

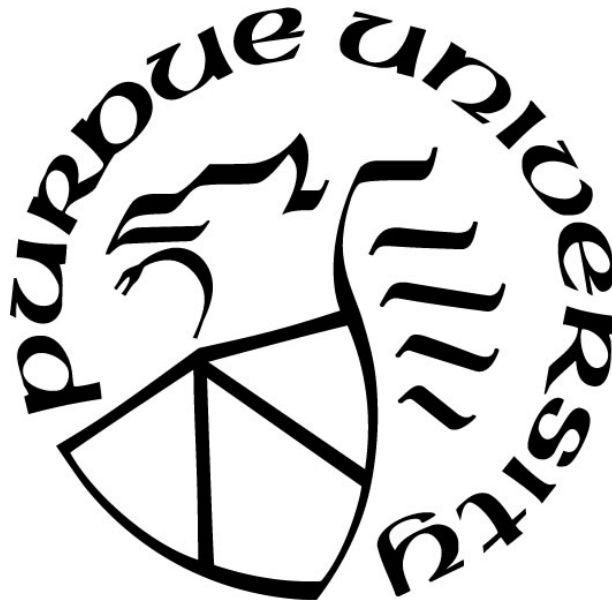
**A BENCHMARK FOR EVALUATING PERFORMANCE IN VISUAL
INSPECTION OF STEEL BRIDGE MEMBERS AND STRATEGIES FOR
IMPROVEMENT**

by
Leslie E. Campbell

A Dissertation

*Submitted to the Faculty of Purdue University
In Partial Fulfillment of the Requirements for the degree of*

Doctor of Philosophy



Lyles School of Civil Engineering
West Lafayette, Indiana
May 2019

**THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL**

Dr. Robert J. Connor, Chair

Lyles School of Civil Engineering, Purdue University

Dr. Mark D. Bowman

Lyles School of Civil Engineering, Purdue University

Dr. Mohammad R. Jahanshahi

Lyles School of Civil Engineering, Purdue University

Dr. Glenn A. Washer

Department of Civil and Environmental Engineering, University of Missouri

Approved by:

Dr. Dulcy Abraham

Head of the Graduate Program

*To my sister, Maureen, for being the mo' to my les' and to my parents, Ward and Cathy, for
thinking that joke was funny all these years*

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Robert Connor, for his support and guidance during this research project. His enthusiasm and encouragement provided the motivation I needed to return to school and pursue this degree. Similarly, I am thankful to my supervisors at the US Army Corps of Engineers, David Lovett, Dr. Christopher Dunn, Jean Vossen, and Mark Gayheart, whose trust and respect made it possible.

I would also like to thank the other members of my committee, Dr. Mark Bowman, Dr. Mohammed Jahanshahi, and Dr. Glenn Washer, for their expert review and contribution to my research.

I would like to give special thanks to Luke Snyder and Julie Whitehead for all their hard work in getting this project started and their ongoing support and willingness to answer any and all questions. I would also like to thank my friends and colleagues at Bowen Laboratory, including Dr. Cem Korkmaz, Dr. Curtis Schroeder, Thomas Bradt, Dr. Jason Lloyd, Dr. Francisco Bonachera Martin, and Adam Cullison.

This research was funded in part by the Indiana Department of Transportation, the Minnesota Department of Transportation, and various state and federal agencies under the Transportation Pooled Fund Program. I also wish to acknowledge the Department of Defense Science, Mathematics, and Research for Transformation scholarship which covered my tuition and education related expenses. Finally, I am forever indebted to the many inspectors that generously volunteered their time and wisdom to complete this project. Without their help, the majority of this research would not have been possible.

And to Matt, I can never say thank you enough for going on this journey with me.

TABLE OF CONTENTS

LIST OF TABLES.....	10
LIST OF FIGURES	13
ABSTRACT.....	17
1. INTRODUCTION	19
1.1 Motivation.....	19
1.2 Research Objectives.....	20
1.3 Development of the National Bridge Inspection Standards.....	21
1.4 Requirements for Bridge Inspection Team Leaders	22
1.5 Evaluating Visual Inspection Capability	23
1.6 Human Factors in Visual Inspection.....	25
1.7 UAS-assisted Bridge Inspection	28
2. PRIOR RESEARCH.....	32
2.1 Whitehead (2015) and Snyder (2015).....	32
2.2 Moore, Phares, Graybeal, Rolander, & Washer (2001).....	35
2.3 Spencer (1996).....	38
2.4 Drury and Watson (2002)	40
2.5 Kulicki, Prucz, Sorgenfrei, & Mertz (1990)	42
3. HANDS-ON VISUAL INSPECTION	45
3.1 Introduction.....	45
3.2 Inspection Scenario.....	45
3.3 Inspector Demographics	49
3.4 Inspector Response Evaluation.....	50
3.5 Inspection Results	52
3.5.1 Crack Detection and False Calls.....	53
3.5.2 Crack Sizing.....	58
3.5.3 Other Defects.....	61
3.6 Crack Length Analysis.....	61
3.6.1 Crack Detection by Crack Length	62
3.6.2 Probability of Detection Curves	66

