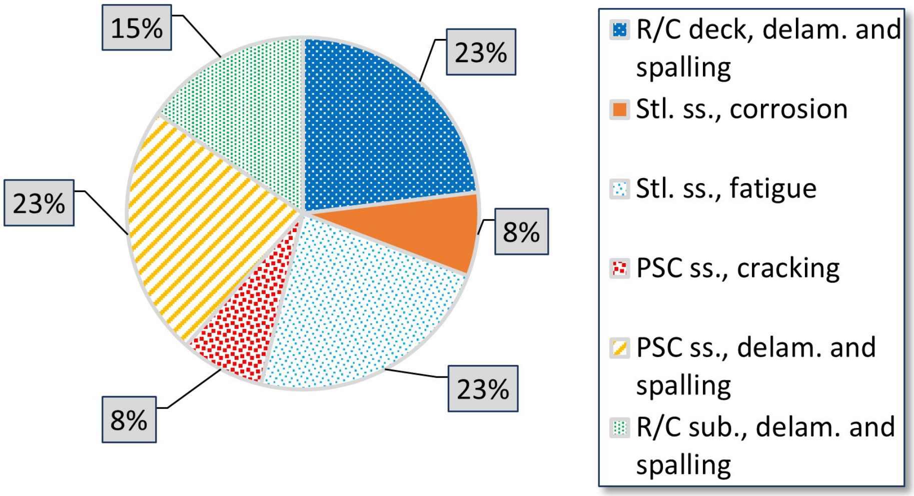


Figure 4.13. Distribution of controlling damage modes for CR 7 bridges without considering the impact damage mode.

The percentage of bridges in good condition eligible for a 72-month inspection interval is shown in Table 4.2. The results show that considering the proposed risk matrix and a CF of *high* (CF 3P), almost 50% of the sample bridges in good condition qualify for an extended inspection interval when the weighted model was used. If the CF was *moderate* (i.e., CF = 2), all the bridges in good condition qualified for the extended interval.

Table 4.2. Proportion of bridges in good condition eligible for 72-month inspection interval.

Model	CF 2 (%)	CF 3 (%)	CF 3P (%)
unweighted	100%	0	23
weighted	100%	0	46

4.4. Statistical Analysis of Component Risk Scores

The results from the back-casting process were analyzed statistically to assess the model performance as compared to the target ranges and the effect of weighting attributes. The analysis focused on the time-dependent damage modes, i.e., those related to corrosion damage. These models were selected for analysis for three reasons. First, the primary deterioration mechanism for highway bridges is corrosion, which affects all bridges in the inventory to varying degrees. Second, the corrosion risk models include the largest number of attributes, making their calibration the most challenging. Finally, risk models for damage modes such as impact damage and fatigue cracking depend mainly on characteristics like ADT, vertical clearance, or construction era. These models are typically consistent over time for a given bridge and rely primarily on engineering decision-making regarding the attributes that control the risk. For example, the likelihood of impact damage is independent of the CR for the superstructure of a bridge. Damage modes associated with corrosion are time dependent and would be expected to have increased risk scores as the CR for the component declines and the bridge ages. The analysis was conducted with the objective of analyzing if the risk scores were consistent with the target ranges for bridge components with CRs of 5, 6, or ≥ 7.

The analysis was conducted on a component level examining components in CR ≥ 7, CR 6, and CR 5 separately. Components in CR 4 or CR 3 were neglected from most of the analysis because these components would be screened from an RBI analysis. The risk scores for components with CR ≤ 4 can be seen in Figure 4.8 and Figure 4.9 which show that the risk scores for these components were similar to CR 5 components.

4.4.1. Back-Casting Results - All Components

The combined results for the components of deck, superstructure, and substructure were analyzed for corrosion-related damage mode of delamination and spalling for R/C decks, R/C substructures, and PSC superstructures, and for corrosion damage / section loss for steel superstructure components. The data set analyzed consisted of the risk scores determined from the individual risk models developed from the six different RAPs.

The results were analyzed to determine the overall distribution of results and to quantify the impact of weighting components. This method of analyzing the results provides insight into the expected results for a larger population of bridges. Additional analysis of this type will also be presented in the section of MC simulations, as previously mentioned. It was assumed in the analysis that the risk scores would be normally distributed about a mean value. In the analysis, the risk scores were sorted into bins with a range of 0.25. For example, risk scores of 1.10, 1.15, and 1.20 were counted in a bin with the range (1.00 < x ≤ 1.25). Risk scores were sorted according to the CR for the subject component. These data were analyzed for unweighted and weighted risk models.

The mean and sample standard deviation of the risk scores assigned to each CR were used to produce normal distribution plots that illustrate the distribution of the risk scores. Cumulative normal distribution curves are presented to show the proportion of a bridge inventory expected to have risk scores that fall within the OF ranges for *remote*, *low*, *moderate*, or *high*. Components in good condition (i.e., CR ≥ 7) were grouped together and components in poor condition (i.e., CR 3, 4) were neglected from the analysis.

The results for all of the components considered in the analysis are shown in Figure 4.14 (A) and Figure 4.14 (B). Figure 4.14 (A) shows the results from the unweighted risk modes. Results for CR 5, CR 6, and CR ≥ 7 are shown separately. The bar chart presents the number of risk scores (i.e., count) falling into each bin on the left ordinate. The right ordinate shows the frequency or proportion of components from a normal distribution based on the mean and sample standard deviation of the data for each CR. This axis is unscaled because the data are normalized such that the integral of each normal curve is equal to 1. The horizontal axis on the bottom shows the OF category, and the horizontal axis on the top of the plot shows the numerical values of the risk scores.

It can be observed in these results that the mean value for CR ≥ 7 bridges (the apex of the normal distribution curve) is larger than 1.0, and these data appear normally distributed. For CR 6 components, the mean value is close to 2.0, and for CR 5 bridges, the mean value is approximately 2.6. Overall, it can be observed that the trend of these data correlate with the CR, i.e., CR ≥ 7 components have lower risk scores as compared with CR 6 components, and CR 6 components have lower risk scores than CR 5 components.

Figure 4.14 (B) illustrates the effect of weighting the CR and CS attributes (C.1 and C.2) for the different components. Qualitatively, it can be observed that the risk scores for CR ≥ 7 bridges are decreased as compared with Figure 4.14 (A), and the risk scores for CR 5 components are increased. This illustrates that the overall effect of the weighting is to provide greater discrimination in the risk scores for CR 5, 6, and ≥ 7 bridges. It can also be observed in Figure 4.14 (A) that the mean value for components with CR 7 and CR 6 are in the *low* range and components with CR 5 are in the *moderate* range. When the model is weighted, the mean values for components with CR ≥ 7 is reduced to being in the *remote* range. It can also be observed that the mean value for components in CR 5 has increased to being closer to the numerical value of 3.0.

The cumulative probability distribution shown in Figure 4.15 quantifies the percentage of the components that could be expected to fall within each category. The cumulative probability graph shows the probability of a randomly selected component being ranked as *remote*, *low*, *moderate*, or *high*. The figure shows the results from the unweighted and weighted models as different line types. The weighting causes those components with CR ≥ 7 to tend toward the lower category, from *low* to *remote*. For components in CR 5, the weighting causes the curve to shift to the right, showing an increased probability that a given CR 5 component would be categorized as *high* and a reduced probability that a CR 5 component would be categorized as *low* or *remote*.

These results illustrate several important points. First, components that are in CR ≥ 7 generally score much lower than components in CR 5. This is not surprising since the CR accounts for a substantial portion of the scoring, so two components with the same attributes in the risk model but different CRs would always score differently. But more importantly, components in CR 7 do not all score in the *remote* category, only 58% of components would score in that range based on the mean and sample standard deviation from these data. As shown in the figure, 42% of the components in CR ≥ 7 were in the *low* or *moderate* category. These components were those with increased risk factors as identified by the individual RAPs, meaning that the risk models are sensitive to loading and design attributes, as well as other condition attributes such as joint condition. For example, if several of the attributes in the risk models were rated as *high*, the risk score would be in the *low* or even *moderate* range. In this way, the models are shown to have a sensitivity to the key attributes identified by the RAPs when applied to real bridges to prioritize them based on risk.

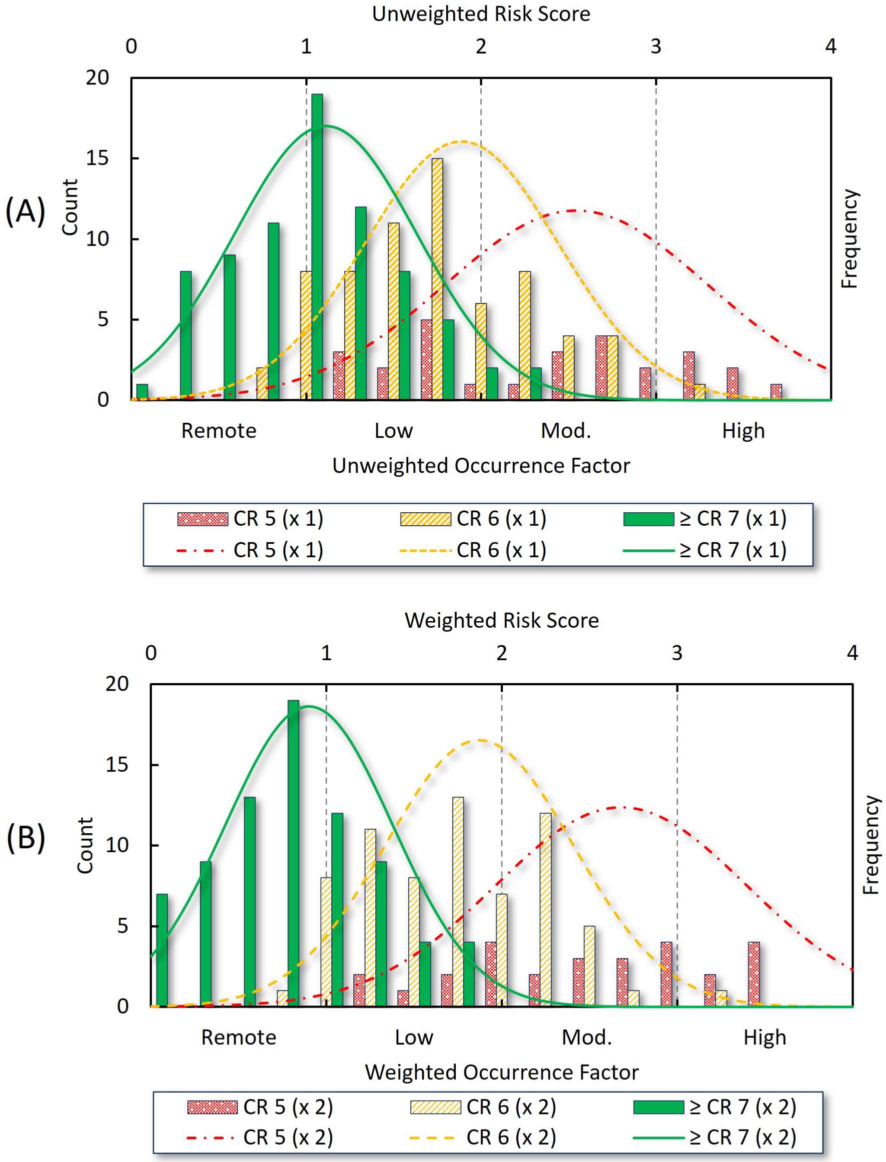


Figure 4.14. Combined results for all components showing risk scores (OF) for sample bridge components in unweighted (A) and weighted (B) models.

Considering the results of weighting quantitively (Figure 4.15) shows some important outcomes from weighting the attributes C.1 and C.2. Based on the statistics from the sample bridges, the weighting increases the proportion of CR ≥ 7 components that would be rated as *remote* from 41% to 58%, meaning a majority of CR ≥ 7 components would be rated in the *remote* range. Recalling the risk matrix shown in Figure 2.4 (B), the components with a *remote* OF could have a 72-month interval when the CF was rated as *high*. For components in CR 5, the weighting has the effect of reducing the proportion of components rated as *low* from 24% to 17%, meaning that 83% of components in CR 5 would rate at least *moderate* risk, and 33% would rate as *high*.

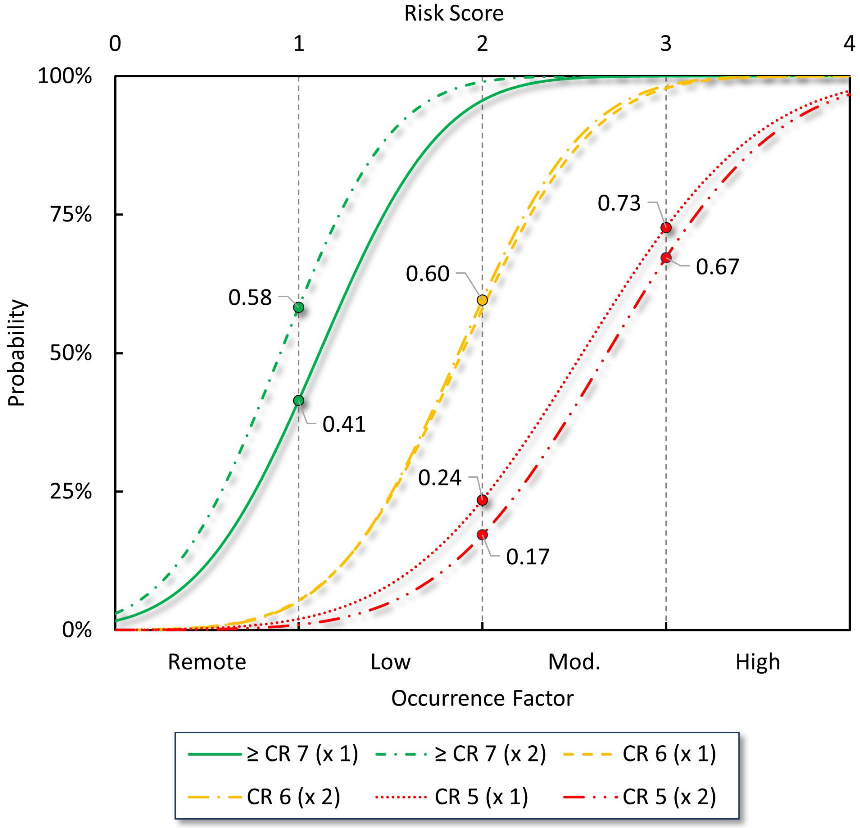


Figure 4.15. Cumulative probability distribution for all components showing results for the unweighted and weighted models.

Components that have CR 6 were essentially unchanged by the weighting. This data indicates that 60% of randomly selected components would have risk scores in the *remote* or *low* range. This is consistent with the Method 1 policy that a bridge with components in CR 6 may be eligible for a 48-month inspection interval. In fact, those CR 6 components rated as *remote* could be eligible for interval of 72-months regardless of the CF and those rated as *low* could qualify for a 72-month interval if the CF was *moderate*, as shown in Figure 2.4 (B), although the NBIS does not allow a 72-month interval for these bridges.

The results indicate the quality of the risk models was improved toward target ranges by increasing the weight of the primary condition attributes by a factor of 2 relative to the other attributes in the model. The data were analyzed similarly for the R/C decks and steel superstructures, and these data are presented below. Similar results were found for PSC superstructures and R/C substructure components. Quantitative values for the mean and standard deviation for all components combined and for deck, superstructure, and substructure components are shown in a summary table at the end of the section.

4.4.2. Back-Casting Results – R/C Decks

The analysis of R/C deck risk scores was completed using data from all 60 bridges in the sample bridge population. The results of the risk scoring from the RAP models were analyzed to assess if the weighting process used with individual components such as the deck, superstructure, or substructure was consistent with the results from all components in the study combined that was presented in the previous section.

Figure 4.16 (A) shows the results for unweighted R/C deck risk models with the risk scores presented as a bar chart and the normal distribution presented as a line plot. The results showed that components in good condition generally resulted in risk scores of less than 2.0, while risk scores for fair condition were greater than 1.0 and less than 3.0. The normal distribution curves illustrate that the mean value for CR ≥ 7 decks was in the range of *remote* (i.e., ≤ 1.0), CR 6 components were rated in the *low* range, and CR 5 components were rated as *moderate* in terms of the OF for the decks.

The results from the weighted models are shown in Figure 4.16 (B). The effect of weighting was to increase the number of CR ≥ 7 deck components that would be rated in the *remote* range and increase the number of CR 5 bridges that would be ranked in the *moderate* risk category.

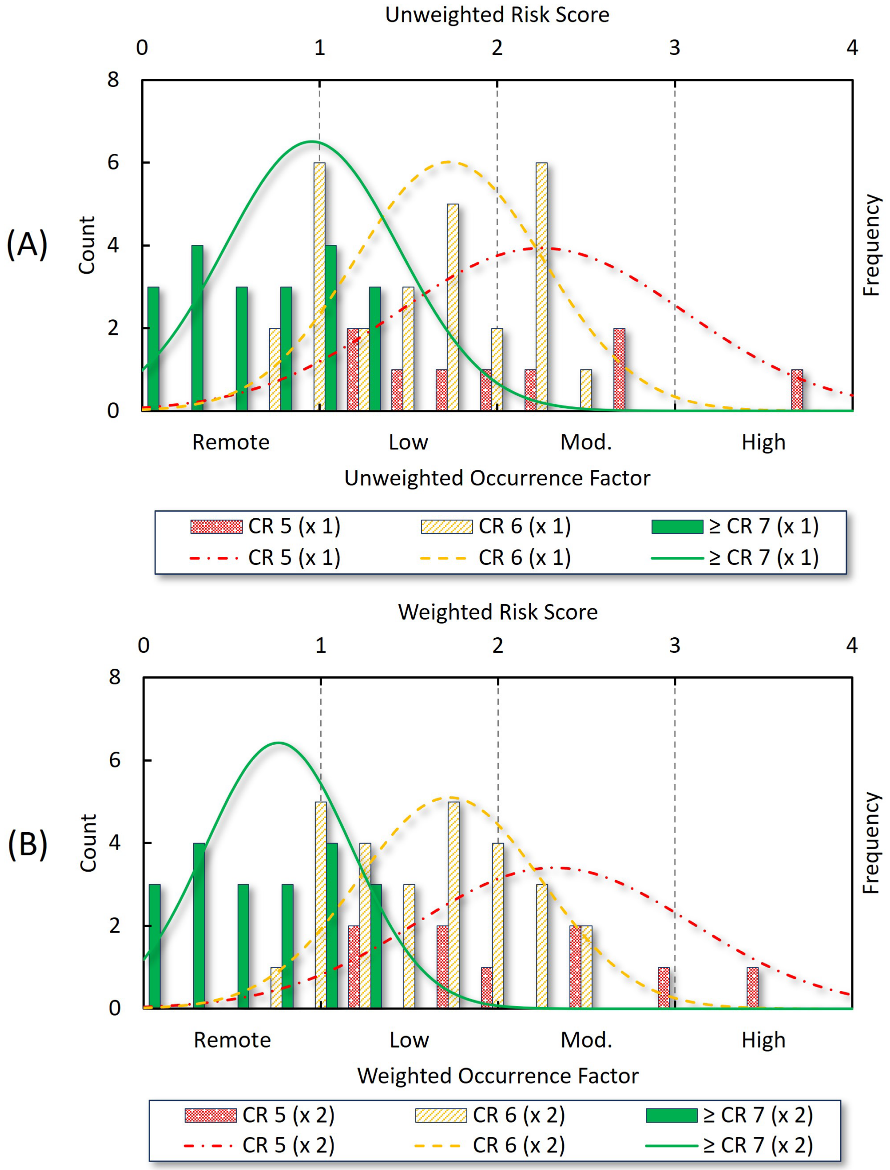


Figure 4.16. Back-casting results for deck components based on unweighted (A) and weighted (B) risk models.

The significance of weighting the risk models can be quantified by examining the cumulative distribution function based on the normal distributions shown in Figure 4.16. This cumulative distribution is shown in Figure 4.17. The effect of weighting the condition attributes was to increase the probability that a CR ≥ 7 component would be categorized as *remote* in term of the relative risk. For example, the unweighted model showed a probability of 54% of being categorized as *remote*. Using the weighted model, the probability of being categorized as *remote* increases to 72%. In other words, almost ¾ of deck components in good condition would be ranked in the *remote* category in terms of relative risk. This is consistent with the engineering judgement that most CR ≥ 7 decks are very unlikely to suddenly become CR 3 bridges. Twenty-eight percent of the decks in CR ≥ 7 would be categorized as *low* likelihood of deteriorating to a CR of 3 in the next 72-month time interval.

On the other hand, components in CR 5 are assessed to have an increased risk of deteriorating to a CR 3 in the next 72-month interval, as shown in the data. About 19% of decks in CR 5 would be categorized as *high* for relative risk of deteriorating to a CR 3 in the next 72-month interval. Only 34% of decks with a CR of 5 would likely be ranked as having a *low* or *remote* OF.

The results are based on a relatively small number of data points - only 60 decks; however, they illustrate that the RAP models were effective in ranking decks relatively and identifying those decks with elevated risk scores. The weighting of the attributes could be further optimized to improve the delineation between bridge components with elevated risk and those with minimal risk.

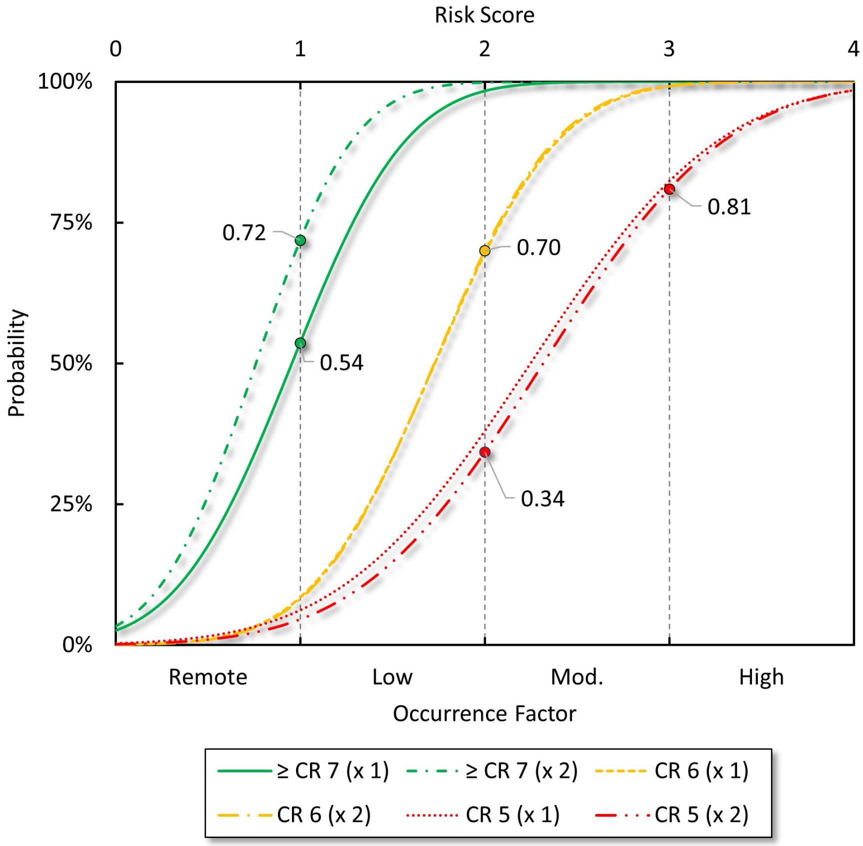


Figure 4.17. Cumulative probability distribution function for unweighted and weighted risk models for deck components.

4.4.3. Back-Casting Results - Steel Superstructure Corrosion Damage

The results for the damage mode of corrosion damage / section loss for steel superstructures were analyzed for the population of 40 steel bridges in the study. The results for the unweighted risk models are shown in Figure 4.18 (A) and the weighted model is shown in Figure 4.18 (B). Results for the steel superstructure corrosion / section loss model were similar to the results for decks and all components combined. The mean OF value for CR ≥ 7 steel superstructures in the unweighted models was in the *low* range. The mean value was reduced to the *remote* range when the risk models were weighted. The mean value for CR 6 steel superstructures was not significantly affected by the weighting, while the mean value for CR 5 steel superstructures was slightly increased.

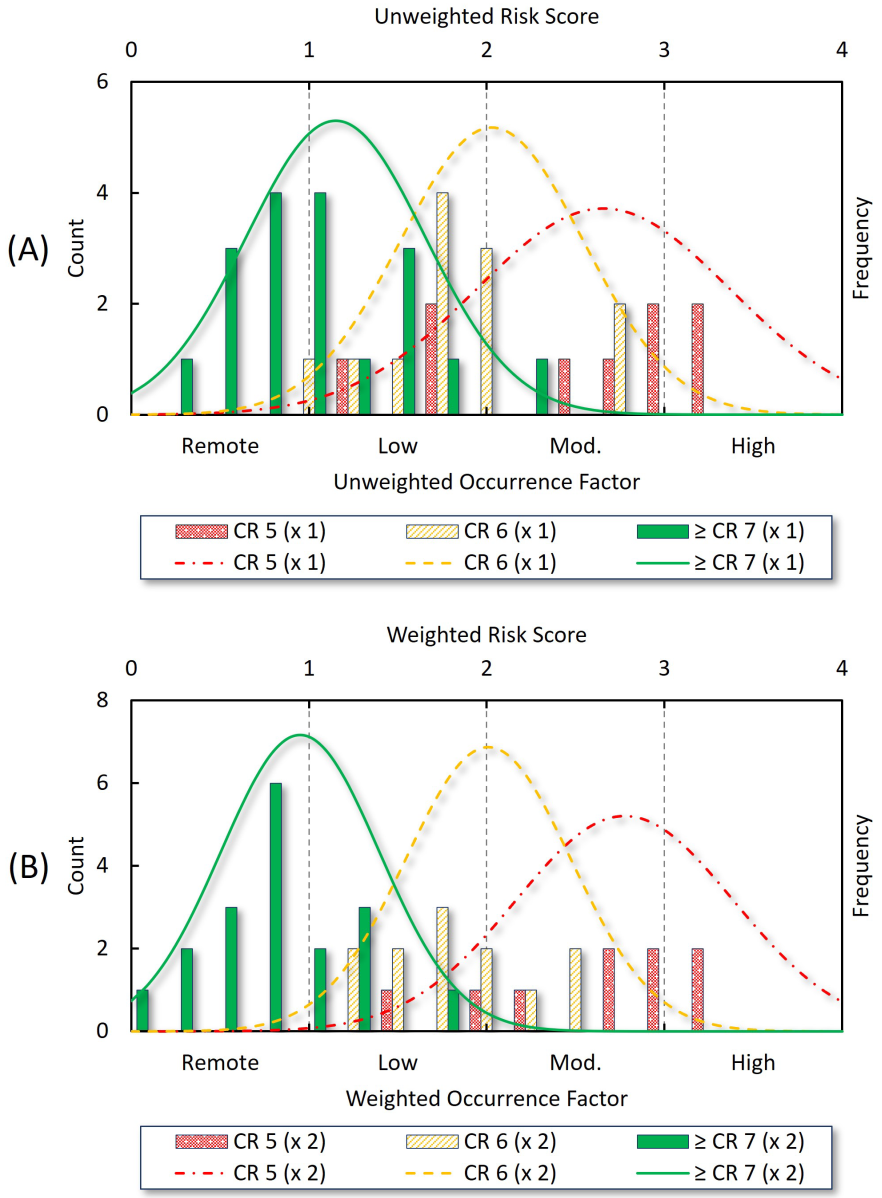


Figure 4.18. Back-casting results for steel superstructure components based on unweighted (A) and weighted (B) risk models.

The quantitative results are shown in Figure 4.19 with data labels showing key transitions between different OF categories. As shown in the figure, the weighting of the risk models increases the likelihood that a randomly selected steel superstructure in CR ≥ 7 would be rated as *remote* from 38% to 54%. The weighting also has a significant impact on CR 5 steel superstructures. In the unweighted model, the likelihood of a randomly selected steel superstructure being ranked as either *moderate* or *high* is 82%, while in the weighted model that likelihood is increased to 90%.

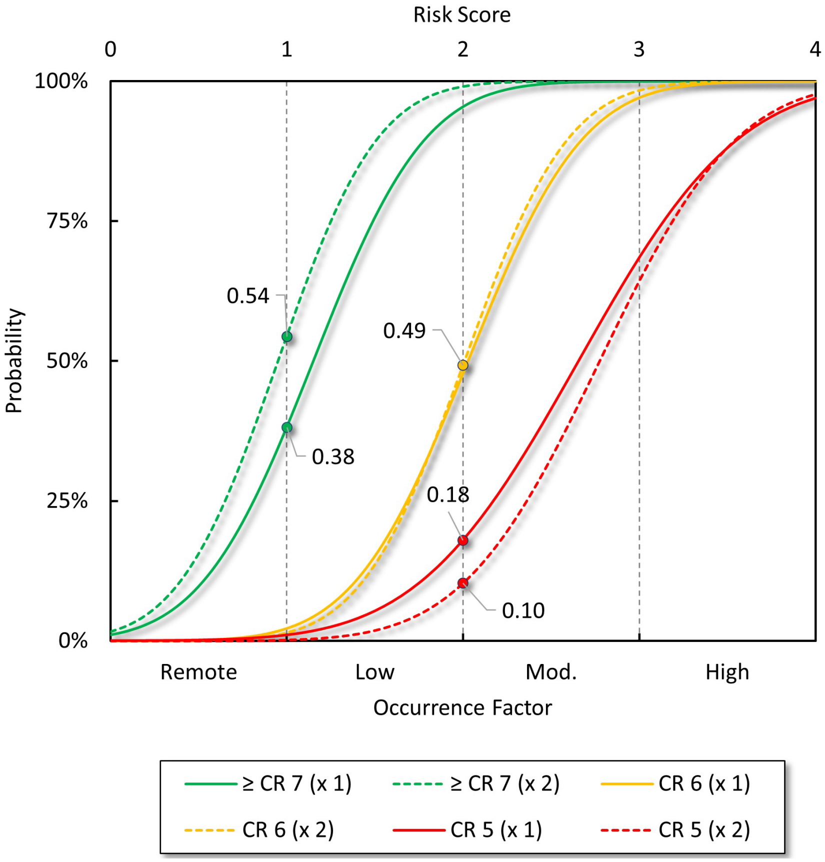


Figure 4.19. Cumulative probability distribution function for weighted and unweighted risk models for steel superstructure components.

The mean and standard deviation values for the four analyses are shown in Table 4.3 and Table 4.4. Table 4.3 shows results for all components combined into one group and for R/C decks. Table 4.4 shows the results for steel superstructures, PSC superstructures, and R/C substructures. The results are presented with the mean value above the sample standard deviation (shown in parenthesis) for CR ≥ 7, 6, and 5 components. Results from the unweighted models are shown in the first row of data and the weighted models are shown in the second row of data. The trends illustrated in the figures above are shown in the quantitative data in the table. For example, the mean values for CR 7 components are reduced by the weighting of the condition attributes of CR (C.1) and the CS (C.2), and the mean value for CR 5 components is increased. The sample standard deviations tended to be reduced for the weighted models compared with the unweighted ones.

For PSC superstructures, there was only a single CR 5 component in the sample bridge population, so statistics are not presented for CR 5 PSC superstructures. It is also notable that the PSC superstructure models indicate that most of the PSC superstructures would be in the *remote* OF rank even in the unweighted models.

Table 4.3. Table showing mean and standard deviation data for all components combined and R/C deck delamination and spalling damage mode.