3.7	Human and Environmental Factors	74
3.7.1	Univariate Regression and Two Samples <i>t</i> -Test.....	76
3.7.2	Multivariate Regression.....	87
3.7.3	Binary Logit Model	91
3.7.4	Vigilance.....	97
3.7.5	Tool Use.....	99
3.7.6	Discussion.....	100
3.8	Recommendations for Visual Inspection.....	102
3.8.1	Equipment.....	102
3.8.2	Training.....	102
3.8.3	Inspection Procedures.....	103
3.8.4	Performance Testing.....	106
3.9	Summary.....	109
4.	INSPECTOR TRAINING	111
4.1	Introduction.....	111
4.2	Observational Skills Training	111
4.2.1	Visual Inspection Process	112
4.2.2	Perception and Recognition.....	114
4.2.3	Attention	115
4.2.4	Memory.....	116
4.2.5	Mental Images	117
4.2.6	Mental Models	118
4.2.7	Judgement and Decision Making.....	120
4.3	Inspection Tools and Techniques Training.....	121
4.4	Summary.....	122
5.	UAS-ASSISTED VISUAL INSPECTION	124
5.1	Introduction.....	124
5.2	Field Inspection.....	125
5.2.1	Inspection Scenario.....	125
5.2.2	Inspector Demographics	128
5.2.3	Inspection Results.....	129

5.2.3.1	Crack Detection and False Calls.....	129
5.2.3.2	Crack Sizing	131
5.2.3.3	Inspection Duration	134
5.2.3.4	NASA Task Load Index	135
5.2.4	Crack Length Analysis	135
5.2.5	Human and Environmental Factors	136
5.2.6	Comparison to Hands-on Inspection	139
5.2.7	Discussion.....	143
5.3	Desk Inspection.....	146
5.3.1	Inspection Scenario.....	146
5.3.2	Inspector Demographics	148
5.3.3	Inspection Results.....	149
5.3.3.1	Crack Detection and False Calls.....	149
5.3.3.2	Crack Sizing	154
5.3.3.3	Inspection Duration	157
5.3.3.4	NASA Task Load Index	158
5.3.4	Crack Length Analysis	159
5.3.5	Human and Environmental Factors	161
5.3.5.1	Univariate Regression and Two Samples <i>t</i> -Test.....	163
5.3.5.2	Binary Logit Model	166
5.3.5.3	Computer System	168
5.3.5.4	Inspector Assessment	171
5.3.6	Comparison to UAS-assisted Field Inspection.....	176
5.3.7	Comparison to Hands-on Inspection	181
5.3.8	Discussion.....	187
5.4	Recommendations.....	188
5.4.1	Equipment.....	188
5.4.2	Training.....	189
5.4.3	Inspection Procedures.....	190
5.5	Summary	191
6.	TRUSS CHORD INSPECTION AND LOAD RATING.....	193

6.1	Introduction.....	193
6.2	Inspection.....	193
6.2.1	Description of Specimen.....	194
6.2.2	Inspection Scenario.....	196
6.2.3	Inspector Demographics	197
6.2.4	Inspector Response Evaluation.....	198
6.2.5	Inspection Results.....	199
6.2.5.1	Task 1	199
6.2.5.2	Task 2	202
6.2.5.3	Task 3	210
6.2.6	Recommendations.....	212
6.3	Load Rating.....	213
6.3.1	Load Rating Scenario	213
6.3.2	Engineer Demographics.....	215
6.3.3	Load Rating Results.....	216
6.3.3.1	Material Strength Assumptions	217
6.3.3.2	Gross and Net Section Area Calculations.....	217
6.3.3.3	Task 1	221
6.3.3.4	Task 2	223
6.3.4	Recommendations.....	226
6.4	Summary	227
7.	SUMMARY AND CONCLUSIONS	230
7.1	Summary of Principal Findings	231
7.1.1	Summary of Hands-on Inspection Findings	231
7.1.2	Summary of UAS-assisted Inspection Findings	232
7.1.3	Summary of Truss Chord Inspection and Load Rating Findings	232
7.2	Recommendations for Future Research	233
	REFERENCES	236
	APPENDIX A. HANDS-ON INSPECTION DOCUMENTS.....	243
	APPENDIX B. UAS-ASSISTED DESK INSPECTION DOCUMENTS	251
	APPENDIX C. TRUSS CHORD INSPECTION DOCUMENTS	268