Model	All Comp. CR ≥ 7	All Comp. CR 6	All Comp. CR 5	Decks CR ≥ 7	Decks CR 6	Decks CR 5
Unweighted Mean (Std. Dev.)	1.11 (0.52)	1.89 (0.55)	2.54 (0.75)	0.96 (0.49)	1.73 (0.53)	2.25 (0.81)
Weighted Mean (Std. Dev.)	0.90 (0.48)	1.87 (0.54)	2.68 (0.72)	0.76 (0.41)	1.73 (0.52)	2.32 (0.78)

The back-casting study provided data-driven analysis of the risk models developed by the RAPs applied to a population of 60 sample in-service bridges. However, the methodology of analyzing historical inspection records to assess the effectiveness of the models was time consuming and arduous. Also, the unique nature of individual bridges requires a sizable number to be analyzed to produce generalized conclusions; validating that conclusion's accuracy is challenging. To calibrate risk models to meet the target ranges, a more efficient process was sought that would provide a systematic methodology to test the risk models, assess the effect of changing the weight or number of attributes in the model, or assess the impact of different criteria used to rate the individual attributes. A systematic, data-driven method was developed to predict the outcomes from the risk models and support implementation of the RBI process and is discussed in Chapter 5.

Table 4.4. Table showing mean and standard deviation data for the corrosion damage mode for steel superstructures, and the delamination and spalling damage mode for PSC superstructures and R/C substructures.

Model	Steel SS. CR ≥ 7	Steel SS. CR 6	Steel SS. CR 5	PSC SS. CR ≥ 7	PSC SS. CR 6	R/C Sub CR ≥ 7	R/C Sub CR 6	R/C Sub CR 5
Unweighted Mean (Std. Dev.)	1.15 (0.45)	2.03 (0.51)	2.66 (0.72)	0.88 (0.54)	2.06 (0.45)	1.31 (0.49)	1.96 (0.61)	2.86 (0.66)
Weighted Mean (Std. Dev.)	0.95 (0.45)	2.01 (0.46)	2.78 (0.61)	0.66 (0.45)	2.16 (0.42)	1.09 (0.49)	1.89 (0.60)	3.02 (0.68)

Chapter 5 Monte Carlo Simulations

5.1. Monte Carlo Simulation Process

Risk models such as the one shown in Table 3.13 consider the rank of each attribute and the criteria for rating the attribute as *high*, *moderate*, or *low*. This data is used to produce a risk score (i.e., OF) using the weighted sum model described by Equation 4. The rank of the attributes provided its initial weight, with attributes ranked *high* rated and scored on a 20-point scale, *moderate* rated on a 15-point scale, etc. However, these initial weights provided by the RAP may need to be modified for the risk model to produce scores that are consistent with engineering judgement and the target ranges described in section 2.1.3. In Chapter 4, it was shown how weighting the CR and CS attributes could be used to adjust the risk scores based on the back-casting results to better match the target ranges. However, the back-casting process is challenging to implement and depends on the characteristics of the bridges selected as sample bridges. Additionally, risk models developed by RAPs can have many different attributes, attribute ranks, criteria for rating the attributes, and total number of attributes in a model. Attributes other than the CR and CS attributes may need to have their weights adjusted to optimize the model and ensure rational results are produced.

There is a challenge with analyzing how a given risk model will perform on actual bridges due to the wide array of potential attributes, criteria, and weights assigned by a RAP, and the variation in the operational environments of bridges in different states. A bridge with several attributes rated *high* may have an elevated risk score (i.e., OF), depending on the number of attributes in the model and the weights and criteria assigned by the RAP, when engineering judgement indicates the OF should be *low* or *remote*. For example, a relatively new CR 8 bridge deck with an element CS of 100% in CS 1, high ADT, and a high rate of deicing chemical application could score in the *moderate* or *high* OF range if there were only a few attributes in the risk model, and/or there was insufficient weight assigned to the CR and CS attributes. Engineering judgement, deterioration data (such as TICR data), and experience all indicate the likelihood of the deck deteriorating to CR 3 in the next 72 months is *low* or *remote* if the deck is currently in CR 8 and CS 1 and the deck is relatively new. There is an almost infinite number of combinations of attributes, ranks (i.e., initial weights), and criteria that could be identified by RAPs. Analyzing the many potential combinations through back-casting was found not to be practical or effective. A more objective and systematic data-driven approach was sought that would allow the effectiveness of the models to be analyzed and demonstrated.

MC simulation is a common method of analyzing multi-variable processes when there is uncertainty in the variables that form the input. The method uses probabilistic theories to combine the results from different input variables and provides a variety of outputs that are possible outcomes given the probabilistic characteristics of the input. The method is frequently used in risk assessment when there is uncertainty in the parameters affecting the level of risk. This approach was used to develop a methodology for analyzing the potential outcomes of the risk models developed by the RAPs. The methodology allows the user to determine weights for attributes, assess the criteria used in the risk model, assess the effect of applying the risk model to families of bridges with similar characteristics, and calibrate a risk model to produce results consistent with the target ranges described earlier in the report.

The structure of the MC simulations used in this research is illustrated in Figure 5.1. The process begins with a RAP developing a risk model for a certain component that includes attributes and criteria, shown as the RAP model in the figure. Probability distributions are then determined for each attribute to describe the likelihood of a given attribute being rated *very high, high*, *moderate*, or *low* according to the criteria from the RAP model and available bridge inventory data or engineering estimates. This data provides the input for the MC simulation. The MC simulation generates risk scores by randomly combining the attributes scores according to the inputted probabilities for each attribute.

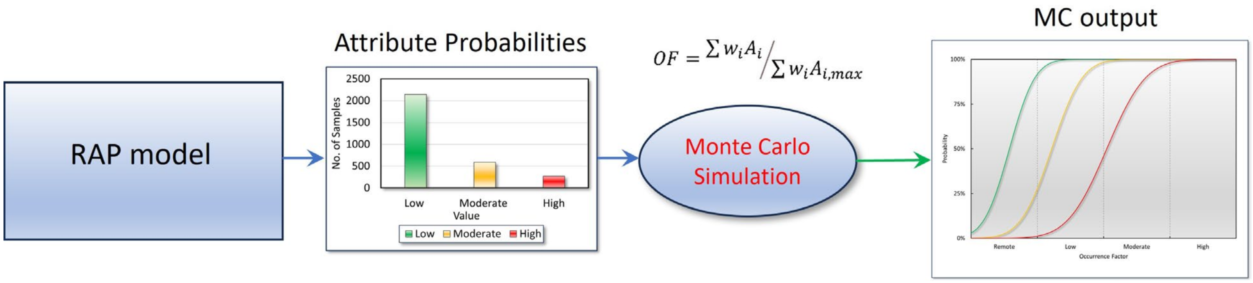


Figure 5.1. Schematic of the MC simulation process applied to a risk model.

This results in 10,000 separate risk scores for each CR (i.e., CR 7, CR 6, and CR 5). The output of the MC simulation is a probability distribution that describes likely outcomes from implementing the risk models on a family of bridges, based on the characteristics of that family as described by the inputted probabilities.

The steps to performing the MC simulations are relatively straight forward:

1. The RAP develops a risk model for a component that identifies attributes that have an impact on the POF, i.e., damage evolving to a point where a component is rated in serious condition (CR 3) during the next 72 months.
2. Criteria for each attribute are estimated by the RAP based on expert judgement. The criteria characterize the attribute's rating as *very high*, *high*, *moderate*, or *low*.
3. The probability of each attribute being rated as *very high*, *high*, *moderate*, or *low* based on its criteria is calculated or estimated for the subject family of bridges being analyzed. The estimate can be made based on available bridge inventory data, point estimation by an engineer or analyst, or by the RAP members through a Delphi process using questionnaires.
    a. If data is available from element-level inspection results, information in the bridge file, past inspection reports, or is an SNBI / Coding Guide item, determine the conditional probability based on frequency. The probabilities should consider the CR of the component, i.e., the conditional probability. For example, a deck in CR 7 is less likely to have more than 5% CS 3 damage and be rated as *high* when compared with a CR 5 deck.
    b. If data is not available for the given attribute, probabilities can be determined based on point estimates. For example, a bridge owner is unlikely to have data recording the concrete cover of bridge decks. However, engineers familiar with a bridge inventory, the evolution of construction specifications in the state, and current policies and specifications, can estimate what proportion of inventory is likely to have low cover. Precision is not required, although obviously the higher quality of the input data, the higher the quality of the output data. If the attribute probabilities are deemed critical, a Delphi process can be used to elicit expert opinion from the RAP panel or other experts in a systematic way.
4. Perform MC simulations using the risk model and the attribute probabilities to determine the mean and standard deviation of the resulting data. The MC simulations use the probability data developed in step 3 for each attribute.

5. Based on the MC simulations, construct cumulative distribution curves to present the MC outputs graphically. These curves can be used to analyze the likely outcome from applying the risk model to the subject family of bridges.

The results produced from the MC simulation were found in the research to be a powerful tool that enables several different critical tasks for developing effective risk models:

1. Calibration of the risk models to determine the appropriate weights for individual attributes to meet target ranges.
2. Comparing risk model results for components with different CRs.
3. Conducting sensitivity studies to assess the thresholds used for the criteria for each attribute.
4. Analyzing the outcome of applying the risk model to families of bridges or portions of families of bridges with similar characteristics.
5. Predict the impact of an extended inspection interval policy on bridge inventories and communicate the effectiveness of the risk models.

For example, the MC simulation can be used to show the effect of weighting the condition-related attributes of CR and CS as compared with weighting attributes like ADT or rate of deicing chemical application. The MC models also provide simple illustrations of how bridge components with different CRs compare one to another.

The outcome from applying the risk model to bridges of the same family, but with different characteristics, can also be assessed using the MC simulation approach. For example, MC simulation can be used to compare how the risk model would rate a population of bridges with high ADT as compared to a population of bridges with low ADT. The following section provides an example of MC simulation results for a bridge deck to illustrate the process and the analysis that can be conducted using this approach.

5.1.1. Example R/C Deck Simulation

The process illustrated in Figure 5.1 was used to analyze an R/C deck model to illustrate how the overall process works. Table 5.1 shows a risk model developed by a RAP for delamination and spalling in a typical R/C deck on a steel superstructure in Wisconsin. There are nine attributes identified by the RAP, including CR, CS, joint condition, etc. For each attribute, the RAP identified criteria that describe a quantity or condition that would indicate an increased impact on the POF. For example, for attribute C.2, Current Element CS, if a deck had wearing surface (Element 510) with > 10% CS 3 damage or element 12 with > 5% CS 3, it would have *high* impact on the POF.

For each attribute in the model, an estimate of the likelihood of that attribute being rated as *high*, *moderate*, or *low* was produced from either bridge inventory data or expert judgement. Most attributes were estimated from bridge inventory data. For example, considering the attribute C.2, Element CS, data for NHS bridges in the subject state were analyzed to determine the probability of a CR 7 deck on a steel bridge meeting the *high* criteria, meaning that the wearing surface element (El. 510) has more than 10% CS 3 or the deck element (El. 12) has more than 5% CS 3. Probabilities were determined for the *high*, *moderate*, and *low* criteria for deck components in CR ≥ 7, CR 6, and CR 5 as shown in Table 5.2. Calculated probabilities were obtained from a simple frequency analysis – i.e., counting the number of decks on steel bridges in CR 7 that met the *high* criteria and dividing by the total number of CR 7 decks on steel bridges. The probabilities are different for CR ≥ 7, CR 6, and CR 5 decks as would be expected. As shown in Table 5.2, CR 7 bridges have a zero or near – zero probability of meeting the *high* criteria based on historical data. Deck components with CR 6 have a 2% likelihood meeting the criteria while CR 5 decks have a more substantial 14% chance.

Table 5.1. Example deck risk model with nine attributes.

Code	Attribute	Rank	Criteria	Rating
C.1	Current CR	High	CR 5 CR 6 CR ≥ 7	High Mod. Low
C.2	Current Element CS	High	Deck (El. 510) CS3 > 10%, or El. 12 > 5% Deck (EL. 510) CS3 1 – 10%, CS2 ≥ 15%, or 1% ≤ El. 12 ≤ 5% Deck (EL. 510) CS 1 or CS2 < 15%, CS 3 < 1%, El. 12<1%	High Mod. Low
C.13	Efflorescence / Staining	High	Deck Element Soffit > 5% Deck Element Soffit 1% ≤ CS3 ≤ 5% Deck Element Soffit < 1%	High Mod. Low
L.1	ADT / ADTT	High	ADT ≥ 20,000 vpd ADT 10,000 – 19,999 vpd ADT < 10,000 vpd	High Mod. Low
L.5	Rate of Deicing Chemical Application	High	Interstate / Urban or ADT > 10,000 vpd Rural, Non-Interstate, 2,000 vpd < ADT < 10,000 vpd Rural, Non- Interstate, ADT < 2000 vpd	High Mod. Low
L.2	Dynamic Loading from Riding Surface	Mod.	Dynamic forces (ADE 9324 CS4) Dynamic forces not a significant consideration	High Low
C.7	Effectiveness of Deck Drainage System	High	Element 9004 Deck drainage: CS 3 or open rails Element 9004 Deck drainage: CS 2 Element 9004 Deck drainage: CS 1	High Mod. Low
D.26	Corrosion Protection Level	High	CP 1 CP 2 CP 3 CP 4	V. High High Mod. Low
C.29	Nondestructive Testing	High	NDT not applied NDT applied	High Low

It should be noted that this probability analysis was completed using Microsoft Excel and existing data from the NBI (https://infobridge.fhwa.dot.gov/) and the FHWA NHS element-level data (https://www.fhwa.dot.gov/bridge/nbi/element.cfm). Most bridge owners will have internal databases used for asset management that contain these data.

Table 5.2. Example probability table for attribute C.2, Current Element Condition State.

Criteria	Rating	CR 7	CR 6	CR 5
Deck Surface (El. 510) CS3 > 10%, or R/C Deck (El. 12) > 5%	High	0%	2%	14%
Deck Surface (EL. 510) CS 3 1 – 10%, CS2 ≥ 15%, or 1% ≤ R/C Deck (El. 12) ≤ 5%	Mod.	10%	23%	42%
Deck Surface (El. 510) CS 1 or CS 2 < 15%, CS 3 < 1%, or R/C Deck (El. 12) < 1%	Low	90%	75%	44%

For data that was not available from inventory or element-level data, a point estimate was used. For example, data for attribute C.13, Efflorescence / Staining of the deck soffit was not available, so the probabilities were estimated based on expert judgement. Most CR ≥ 7 decks are unlikely to have significant soffit damage while a significant proportion of CR 5 decks may have soffit damage. A conservative point estimate was made of the probability of a bridge deck meeting the *high*, *moderate*, or *low* criteria as shown in Table 5.3. It was estimated that 3% of decks rated in CR 7, 5% of CR 6 decks, and 20% of CR 5 decks may meet the *high* criteria.

Table 5.3. Probability estimate used to describe C.13, Efflorescence / Staining for bridge decks.

CR 7 (%) [H / M / L]	**CR 6 (%) [H / M / L]**	**CR 5 (%) [H / M / L]**
[3 / 7 / 90]	[5 / 10 / 85]	[20 / 20 / 60]

Estimates for each attribute were developed from inventory data or by engineering judgement. The probability values provided the input data for the MC simulations. Additional information on developing probability data for MC simulation inputs is discussed in Section 5.3.

Example results for the MC simulation are shown in Figure 5.2 (A) and Figure 5.2 (B). Figure 5.2 (A) shows the probability distribution from the MC simulations based on the risk model shown in Table 5.1. The bar chart illustrates the number of MC simulations resulting in the value represented by each column or bar. The line plot shows the probability distribution function based on the mean and standard deviation of the data represented in the bar chart. As shown in the figure, the MC simulations produce normally distributed results represented by the bar chart and modeled by the line plot. From these data, the cumulative probability distribution is determined as indicated by the arrow in the figure.

Figure 5.2 (B) shows three cumulative distribution curves produced from the data shown in Figure 5.2 (A). The cumulative probability distribution curves quantify the probability of a randomly selected deck being rated as having *remote*, *low*, *moderate*, or *high* OF based on the simulations. For example, the data shows that about 76% of CR 7 decks would be assessed as having *remote* likelihood and approximately 24% would be assessed as *low*. Approximately 31% of deck components rated as CR 6 would be rated as *remote* with most others rated as *low*. Deck components in CR 5 would be rated as *low* or *moderate*. In this way the MC results shown in Figure 5.2 (B) quantify the outcomes from the risk model being applied to a population of actual bridges with characteristics typical of the bridge population on which the analysis is based.

Components that present uncommonly high POF as compared with typical bridges are not captured by the MC simulations because their attributes do not match the “typical” values used to form the model. For example, a deck in CR 7 with more than 5% CS 3 damage would be unusual and would have an increased risk as compared with “typical” CR 7 decks. The damaged deck is captured by the risk model but is not included in the MC simulation, as will be discussed in the following section.

5.2. Identifying Components with Elevated Risk

The MC simulation is based on typical attributes qualities found in the bridge inventory and the associated probabilities such that the MC outcome reveals typical results. In this way the bridge owner can assess what the typical results would be for a given family of bridges, but *not the specific results for an individual bridge*. For example, it is unlikely that a bridge deck with CR 7 would have more than 5% CS 3 damage, as previously mentioned. But if that were the case, then the risk score may be higher than any values predicted by the MC simulation.

To illustrate this effect, consider a bridge deck with various levels of current damage or potential for damage using the risk model shown in Table 5.1. The damage in the deck is described by the condition attributes C.1, Current CR, and C.2, Element CS, and C.13, Efflorescence / Staining. The potential for damage is described by the loading and design attributes such as ADT, Rate of Deicing Chemical Application, Effectiveness of Deck Drainage, etc.

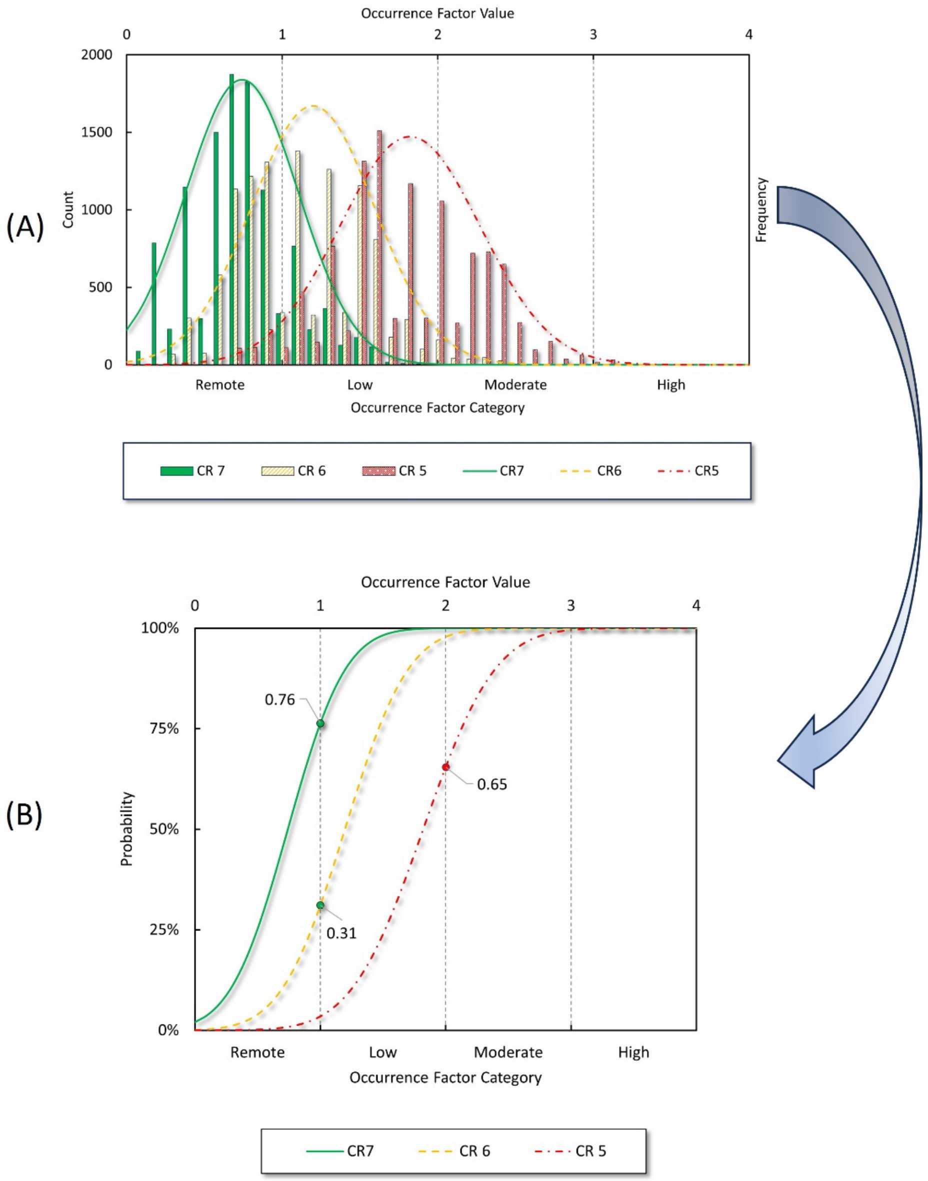


Figure 5.2. Example MC simulation results for CR 7, CR 6, and CR 5 bridge decks showing probability distribution (A) and the cumulative probability distribution (B).

Three different scenarios are shown in Table 5.4. Scenario 1 presents a deck with current damage and relatively low potential for future damage. Scenario 1 is a deck with more than 5% CS 3 damage in the deck element, damage in the soffit of the deck, low ADT, a low rate of deicing chemical application, typical corrosion protection (i.e., ECR with normal cover, CP 2), good deck drainage, and NDT testing applied to the deck.

Scenario 2 is a deck *without* current damage in the deck or soffit, but with other attributes that indicate the potential for damage is increased as compared with other decks. For this case, the CS attribute C.2 is rated as low, there is no damage in the soffit of the deck, but the deck is exposed to high ADT, high rate of deicing chemical application, poor deck drainage, and no NDT testing.

Scenario 3 illustrates a deck with *both* current damage to the deck and high potential for damage. Scenario 3 includes CS 3 damage of greater than 5% in both the deck and the soffit, high ADT, high rate of deicing chemical application, and poor deck drainage.

The resulting OF values calculated from the risk model for each scenario are listed in Table 5.4 and shown graphically in Figure 5.3. These data show a deck in CR 7 with damage (scenario 1) scores in the *low* OF range (1.33). For scenario 2, where the potential for damage is high but damage has not yet occurred, OF values are also rated in the *low* range (1.96), but any damage in the deck would push that result from *low* to *moderate*. For deck components with both current damage *and* attributes that indicate a high potential for deterioration, a CR 7 component scores in the *high* range. These values are also increased for a CR 6 and CR 5 decks.

Table 5.4. Example scenarios for decks with damage and the resulting OF values.

Scenario No.	Scenario Description	CR 7 (OF)	CR 6 (OF)	CR 5 (OF)
1	Deck with CS 3 > 5% damage in deck and soffit, no efflorescence or staining, low ADT, low rate of deicing chemical application, good deck drainage, no dynamic loading, and NDT applied.	1.33	1.69	2.04
2	Deck without deck or soffit damage, high ADT, high rate of deicing chemical application, poor deck drainage, dynamic loading on deck, and no NDT applied	1.96	2.31	3.38
4	Bridge deck with CS 3 > 5% damage in deck, soffit damage, High ADT, high rate of deicing chemical application, dynamic loading on deck, poor deck drainage, and no NDT	3.02	3.38	3.73

Figure 5.3 shows the cumulative probability distribution curves based on the conditional probabilities for each attribute for a deck in CR 7, CR 6, and CR 5. The results for the three different scenarios are shown as individual points on the figure with the ordinant values (y-axis) chosen arbitrarily to provide clarity in the figure. The points are color-coded to indicate the CR of the deck as CR 7 (green), CR 6 (yellow), or CR 5 (red). As shown in the figure, the original MC simulation that produced the curves did not predict any CR 7 bridges would score in the *moderate* or *high* range, with most CR 7 decks being assessed in the *remote* range. This is because the likelihood of a CR 7 deck having a significant amount of CS 3 damage is small given the typical probabilities for the overall inventory of bridges. However, were the deck to be atypical and have significant deck damage (i.e., scenario 1), the OF value is increased to 1.33 and the resulting OF category is *low*.

If the potential for damage is high (i.e., scenario 2), the OF value is also increased. For scenario 2, the deck OF for the CR 7 deck is 1.96 (*low*), a value greater than any predicted by the MC simulation. This is due to the low likelihood that all the attributes associated with the potential for damage would be rated as *high* for an individual deck in the MC simulation. Regardless of the likelihood of this situation, the risk model assesses the elevated risk associated with the high potential for damage.

The highest risk scores are obtained when the deck has both damage and high potential for damage (i.e., scenario 3). For scenario 3, the OF for the CR 7 deck is elevated to 3.02, indicating a *high* OF category. For all three scenarios, the OF values for CR 6 and CR 5 bridges are also elevated, as would be expected.

This example illustrates the objective of the risk model to identify the increased risk that may be present if there is atypical damage (scenario 1), atypical potential for damage (scenario 2), or both (scenario 3). In this way, the example illustrates the approach of using a MC simulation to produce expected or typical results for a family of bridges. The atypical component with unusually high damage or potential for damage can be identified because its risk score is greater than would be expected for the typical bridge represented by the MC simulation. When applied to actual bridges where a CR 7 deck would be expected to have a *remote* OF, a bridge with atypical characteristics is appropriately assessed as having increased risk as indicated by the OF being categorized by *low*, *moderate*, or *high*. The example illustrates how the MC simulation can be used to identify those bridges that present *elevated risk and require shorter inspection intervals* and those that *do not have elevated risk. This is precisely the objective of the risk analysis.*

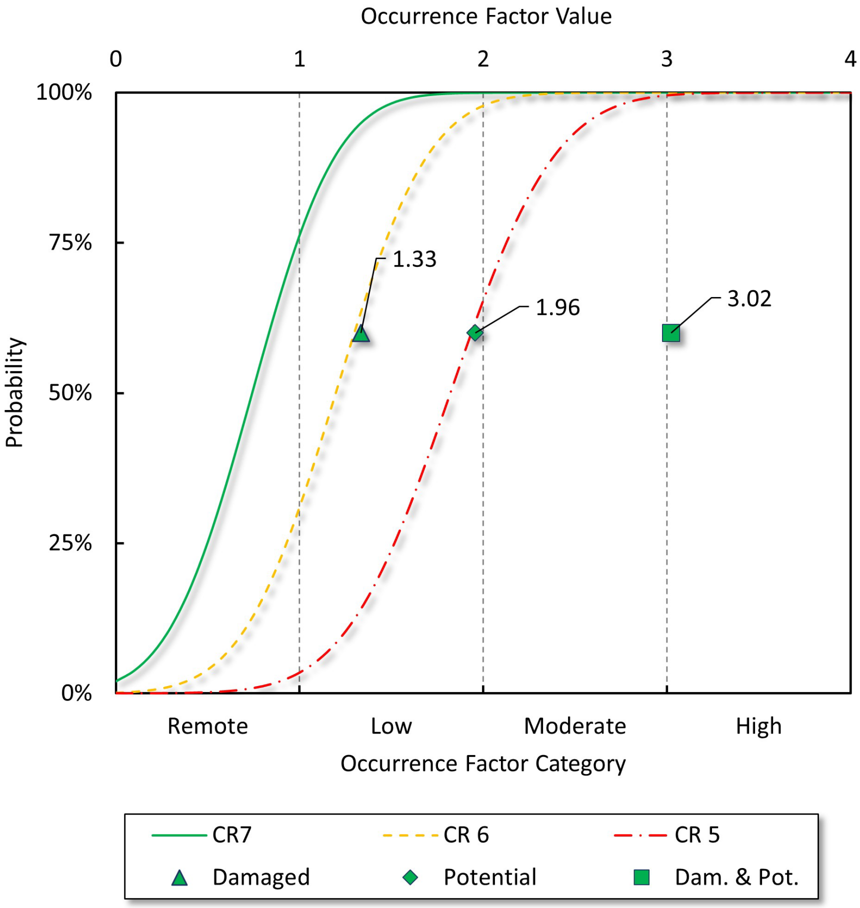


Figure 5.3. Results for CR 7 deck with different damage scenarios plotted with cumulative probability distribution from MC simulation results.

Different scenarios can be studied probabilistically using the MC simulations by setting certain probabilities at 100%. To illustrate this feature, five scenarios are considered as shown in Table 5.5. The analysis considers the original MC model with typical probabilities for attributes. Three scenarios consider increasing levels of damage with the deck element having CS 3 > 5%, both the deck and the deck soffit having CS 3> 5%, finally the deck and soffit having CS 3 > 5% and high ADT. Finally, a scenario is considered in which the attributes other than condition are rated *high*.

Table 5.5. Scenarios for probabilistic analysis of a risk model for decks.

Scenario No.	Scenario Description
1	CR 7 deck with original probabilities for attributes, deck CS 1, typical ADT
2	CR 7 deck with deck element (El. 12) CS 3 > 5%, soffit CS 1, typical ADT
3	CR 7 deck with deck element (El. 12) CS 3 > 5%, soffit damage CS 3 > 5%, typical ADT
4	CR 7 deck with deck element (El. 12) CS 3 > 5%, soffit damage CS 3 > 5%, and High ADT
5	CR 7 deck with deck element (El. 12) CS 1, soffit CS 1, high ADT, high deicing, high ponding, dynamic loading, and no NDT

The results of this analysis are shown in Figure 5.4. As the damage in the deck increases, the distribution curve is shifted to the right such that when CS 3 damage is present at a level of greater than 5% in the deck, only about 12% of decks could still be rated in the *remote* range with most decks being rated in the *low* range. If there is damage in both the deck and the soffit, most CR 7 decks would be rated as *low* with approximately 33% being rated in the *moderate* range. Finally, if the deck were exposed to high ADT, most of the CR 7 decks would be rated in the *moderate* range (approximately 67%). The figure also shows the results of having CS 1 in the deck and soffit, but attributes that address the potential for damage rated *high*. This scenario considers a deck with high ADT, high rate of deicing chemical, ponding or poor deck drainage, dynamic loading on the deck, and no NDT applied. Most decks with these precursors to damage would be rated in the *low* category (67%).

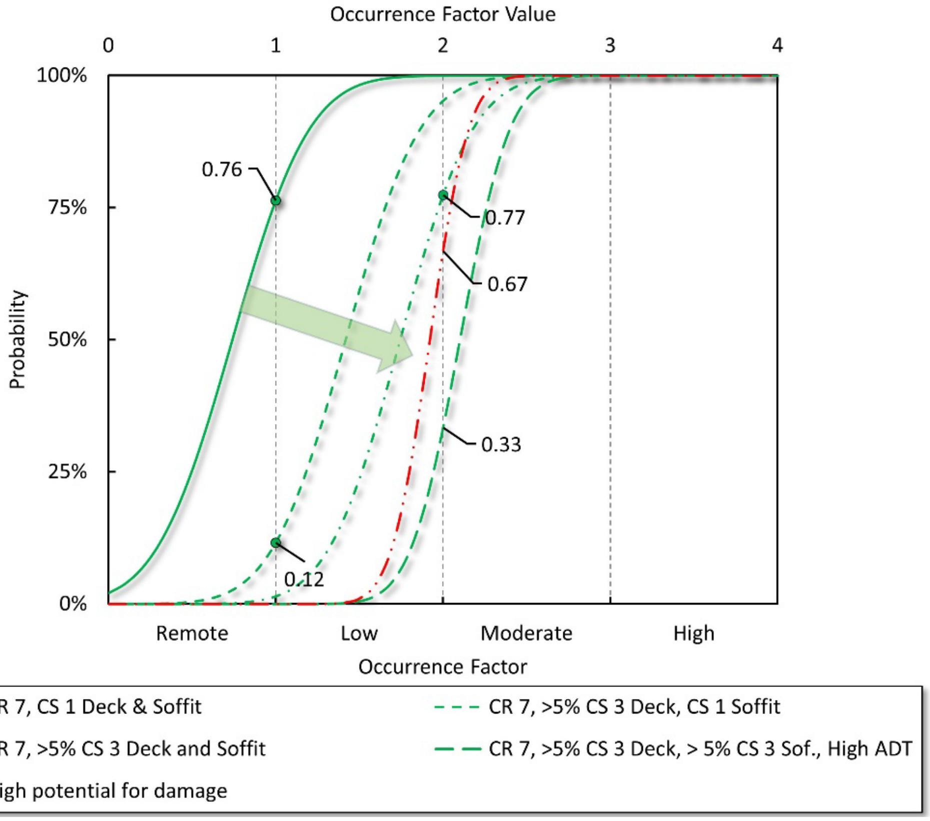


Figure 5.4. MC simulation results for decks with increasing levels of damage as shown in Table 5.5.

These examples illustrate how the MC simulations based on conditional probabilities for the attributes identified by the RAP form a model that can be used to analyze the results of applying the model to a population of bridges and test different potential scenarios. However, risk models can have different numbers of attributes, and the attributes can have different relative weights. To make the MC simulation procedure described above practically implementable, an understanding of the process's sensitivity is needed so engineers can analyze how the types, number, and weights of attributes affect the outcome.

5.3. Probabilities for MC Simulation

This section presents the application of the MC simulation process to an example risk model from Wisconsin shown in Table 5.1. The section describes the different steps in producing the MC simulation in terms of obtaining probabilities to describe the attributes and provide input for the simulation.

The different deck risk models developed by the RAPs under the project shared many common attributes such as the CR and CS, rate of deicing chemical application, and ADT. The Wisconsin model was unique in having an attribute to consider the reduced risk from performing NDT as part of the condition assessment of the deck. Performing NDT on the deck provides additional insight on the condition of the deck by detecting subsurface damage not observable in a routine visual inspection. The attribute in the risk model represents this effect by reducing the overall risk score, and the MC simulation was used to assess how including this attribute in the risk model would affect the likelihood of a CR 7 deck being rated *remote, low, moderate, or high*.

5.3.1. Assessment of Probabilities

As previously mentioned, bridge inventory data or element level inspection results can be used for most of the attributes listed in the risk models to determine the probabilities for each attribute. For situations where there is no available data for a given attribute, expert judgement can be used. In addition, the interaction or coupling of different attributes must be considered when estimating the probabilities for different criteria. Two common types of coupling are CR and CS, and ADT values used for both considering the increase rate of deterioration that typically results from high traffic, and the application of deicing chemicals based in part on the ADT level. For the former, the researchers analyzed element-level data for bridge components that were CR ≥ 7, CR 6, and CR 5, and the MC simulation analyzes these different CRs individually. For the latter, the algorithms within the MC simulation need to be appropriately adjusted to consider the coupling of the ADT attribute and the rate of deicing chemical application. For example, if bridges with ADT > 10,000 vpd are rated as *moderate* for the attribute L.1, ADT / ADTT attribute, and bridges with ADT > 10,000 vpd are assumed to have the application of deicing chemicals of *high,* the MC models must consider this interaction to produce a reliable result.

5.3.2. R/C Bridge Deck Probabilities

The analysis of RC bridge decks was conducted based on the Wisconsin risk model shown previously in Table 5.1. The risk model included nine attributes. Most of the attributes were rated on a *low*, *moderate*, and *high* rating scale. There were two attributes what were rated on a *high-low* basis. Bridge decks subjected to dynamic loads resulting from the "bump at the end of the bridge" were rated as either *high* or *low*. Additionally, there was an attribute to assess the reduction in risk from performing NDT of the deck as part of the condition assessment.

The element-level data for NHS bridges were analyzed to estimate probabilities for corrosion damage in bridge decks. The attribute C.2, Current Element Condition State included both the deck element (El. 12) and the wearing surface (El. 510). The data for NHS bridges were used to provide estimates of the probabilities of a randomly selected deck being rated as *high*, *moderate*, or *low*. The analysis utilized different criteria for the deck element and the wearing surface element as shown in Table 5.6. For the rating of *high*, the deck element had a threshold of greater than 5% in CS 3, while the wearing surface element had a threshold of greater than 10%. The table shows the resulting probabilities for CR ≥ 7, CR 6, and CR 5 R/C decks.

Table 5.6. Probability data for attribute C.2, Current Element Condition State.

Attribute	Rank	FINAL	Rating	CR ≥ 7	CR 6	CR 5
C.2 Current Element CS	High	Wearing surface (El. 510) CS 3 > 10% , or Deck (El. 12) > 5%	High	0%	2%	14%
C.2 Current Element CS	High	Wearing surface (El. 510) CS 3 1-10%, CS 2 ≥ 15%, or 1% ≤ Deck (El. 12) ≤ 5%	Moderate	10%	23%	42%
C.2 Current Element CS	High	Wearing surface (El. 510) CS 1 or CS 2 <15%, CS 3 < 1%, Deck (El. 12) < 1% CS 3	Low	90%	75%	44%

5.3.2.1. Average Daily Traffic Analysis

The ADT values for Wisconsin were determined from an analysis of the 2022 NBI data. The analysis considered state-owned bridges to provide a conservative estimate of traffic levels, since state - owned bridges would typically have the greatest number of vehicles as compared with locally owned bridges. The NBI data for Wisconsin was analyzed for bridges with steel superstructures specifically. To conduct the analysis, the NBI data was reduced to only those bridges with steel superstructures and basic configurations (stringer, stringer and floor beams, and box beams). For bridge decks, there were two attributes that considered the ADT levels in the analysis. Loading attribute L.1, ADT / ADTT, considered bridges with ADT of 20,000 vpd or more as high, and bridges with 10,000 – 19,999 vpd as moderate. This attribute represents the increased rate of damage accumulation that tends to accompany high ADT levels. The attribute L.5, Rate of Deicing Chemical Application, considers the ADT level and if the bridge is in an urban area or on an interstate roadway. Bridges with ADT > 10,000 vpd or located on interstates or in urban areas are rated as *high*. Bridges not located on interstates or in urban areas ADT of between 2,000 vpd and 10,000 vpd are rated *moderate*, and those with less than 2,000 vpd are rated *low*.

The attribute criteria for L.1 and L.5 are not independent, and therefore, the relationship between the attribute criteria needed to be considered for the MC simulation. If a bridge's ADT levels were identified as *high* or *moderate* for L.1, that same bridge would have to be rated as *high* for L.5. Additionally, the service level or functional classification of the bridge needed to be considered for analyzing probabilities for attribute L.5.

To address the requirement for L.5 that any interstate or urban roadway be rated as *high* and any bridge with ADT > 10,000 vpd also be rated *high*, a more detailed study of the NBI data was completed. To estimate the number of bridges that could be characterized as an interstate or urban bridge, Coding Guide Item 26, "Functional Classification of Inventory Route," was analyzed. The values of Coding Guide Item 26 were analyzed to identify those bridges listed as "Principal Arterial – Interstate, Principal Arterial-Other" for both urban and rural areas (codes 01, 02, 11, 12), as well as "Other Principal Arterial" and "Minor Arterial" (codes 14, 16) as "Interstate / Urban." All other functional classifications were identified as "Rural, non-Interstate." Since loading attribute L.1 would rate any bridge with greater that 10,000 vpd as either *moderate* or *high*, and loading attribute L.5 would indicate any bridge with > 10,000 vpd as *high*, any bridge that the MC simulation randomly selected as *moderate* or *high* for L.1 would necessarily be *high* under L.5. However, there are also bridges that are urban or interstates that have less than 10,000 vpd and should be ranked as *high* for L.5; these were identified through the analysis.

Table 5.7 lists the probabilities based on ADT data to assess if a given bridge should be rated according to L.1 and L.5 for the MC Simulation. A simple "if" statement was used to define L.5 based on L.1, meaning that L.5 was *high* if L1 is defined as *moderate* or *high.* Among the remaining bridges, it was determined that the likelihood of a given bridge being less than 10,000 ADT and being an interstate/urban bridge meeting the definition of L.5 *high* rank was 39%, and the likelihood of the rank of *moderate* and *low* was 24% and 37%, respectively.

Table 5.7. ADT probabilities used for the MC simulation of the bridge deck risk model.

ADT Level	Rank	Probability (%)	L.5 - L.1 (%)
ADT ≥ 20,000	High	20	-
ADT 10,000 – 19,999	Moderate	22	-
ADT < 10,000	Low	58	-
Interstate / Urban or ADT > 10,000	High	65	39
Rural, Non-Interstate, 2,000 < ADT < 10,000	Moderate	16	24
Rural, Non- Interstate, ADT < 2000	Low	22	37

5.3.2.2. Point Estimates

Probability to describe attribute criteria can be determined from existing bridge inventory data in many cases, but there are other cases where the data may not be available. For these cases, an estimate is needed to determine the probabilities. This section describes some of the estimates made in analyzing the deck risk model to illustrate estimating probabilities based on engineering judgement.

There were several attributes in the risk model that did not have data available in the NHS bridge element database. This included C.13, Efflorescence / Staining, C.7, Effectiveness of Deck Drainage System, D.26, Corrosion Protection Level, L.2, Dynamic Loading from Riding Surface, and C.29, Nondestructive Testing. For these attributes, engineering estimates were used to provide input data for the MC simulations.

A point estimate based on engineering judgement was used for the attribute of efflorescence and rust staining (C.13). The Manual for Bridge Element Inspection (MBEI) describes efflorescence with rust as CS 3, and this definition was applied here. It was assumed that the likelihood (i.e., probability) of efflorescence with rust staining of the deck soffit would vary based on the CR of the deck component. The attribute criteria indicated that the rating of *high* for this attribute was defined as having greater than 5% of the deck soffit assigned CS 3 (CS 3 > 5%). The rating of *moderate* was defined as 1% to 5% of the deck soffit (1%≤ CS 3 ≤ 5). It is unlikely that a CR 7 deck would have a significant amount of rust-stained efflorescence on the deck soffit. It was estimated that not more than 3% for CR 7 decks were likely to meet the criteria to be rated *high,* and not more than 7% were likely to be rated as *moderate.* The resulting probability vector was [3, 7, 90]. The estimated values are represented by a bar chart shown in Figure 5.5 (A). This is more likely for CR 6 bridges, but still relatively uncommon, and it was estimated that not more than 10% of CR 6 decks have a significant amount of efflorescence, and only a portion of those would have rust staining. It was estimated that 5% of the population of CR 6 components would be rated as *high* and 10% could be rated *moderate*, and the resulting probability vector was [5, 10, 85]. For deck components in CR 5, it was assumed that up to 40% of this population might have some efflorescence in CS 3, but no more than 20% were likely to be affected at the *high* level and another 20% at the *moderate* level.

For Corrosion Protection Level (D.26), it was estimated that only a small portion of the existing inventory of bridges would have low cover or bare reinforcing steel and no overlay or other corrosion protection. It was assumed that only 5% of the inventory would have CP 1. A sizable portion of the inventory is likely to have either normal cover with epoxy coated reinforcing (ECR) and be rated as CP 2, or normal cover, ECR, and an overlay, and be rated CP 3. There would only be a small portion of the inventory that would have ECR, normal cover, an overlay, and a sealer applied (i.e., CP 4). It was assumed that the CP level was not a function of the CR for the component. Therefore, the distribution shown in Figure 5.5 (B) was assumed for this attribute with a probability vector [5, 40,45,10] for CR 7, 6, and 5 decks.

Similar estimates were made for C.7, Effectiveness of Deck Drainage System, L.2, Dynamic Loading from Riding Surface, and C.29, Nondestructive Testing. Table 5.8 lists the probability values used for the analysis of the risk model. The MC simulations were conducted separately for CR 7, CR 6, and CR 5 components.

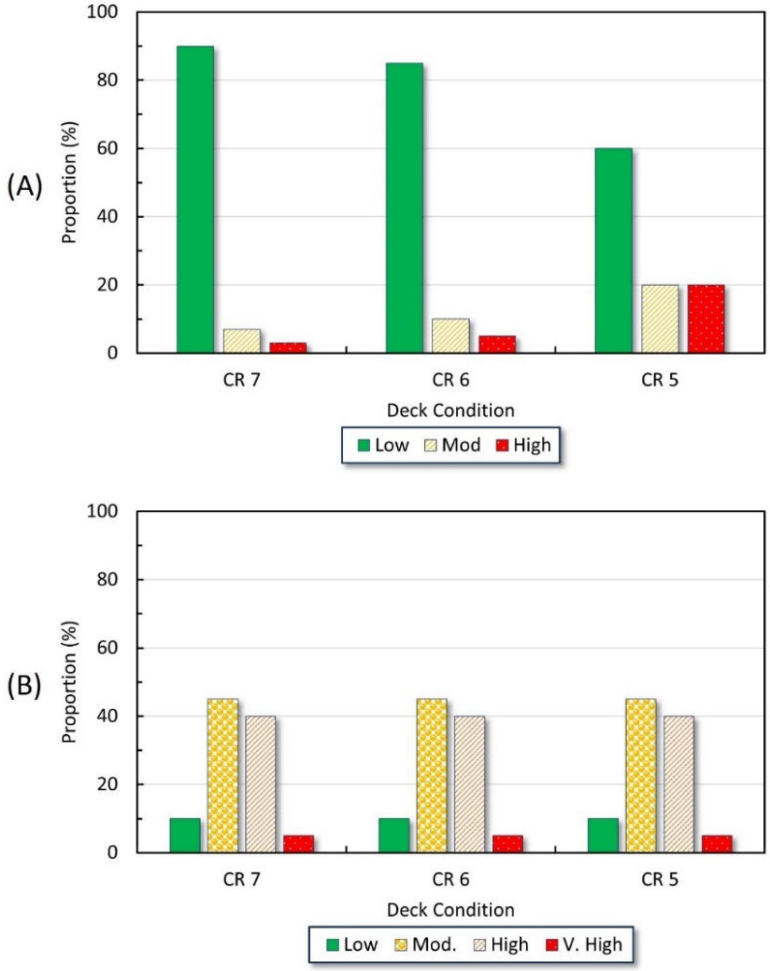


Figure 5.5. Probability distribution estimates for the attributes of efflorescence and staining (A) and corrosion protection level (B).

5.4. Weighting the Risk Models

The probabilities shown in Table 5.8 were used to perform MC simulations to determine the effect of weighting the condition attributes. The effect of weighting the condition attributes differently was studied using different potential weighting schemes and the effect on the resulting average value derived from the simulations. The condition attributes were the focus of the study for two reasons. First, the results from the back-casting illustrated that different weights for the condition attributes generally improved the quality of models based on the assumption that risk would increase as the CR for a given component decreases. Second, from a practical standpoint, it would be expected that actual damage represented by the condition attributes would have a more significant effect on the relative risk as compared with a loading attribute or a design attribute, each of which may contribute to rate of deterioration.

Table 5.8. Table showing probability values for MC Simulation for bridge decks in WI.

Att. No.	Att. Name	CR 7 (%) [H/M/L]	CR 6 (%) [H/M/L]	CR 5 (%) [H/M/L]
C.1	Current CR (fixed)	[0/0/100]	[0/100/0]	[100/0/0]
C.2	Current Element CS	[0/10/90]	[2/23/75]	[14/42/44]
C.13	Efflorescence / Staining	[3/7/90]	[5/10/85]	[20/20/60]
L.1	ADT	[20/22/58]	[20/22/58]	[20/22/58]
L.5	Rate of Deicing Chemical Application	[39/24/37][1]	[39/24/37][1]	[39/24/37][1]
C.7	Effectiveness of Deck Drainage System	[1/9/90]	[5/15/80]	[10/30/60]
D.26	Corrosion Protection Level	[5/40/45/10]	[5/40/45/10]	[5/40/45/10]
L.2	Dynamic Loading from Riding Surface	[10/90]	[10/90]	[10/90]
C.29	Nondestructive Testing	[30/70]	[30/70]	[30/70]

[1] Probabilities for L.5 for those bridges not rated as *moderate* or *high* for attribute L.1.

The MC results for the original risk model for decks are shown in Figure 5.6. The figure shows two plots. A column or bar plot shows the distribution of results from MC simulations for CR 7, CR 6, and CR 5 decks. The ordinate (i.e., y-axis) on the left shows the count from the simulation for each 0.1-width bin of data represented by the columns. The abscissa on the bottom of the plot shows the OF category, and the abscissa at the top of the plot shows the OF numerical values. The second plot shown with curves is the normal distribution for the data from the MC models based on the mean and sample standard deviation. The ordinate on the right is unscaled and shows the normalized frequency for the normal plot. The mean value for the normal distribution is the apex value for each curve. It can be observed in the figure that the mean OF values increase as the CR decreases. It can also be observed from the column plot that the MC results appear normally distributed. It should be noted that the results from the weighted sum model are not continuous because only certain values of the OF are possible when the attribute scores are summed. As a result, the appearance of the column plot depends on the width of the bins assigned to the data and how the bin width interacts with the OF values. For example, a gap appears in the *low* OF category because there are values that cannot be produced by the weighted sum model.

The cumulative distribution functions from the MC simulations are shown in Figure 5.7. The figure shows the distribution function for the unweighted risk model and a model that is weighted by multiplying the CR and CS by a factor of 2. For decks with CR 7, the unweighted model estimates about 61% of CR 7 decks would fall in the *remote* category for the OF. When the model is weighted, the likelihood of a CR 7 deck being rated in the *remote* category was increased to 76%.

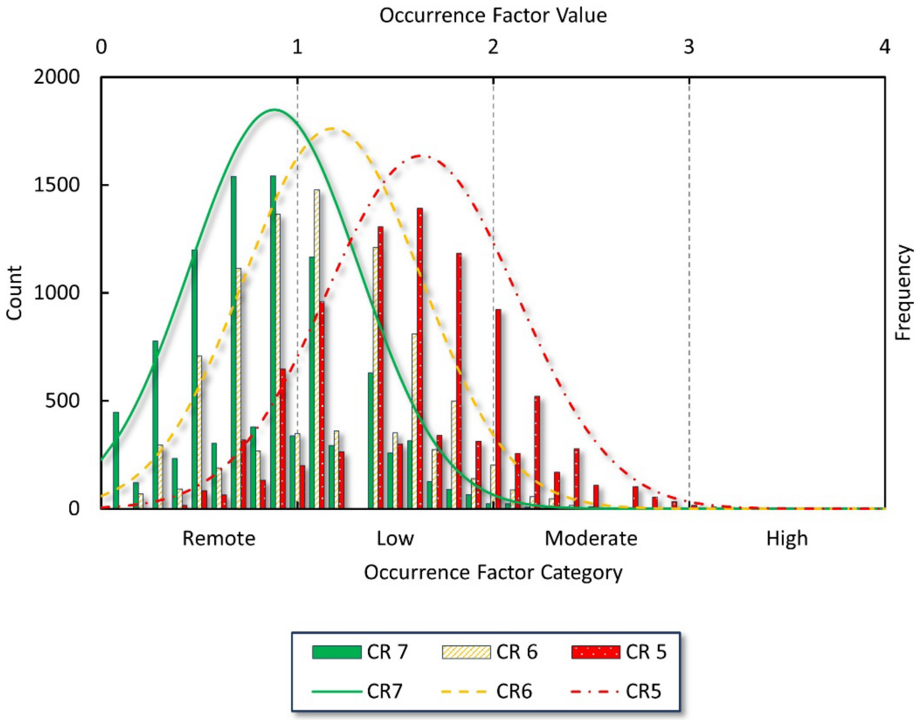


Figure 5.6. Results of MC simulations for decks with CR 7, CR 6, and CR 5.

When the model is weighted, it has the effect of shifting the CR 7 data to the left and the CR 5 data to the right as shown in Figure 5.7. When the attributes C.1 and C.2 are increased from 20 points to 40 points, the total number of points in the model is increased by 40 points. Since components in CR 7 score 0 points for the CR attribute, the proportion of the available points scored by a CR 7 component is reduced. Since the proportion of the CR attribute is increased in the weighted sum model, CR 5 components scored higher, and therefore, the cumulative distribution curve shifts to the right. This is significant because it provides a means of calibrating the model that applies not only to the CR and CS attributes, but to any other attribute as well.

To illustrate the effect of increasing the weights of individual condition attributes on the overall results of the risk model, different weightings of CR and CS attributes were considered. Although any of the attributes in the model can be weighted in a certain way to better represent engineering judgement and to meet the target ranges, the condition attributes of CR and CS link the risk models to the standard methods of condition assessment that are widely understood. Additionally, the sensitivity studies of the back-casting data indicated weighting these attributes improved the quality of the model when compared with the target ranges, as previously mentioned. To assess the effect of weighting CR and CS, an MC simulation model was prepared with different weighting scenarios. The weighting scenarios included the original, unweighted model and models with the attributes C.1 and C.2 multiplied by 1.5, 2.0, 2.5, 3.0, 3.5, and 4. The results are shown in Figure 5.8 which illustrates how increasing the multiplier for C.1 and C.2 affects the mean value of the MC simulations for deck components based on the risk model.

These data demonstrate how increasing the weight of the condition-related attributes reduces the mean value of the risk model for bridge components in CR 7 and increases the mean value for components in CR 5. Components in CR 6 change only a small amount. These data, along with parametric studies of the

MC simulation process presented later in the report, provide guidance to users on how to calibrate the risk models to be consistent with engineering judgement and the target ranges described in section 2.1.3.

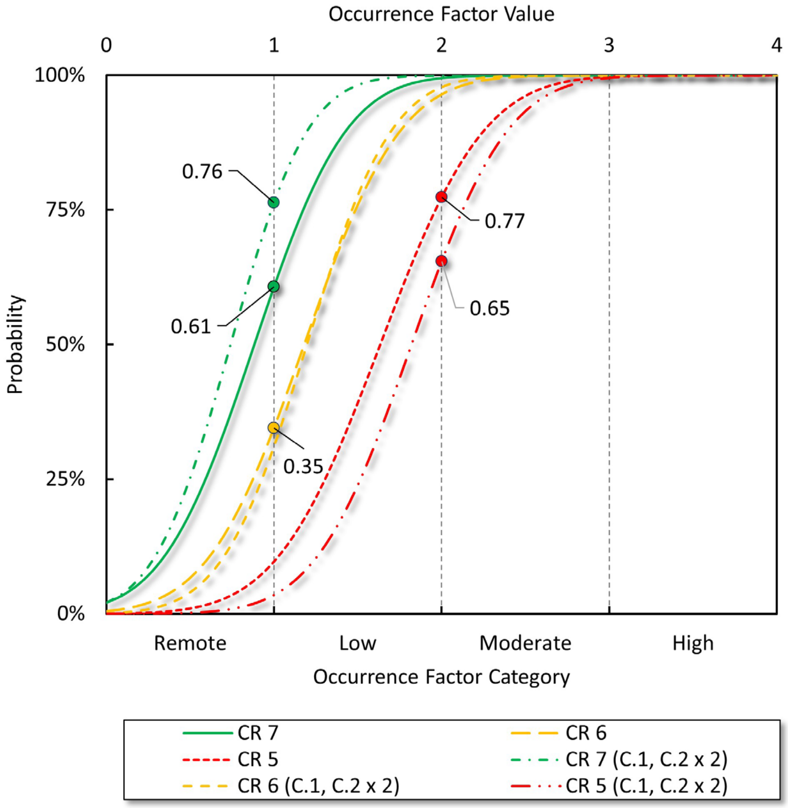


Figure 5.7. Cumulative distribution function based on MC simulations for unweighted and weighted models.

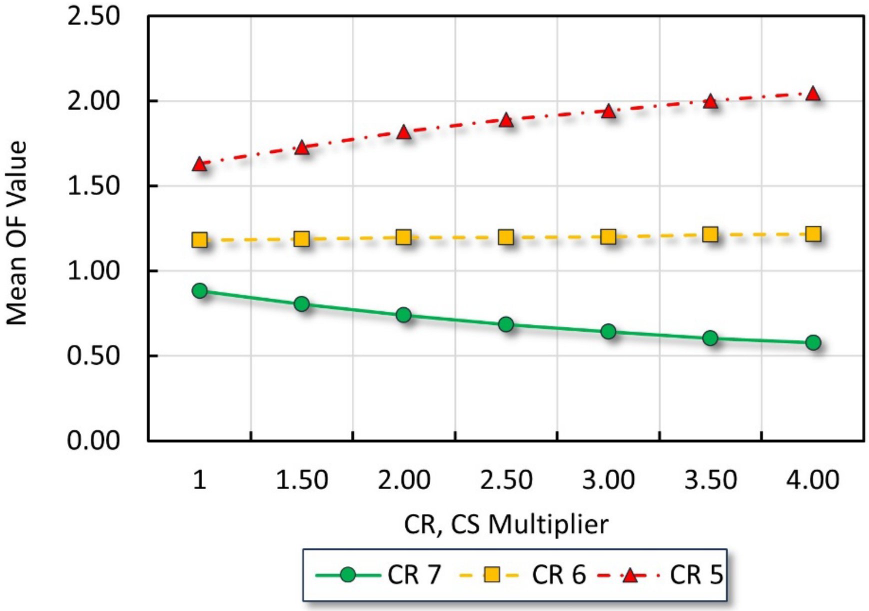


Figure 5.8. Effect of weighting on the mean values of the OF for CR 5, CR 6, and CR 7 deck components.

5.4.1. Effect of ADT on Example R/C Deck Model

An important question for most engineers would be how high ADT bridges would compare with low ADT bridges when the risk model is applied. This depends on if the risk model includes an ADT attribute, one with criteria that depends on ADT, or one correlated with ADT. The deck model has an ADT (L.1) attribute and an attribute with criteria that depend on ADT, L.5, Rate of Deicing Chemical Application, as previously described.

To analyze the effect of high ADT, MC simulations were completed assuming that all the bridges had an ADT of greater than 20,000 vpd to represent high ADT bridges. A second MC simulation assumed that all the bridges have ADT of less than 10,000 vpd to represent low ADT. In this way the effect of high ADT on model results can be quantified. The results are shown in Figure 5.9, which shows the cumulative probability distribution for decks with high ADT and low ADT. These data can be compared to Figure 5.7 weighted results, which represents the expected results for decks overall based on ADT data from the NBI.

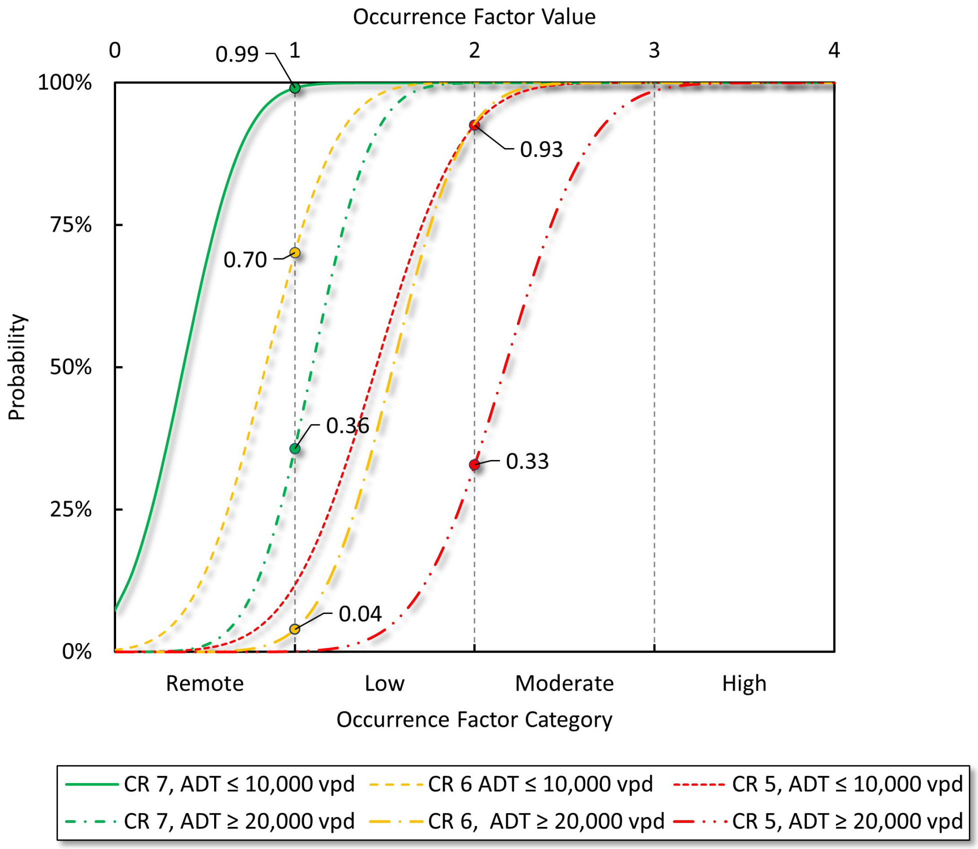


Figure 5.9. Cumulative distribution of results from MC simulations for low and high ADT bridges.

As the data shows, for high ADT bridges, only 36% of CR 7 components would be rated as having *remote* likelihood, while for low ADT bridge, 99% would be rated as having a *remote* likelihood. Recall that based on the proposed risk matrix, components rated as *remote* may be eligible for a 72-month interval when the CF is *low*, *moderate*, or *high* (CF 1, CF 2, or CF 3). Components rated with a *low* OF are only eligible for a 72-month interval if the CF is *moderate* or *low* (CF 1 or CF 2).

Most CR 6 bridges with high ADT would be rated as having a *low* OF, meaning that if the CF was high, the assigned inspection interval would be 48 months. This fits with the NBIS requirements for Method 1, which states that components in CR 6 can have a 48-month interval regardless of the ADT level on the bridge.

High ADT reduces the likelihood of the deck being rated as *remote*, meaning that when other attributes are increased (for example, the deck has poor drainage or soffit damage) the OF would be *low* rather than *remote*. Importantly, the model shows that simply having high ADT would not prevent a deck from being rated as *remote*, in fact, about 40% of decks would still be rated as *remote* considering both CR 6 and CR 7 decks.

This result indicates that the risk model is sensitive to the effect of increased ADT on the deterioration pattern of decks, with increased ADT reducing the likelihood of a particular deck to be rated as *remote* as compared with the overall population of decks (Figure 5.7).

5.4.2. Application of NDT

The application of NDT technologies for RBI has the assumed effect of reducing the uncertainty in the condition assessment for a given component. Practically speaking, NDT technologies are primarily applied to bridge decks of in-service bridges. While other components are sometimes subjected to NDT, such as steel members with section loss or a potential for cracking, these applications are not widespread. Bridge decks are most commonly tested with technologies such as infrared thermography or ground penetrating radar used to assess the condition and potential for future damage, respectively.

The Wisconsin deck model included an attribute to consider if a given bridge had been subjected to NDT. Wisconsin currently has a policy to assess bridge decks with Infrared Thermography. It was assumed for the previous analysis that 70% of bridge decks in the inventory had been assessed with NDT. The rank of the attribute was *high*, meaning 20 points was assigned to any bridge that was <u>not</u> assessed with NDT, raising its relative risk score as compared with a component that had undergone NDT. This parameter for NDT was analyzed for two purposes. First, to assess how the inclusion of NDT affects the model in terms of overall results, and second, how a different ranking affects the outcome of the risk model. For example, if the weight of the NDT attribute was 10 points instead of 20 points.

To illustrate the impact of having an NDT attribute, the MC simulation was conducted with the weighted model (CR, CS x 2) assuming 90% of the decks in the inventory were assessed with NDT. A second simulation was conducted assuming only 10% of the decks were assessed with NDT. The overall results are shown in Figure 5.10 that shows the cumulative probability distribution for the two scenarios. The difference between the two scenarios is significant because many more CR 7 decks would have remote likelihood if NDT was applied compared with decks without NDT. When very few decks (10%) have been assessed with NDT, then only 55% of the decks would be ranked as *remote,* while if 90% of the decks had NDT, 83% of the decks would be rated in the *remote* category. Regardless of the application of NDT, decks in CR 7 fall primarily in the *low* to *remote* range, aligning with the target ranges.

The second question would be how the weight of the NDT attribute affects the results from the overall model. For example, if the RAP had ranked the influence of NDT to be *low*, then that attribute would only be assigned 10 points, and therefore have less of an overall effect on the model. The results of the analysis with the NDT attribute ranked *low* and therefore assigned only 10 points is shown in Figure 5.11. As shown in the figure, the proportion of the CR 7 bridges that would be categorized as *remote* would be 69% when only a few bridges (10%) are subjected to NDT and 82% when most of the decks (90%) are assessed with NDT. These data illustrate the smaller effect on the risk score of a particular attribute being ranked *low* rather than *high* by the RAP.

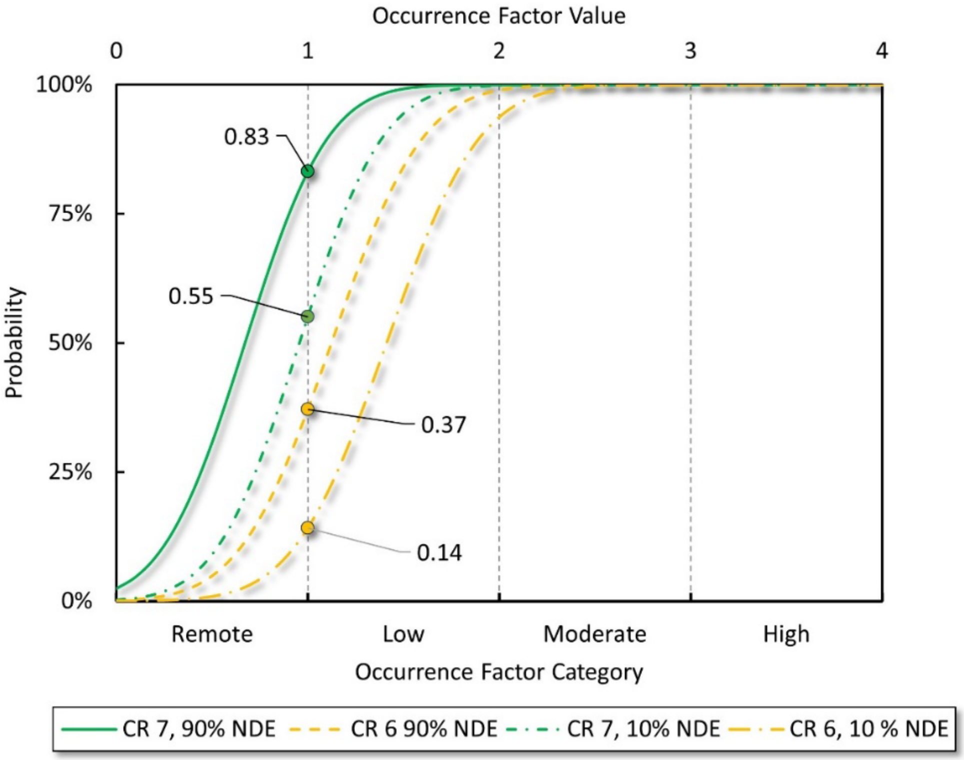


Figure 5.10. MC simulation results showing effect of NDT on OF values.

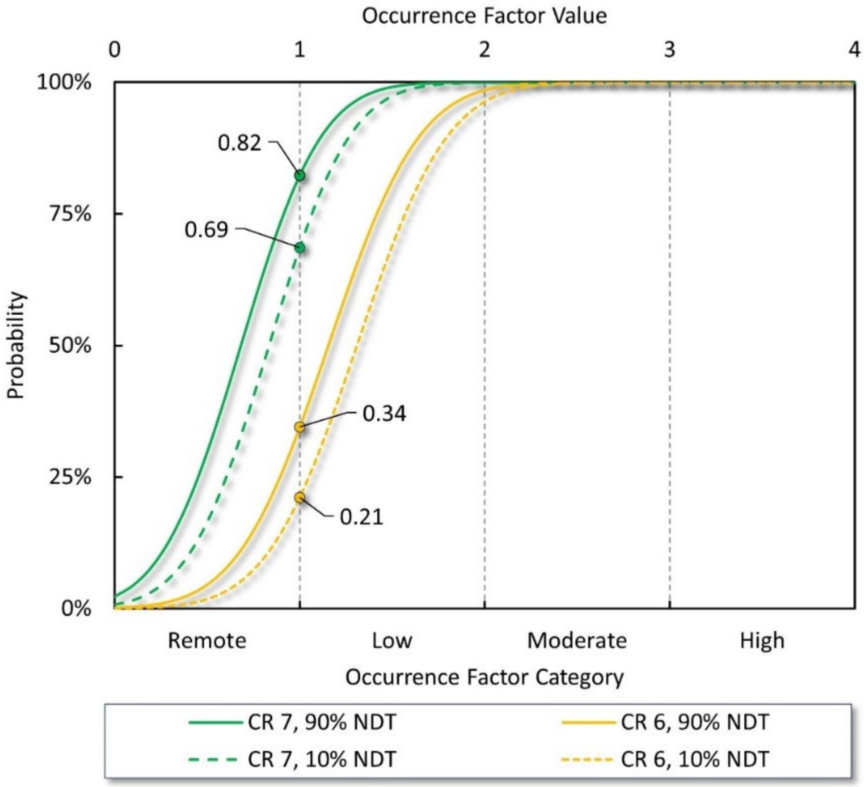


Figure 5.11. MC results for 90% NDT and 10% NDT with NDT at 10 points.

5.4.3. MC Simulation of Fatigue Cracking Damage Mode

A model for the damage mode of fatigue cracking was shown in Table 3.14. The risk model has only six attributes, including the CR and CS and attributes related to the fatigue category of details on a bridge, the year the bridge was constructed, and the ADT/ADTT level for the bridge. These attributes were common to most of the fatigue cracking risk modes developed by RAPs in the study. The risk models for fatigue presents a challenge because there are very few attributes, and the presence of fatigue-sensitive details and year of construction could result in an elevated risk score for a very low ADT bridge with a low likelihood of fatigue cracking based on expert judgement. Therefore, this risk model for fatigue cracking was examined by MC simulation to assess if the model would meet the target values and produce a reasonable result when applied to a bridge inventory. Inventory statistics and point estimates were used to estimate the likelihood of each attribute. This included an assessment of the criteria for ADT presented in Table 3.14 that indicates that ADT greater than 10,000 vpd would be rated *high*. It was found that 60% of the steel bridges in the subject bridge inventory would be rated as *high* based on the ADT values chosen. MC simulations were used to assess the effect of modifying the ADT criteria to be based on the ADT statistics for the subject inventory rather than fixed values.

The MC simulation produced results as shown in Figure 5.12. The figure presents results for the original risk model criteria as shown in Table 3.14. The MC simulation results indicated that a randomly selected bridge would have likelihood of about 11% of being rated with a *remote* OF, and a 21% likelihood of being assessed as *moderate*. Most of the CR 7 bridges would be assessed as having a *low* OF. These data somewhat match the target values, although a very small percentage of steel bridges would be rated as *remote* for the fatigue cracking damage mode.

However, the criteria assigned of 10,000 vpd being rated as *high* represents a relatively low number of trucks per day. For example, if the average proportion of truck traffic is 7% of ADT, then only 700 trucks would be traversing bridges with an ADT of 10,000 vpd. Since fatigue cracking is driven by truck loading rather than passenger vehicles, increased thresholds for rating the ADT attribute were explored. To develop rational thresholds for the ADT attribute, statistics on the ADT levels for steel bridges in the subject bridge inventory were studied. The criteria for rating the ADT attribute based on inventory statistics assigned the 75th percentile of ADT to the rating of *high* and the 50th percentile of ADT to the rating of *moderate*. Bridges with ADT less than the 50th percentile were assigned the rating of *low*.

The results of using these adjusted criteria for the ADT attribute are shown in the figure and marked as CR 7 adj., CR 6 adj., and CR 5 adj. Making this adjustment to the model increases the likelihood of a steel superstructure being rated in the *remote* or *low* OF category to 90%.

As an example of scoring a specific steel bridge, a CR 7 steel superstructure with high ADTT (> 75th percentile), Category D fatigue details, and built before 1975 was scored. This component was assumed to not have any CS 2 (arrested) cracks, and secondary members were bolted connections. This risk score for this steel superstructure is shown on the figure as an asterisk falling in the *moderate* OF range (note the vertical axis value is arbitrary for this data point). With this rating of the OF, a maximum of a 48-month inspection interval could be achieved. If the fatigue category of details were Category C rather than Category D, the risk score would be reduced to the *low* range. Different combinations of attributes would produce different results, obviously, but these data indicate the risk model reflects the target ranges and identifies a component with elevated risk. Specifically, a steel bridge built before 1975, when fatigue resistance requirements were first introduced in specifications, with high ADT and fatigue sensitive (Category D) details, had an increased risk as compared with a bridge low ADT or with Category C details.

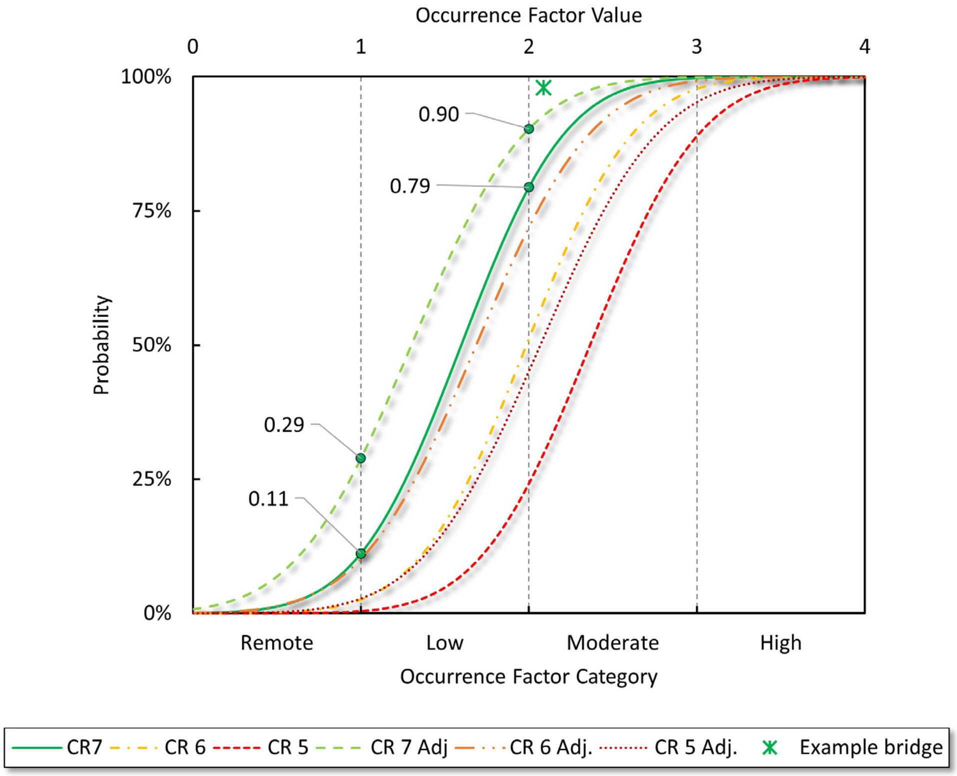


Figure 5.12 Example fatigue cracking damage mode MC simulation showing different ADT / ADTT thresholds and an example bridge with OF = moderate.

Some RAPs may consider that any structure that has arrested cracks should be screened from the RBI procedure or given more weight in the model. When the same adjusted model shown in Figure 5.12 was simulated without the attribute of cracking, the resulting curve is shown in Figure 5.13. For this model, the likelihood that a steel superstructure would be assessed as having *low* or *remote* OF is 78%. The result for the same example bridge with Category D details and built before 1975 is plotted as an asterisk in the figure showing a rating of *moderate*, though the OF value is greater when compared with the data shown in Figure 5.12. Because the risk model shown in Table 3.14 has very few attributes, a different approach of using a categorical model could also be used in lieu of scoring the risk model attributes.

5.5. Conclusions from MC Simulation

The results in this chapter illustrate how the MC simulation can be used to analyze risk models developed by an RAP. Using available bridge inventory and condition data, the potential outcomes from the risk model can be simulated. It was shown that the MC simulation can be used to demonstrate that the risk models will identify components with elevated risk and illustrate the effect of increasing damage on the risk model outcomes. The MC simulation approach can also be used to adjust the weights of the attributes, analyze the results, and compare the results with the target ranges. In this way, the criteria and attribute weights can be adjusted using available data, and the impact of adjusting the criteria and weights can be analyzed and compared with engineering judgement and the target ranges provided in section 2.1.3. It was also shown how the MC simulation approach can be used to assess the potential outcomes of the risk model when applied to portions of a bridge inventory with particular characteristics, such as high ADT bridges, or bridges subjected to NDT. The methodology based on MC simulations can provide an effective tool for analyzing risk models and communicating their effectiveness.

Probability data associated with the risk models from all six RAPs was developed during the research. These data provide inputs, along with point estimates (as needed), for the MC simulation of the individual risk models. This data is provided in Data Supplement C, *Probability Tables*.

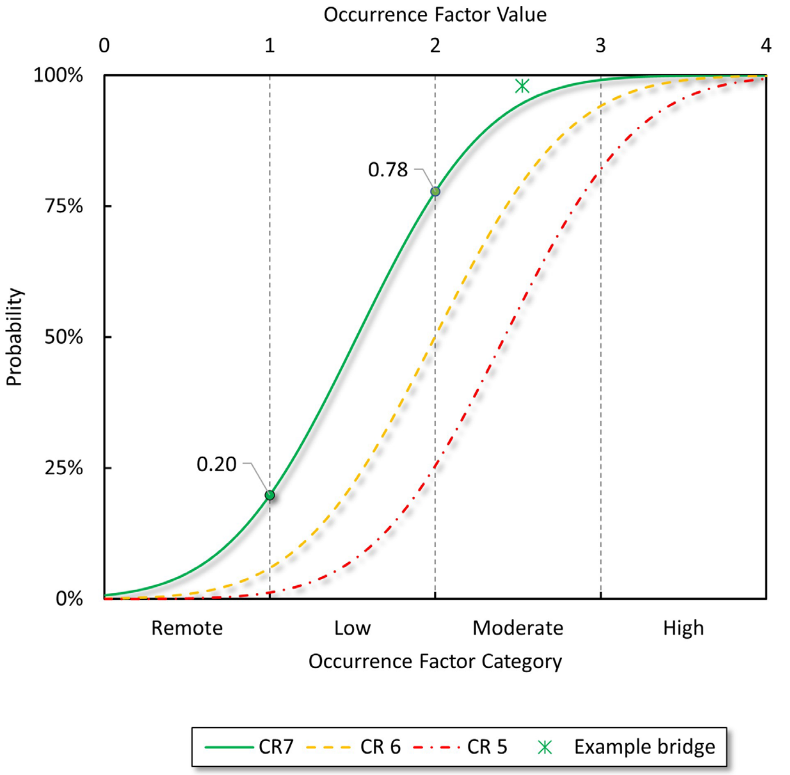


Figure 5.13 Fatigue cracking risk model without a CS attribute.

Chapter 6 Consequence Factor

This section of the report will describe the CF used to estimate the expected consequences of serious damage (CR 3) in a bridge component. The CF is a qualitative categorization that evaluates the outcome resulting from a bridge element failure (i.e., CS 3) due to a given damage mode. This factor depends on the specific scenario of the damage mode, component type, and the service level of a bridge.

The CF is divided into four levels: low, moderate, high, and severe intended to consider the bridge's safety and serviceability. Table 6.1 shows a general description of each level of the CF from the NCHRP 782 report. This table provides the framework for the CF, but specific criteria were not provided. A series of different scenarios that linked certain damage modes with attributes that should be considered were provided that further describes the framework for specific damage modes such as spalling of R/C decks and substructures, and loss of load capacity in primary members. These tables describe example situations for *low*, *moderate*, *high*, and *severe* CFs, and the parameters suggested for assessment. For example, parameters to assess the CF for a deck might include the ADT for the deck and/or the feature under the deck, the level of service of the deck (feature carried), and if the deck has stay-in-place forms to contain any spalling concrete. The attributes for the CF described in the NCHRP 782 report also included load capacity, level of redundancy, and composite construction.

Table 6.1. General description of the CF categories.

Level	Category	Consequence on Safety	Consequence on Serviceability	Summary Description
1	Low	None	Minor	Minor effect on serviceability, no effect on safety
2	Moderate	Minor	Moderate	Moderate effect on serviceability, minor effect on safety
3	High	Moderate	Major	Major effect on serviceability, moderate effect on safety
4	Severe	Major	Major	Structural collapse/loss of life

Consequence scenarios were presented to several of the RAPs to determine if there were other significant factors that should be considered in determining the appropriate CF for different damage modes and different components. The RAPs identified the factors listed in Table 6.2. Additional potential attributes identified included traffic speed, number of spans and span length, and the construction material.

Assuming the CF would typically be a function of bridge redundancy, load capacity, ADT and the feature under the bridge, the CF framework was further developed through the research. Although this factor was originally proposed to be rated using a weighted sum model, this was not found to be a suitable approach because two of the key factors that would affect the CF, structural redundancy, and load capacity, are not sufficiently interrelated to be suitable for a weighted sum scoring process. Structural redundancy is independent of any of the other attributes that might be considered for the CF, making a weighted sum model ineffective for rating the CF.

Consequently, a categorical model was developed for the CF that would address the consequences of a component in a bridge deteriorating to a CS 3. The categorical model considers the load capacity of the structure, redundancy, serviceability, and the feature under the bridge. The CF to be assigned for a given damage mode is the highest relevant CF category. In this way, bridges can be categorized based on available data, expert judgement, and existing policies and procedures. All of the data for the CF is available from bridge inventory data.

Table 6.2. CF attributes identified during the initial RAP meetings.

Attributes (1)	Attributes (2)
ADT/ADTT	Traffic speed
Number of Spans	Feature Under
Length	Feature Carried
Number of Lanes	Stay-in-Place Forms
Construction Material	Redundancy
Composite Construction	Load Carrying Capacity / Rating

Figure 6.1 shows the attributes for the redundancy and load capacity attributes. The redundancy factor assigns non-redundant structures, such as NSTMs, as having a *severe* consequence, while redundant structures are placed in the category of *moderate* or *low*. Structures that require analysis to establish redundancy, or that are subject to state policies are assigned a *high* CF. For example, bridges with only three members, bridges with large beam spacings, or bridges with redundancy established through nationally recognized methods could be assigned a CF of *high* for damage modes that affect the superstructure.

The load capacity attribute categorizes bridges in terms of the inventory load rating for the bridge, with the *severe* category assigned to bridges with an inventory LRF of less than 1.0. Bridges with a LRF of greater than 1.2 are assumed to have reserve capacity and are categorized as having a CF of *low*. Bridges with an LRF of 1.0 to 1.2 are categorized as having *moderate* CF for most cases. Other factors such as frequent exposure to overloads or permit load limits could elevate the CF for a certain bridge to *high* based on engineering judgement. For example, bridges with a routine permit load rating (SNBI item B.LR.08) of "B," meaning some permit loads are restricted, could be assigned a *high* CF. This is based on the rationale that there is an increased chance of overload, resulting in a more significant consequence as compared with bridges with reserve strength. If the bridge has a LRF of less than 1.0, the CF is rated as *severe* based on the rational that there would be little reserve strength is the case of a member losing load carrying capacity.

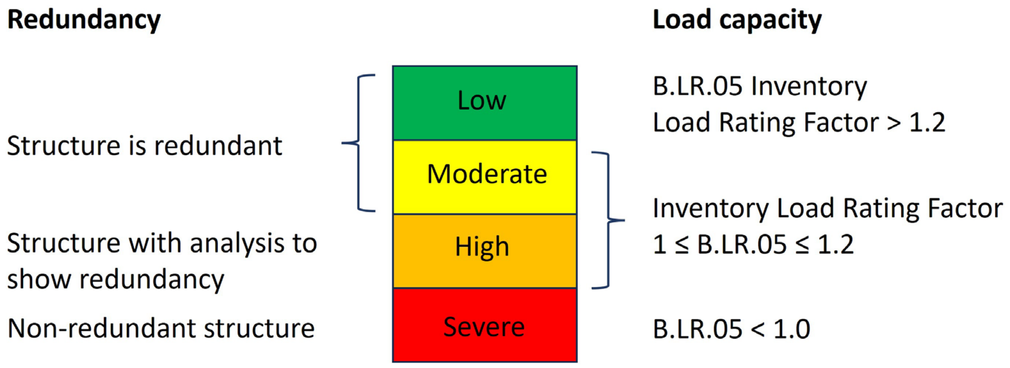
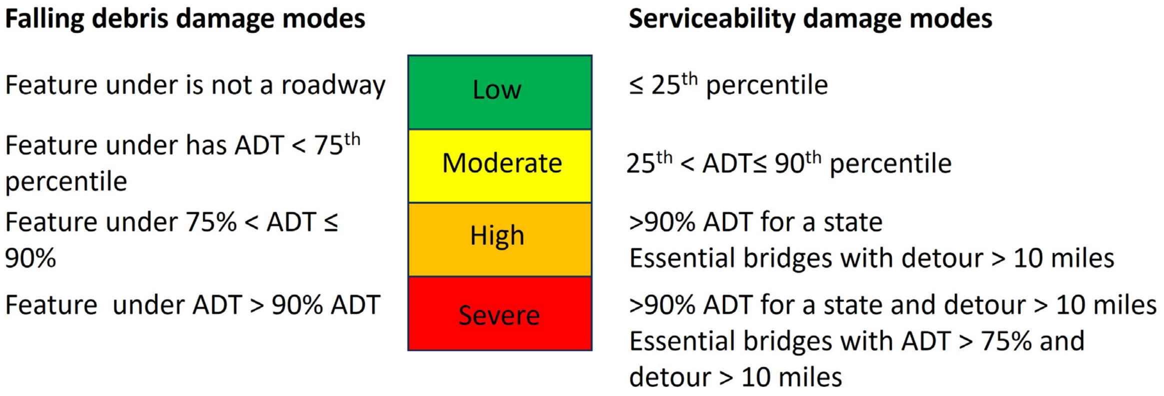


Figure 6.1. CF attributes for redundancy and load capacity.

Factors that impact the serviceability of a bridge are illustrated in Figure 6.2. The serviceability factors consider the traffic volume on the deck to determine the potential consequence resulting from a lane closure due to either loss of a primary member in a redundant bridge, or serious damage in a bridge deck that affects traffic safety and serviceability of the deck. The levels of ADT shown in the figure were subjectively selected. The ADT levels vary significantly between different states, and as such, this attribute was described in terms of percentiles. The goal was to identify bridges where lane closures present a significant consequence in transportation efficiency. The factor also considers the detour length for bridges that are essential to the transportation network.

Finally, the CF includes attributes that consider the possibility that a superstructure or deck component that has deteriorated to a CR 3 may result in concrete falling into traffic below the bridges. This factor considers that two events must occur for debris falling from a bridge to impact a car directly. First, debris must fall from the bridge, and second, a vehicle must be present to be impacted by it. Consequently, this factor considers the feature under the bridge, and the ADT level on the feature under the bridge. Again, the ADT level is expressed in percentile terms and was subjectively chosen. Users may wish to assign different percentile values for ADT based on engineering judgement and experience with their inventories. Another factor that may be suitable for consideration under this attribute is composite design and existence of cross-frames. For a structure with no cross frames and non-composite design, there is a possibility for a failed superstructure member to fall into traffic below the structure. For example, a non-composite prestressed box girder bridge member may collapse into the roadway below.



Falling debris damage modes		**Serviceability damage modes**
Feature under is not a roadway	Low	≤ 25th percentile
Feature under has ADT < 75th percentile	Moderate	25th < ADT≤ 90th percentile
Feature under 75% < ADT ≤ 90%	High	>90% ADT for a state Essential bridges with detour > 10 miles
Feature under ADT > 90% ADT	Severe	>90% ADT for a state and detour > 10 miles Essential bridges with ADT > 75% and detour > 10 miles

Figure 6.2 CF attributes and categories for falling debris and serviceability.

The highest relevant CF for a given damage mode should be used in the risk analysis. It should be noted that most of the criteria for the CF attributes are based on engineering judgement, and users may wish to modify the factors appropriately for their bridge inventories.

Chapter 7 Conclusions

The report includes examples and analysis of the risk models, criteria and commentary for attributes, and methodologies for analyzing and calibrating the risk models. Case studies in the form of RAP models developed by different states, revised through the research, and analyzed against bridge inventory data, are included in the report and the appendices. Commentary supporting the new attributes were combined with existing commentary. These elements of the report form guidance for implementing RBI, the research's main objective. A summarized handbook describing the process for RBI implement is included in Appendix A.

The conclusion from the study were as follows:

- The methodology developed in the research for analyzing risk models using MC simulation provides a tool for calibrating the weights of attributes in risk models, adjusting criteria for attributes, and forecasting the outcome of the risk models when applied to the subject inventory of bridges. It was shown that this methodology can be used to demonstrate the risk model's ability to identify bridge components with elevated risk. The methodology based on MC simulations can provide an effective tool for analyzing risk models and communicating their effectiveness.
- It was found that there was consistency in many of the damage modes and attributes identified by RAPs. In this research, six different RAPs assessed risk factors consistently, with differences related to their inventory of bridges and environment. This conclusion supported the concept of the RAPs as an effective tool for risk analysis.
- It was found that the risk models developed by the RAPs were consistent with target ranges based on NBIS requirements. Given that the RAPs were from six different states and were composed of individuals with diverse backgrounds and experiences, the consistency of the RAP outcomes with the target ranges was a significant finding.
- The original back-casting procedure envisioned for the research had limited effectiveness due to inconsistencies in inspection data format and content and changing inspection requirements. This was complicated by the diversity of inspection practices between different states. A single state analyzing their own inspection data would have more success. Additionally, it was found that insights into the quality of the risk models could not be obtained from simply reviewing inspection results, although the process was useful for determining if any significant risk factors had not been included in the risk models.

7.1. Recommendations

- RAPs should prepare for their meetings by obtaining certain statistics describing their bridge inventories that would be expected to arise during the consideration of criteria for damage modes. Examples of these statistics include Average Daily Traffic (ADT) and Average Daily Truck Traffic (ADTT) levels, a means of estimating areas of the state where deicing chemical use is relatively high, and some statistics on typical element-level condition state quantities for bridge components with different condition ratings. This would reduce efforts following the RAP meeting to formulate criteria threshold values and improve the process's efficiency.
- The original risk matrix proposed in the NCHRP 782 report should be modified as follows:
  - The matrix location [R1,1] should be 72 months, since 96-month intervals are not allowable under the NBIS.

    - The position R1,3 was originally indicated as a 48-month interval but is proposed as a 72-month interval to provide appropriate granularity to sort bridges into different “bins” in terms of risk.
- Special inspection procedures to collect the necessary data to support attribute criteria will be needed if element-level inspection data is unavailable for a bridge. These inspection procedures should document the relevant element-level data to address the criteria developed by the RAP. The procedures could be limited to only those bridges assigned extended inspection intervals using the Method 2 approach.

References

AASHTO. (2019). *Manual for Bridge Element Inspection, 2nd Edition*. Washington, D.C.: American Association of State Highway Transportation Officials.

Allison, P. D. (2010). *Survival analysis using SAS: A practical guide*: SAS Institute.

API. (2016). *API Recomended Practice 580, Risk-Based Inspection, 3rd edition*. Washinton, D.C.: American Petroleum Institute.

FHWA. (1988). *Technical Advisory: Revisions to the National Bridge Inspection Standards (NBIS)*. Washington, D.C.

FHWA. (1995). *Recording and Coding Guide for the Structure Inventory and Appraisal of the Nation's Bridges*. (FHWA-PD-96-001). Washington, D.C. : Federal Highway Administration

FHWA. (2022a). National Bridge Inspection Standards. In *23 CFR 650* (Vol. 87 FR 27396, pp. 42). Washington, D.C.: Federal Registrar.

FHWA. (2022b). *National Bridge Inspection Standards Inspection Interval Guidance*. Washington, D.C.: FHWA

FHWA. (2022c). *Specifications for the National Bridge Inventory*. (FHWA-HIF-22-017). Washington, D.C.: Federal Highway Administration

Gunaydin, H. M. (2006). The Delphi method. *Optimization Group.* Retrieved from https://www.researchgate.net/profile/Huesnue-Guenaydin/publication/308653253_The_Delphi_Method/links/57ea495008aef8bfcc9901ff/The-Delphi-Method.pdf

Hohmann, E., Brand, J. C., Rossi, M. J., & Lubowitz, J. H. (2018). Expert opinion is necessary: Delphi panel methodology facilitates a scientific approach to consensus. In (Vol. 34, pp. 349-351): Elsevier.

Hosmer Jr, D. W., Lemeshow, S., & May, S. (2008). *Applied survival analysis: regression modeling of time-to-event data* (Vol. 618): Wiley-Interscience.

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association, 53*(282), 457-481.

Kiral, I., Kural, Z., & Çomu, S. (2014). *Risk identification in construction projects: Using the Delphi method.* Paper presented at the 11th International Congress on Advances in Civil Engineering, Istanbul, Turkey.

Nasrollahi, M., & Washer, G. (2015). Estimating Inspection Intervals for Bridges Based on Statistical Analysis of National Bridge Inventory Data. *Journal of Bridge Engineering, 20*(9), 04014104. doi:doi:10.1061/(ASCE)BE.1943-5592.0000710

Parr, M. J., Connor, R. J., & Bowman, M. (2010). Proposed method for determining the interval for hands-on inspection of steel bridges with fracture critical members. *Journal of Bridge Engineering, 15*(4), 352-363.

PTI/ASBI. (2019). Specification for Multistrand and Grouted Post-Tensioning, M50.3-19. In (Vol. M50.3-19, pp. 64). Farminton Hills, MI: Post-Tensioning Institute, American Segmental Bridge Institute.

Washer, G., Connor, R., Ciolko, A., Kogler, R., Fish, P., Forsyth, D., & Nasrollahi, M. (2014). *NCHRP 782: Proposed Guideline for Reliability-Based Bridge Inspection Practices*. Retrieved from Washington, D.C.

APPENDIX A. HANDBOOK FOR IMPLEMENTATION OF RISK-BASED INSPECTION (RBI)

Appendices are available as a separate file.

APPENDIX B. NSTM HANDS-ON INSPECTION INTERVAL ASSESSMENT

Appendices are available as a separate file.

APPENDIX C. RISK MODELS FROM THE RESEARCH

Appendices are available as a separate file.

APPENDIX D. Attribute Index and Commentary

Appendices are available as a separate file.