APPENDIX D. TRUSS CHORD LOAD RATING DOCUMENTS	274
---	-----

LIST OF TABLES

Table 1.1 Probability of detection link functions.....	24
Table 3.1 Inspector demographics for hands-on inspections.....	50
Table 3.2 Summary of results from hands-on inspections.....	53
Table 3.3 Hands-on inspection results by inspector	54
Table 3.4 Error analysis for crack length measurements from hands-on inspections	61
Table 3.5 Probability of detection crack sizes considering all cracks	67
Table 3.6 Probability of detection by crack length (all cracks)	70
Table 3.7 Probability of detection crack sizes considering only the out-of-plane cracks.....	71
Table 3.8 Descriptive statistics for dependent and independent variables	75
Table 3.9 p -values for the univariate linear regression analyses	77
Table 3.10 Results from multivariate regression analysis predicting detection rate	89
Table 3.11 Results from the multivariate regression analysis predicting false calls	89
Table 3.12 Results from the binary logit model estimating probability of detection	95
Table 3.13 Marginal effects for the parameters in the binary logit model	96
Table 3.14 Summary of results for inspector qualification based on unweighted detection rate	107
Table 3.15 Summary of results for inspector qualification based on weighted detection rate ...	107
Table 3.16 Summary of results for inspector qualification based on crack type detection rate .	108
Table 3.17 Summary of results for inspector qualification based on inspector rating	108
Table 3.18 Summary of results for inspector qualification based on probability of detection...	109
Table 4.1 Description of the visual inspection process unit	113
Table 4.2 Description of the perception and recognition unit	115
Table 4.3 Description of the attention unit	116
Table 4.4 Description of the memory unit.....	117
Table 4.5 Description of the mental images unit	118
Table 4.6 Description of the mental models unit.....	119
Table 4.7 Description of the judgement and decision making unit	121
Table 5.1 Inspector demographics for UAS-assisted field inspections	129

Table 5.2. UAS-assisted field inspection results by inspector.....	129
Table 5.3 Error analysis for crack length estimates from the UAS-assisted field inspections ...	134
Table 5.4 UAS-assisted field inspection durations by inspector	134
Table 5.5 NASA-TLX workload scores for the inspectors and pilot	135
Table 5.6 Wind speeds by inspection day.....	137
Table 5.7 Average performance during UAS-assisted field and hands-on inspections.....	140
Table 5.8 Difference in inspection performance during UAS-assisted field and hands-on inspections for a single inspector.....	142
Table 5.9 Inspector demographics for UAS-assisted desk inspections	149
Table 5.10 Summary of results from UAS-assisted desk inspections	150
Table 5.11 UAS-assisted desk inspection results by inspector.....	150
Table 5.12 Average UAS-assisted desk inspection performance by video set.....	154
Table 5.13 Error analysis for crack length estimates from UAS-assisted desk inspections	157
Table 5.14 UAS-assisted desk inspection durations by inspector	158
Table 5.15 NASA-TLX workload scores for the UAS-assisted desk inspectors	159
Table 5.16 Descriptive statistics for dependent and independent variables	162
Table 5.17 <i>p</i> -values for the univariate linear regression analyses	163
Table 5.18 Results from the binary logit model estimating probability of detection	167
Table 5.19 Marginal effects for the parameters in the binary logit model	167
Table 5.20 Usefulness ratings for video playback features	170
Table 5.21 Inspector recommendations for improving UAS-assisted desk inspections.....	176
Table 5.22 Average performance during UAS-assisted desk and field inspections.....	178
Table 5.23 Difference in inspection performance during UAS-assisted desk and field inspections for two inspectors.....	181
Table 5.24 Average performance during UAS-assisted desk and hands-on inspections.....	183
Table 5.25 Difference in inspection performance during UAS-assisted desk and hands-on inspections for two inspectors.....	186
Table 6.1 Inspector demographics and inspection conditions	198
Table 6.2. Summary of Task 1 results	200
Table 6.3 Summary of reference areas	203
Table 6.4 Summary of Task 2 results	204

Table 6.5 Measurement statistics for the cover plates at Cross Section 1	206
Table 6.6 Measurements statistics for the cover plates at Cross Section 2.....	206
Table 6.7 Percent error for cover plate area estimates.....	207
Table 6.8 Descriptive statistics for Member A at Cross Section 1	208
Table 6.9 Descriptive statistics for Member B at Cross Section 1	208
Table 6.10 Descriptive statistics for Member A at Cross Section 2	208
Table 6.11 Descriptive statistics for Member A at Cross Section 2	209
Table 6.12 Descriptive statistics for the truss chord at Cross Sections 1 and 2.....	209
Table 6.13 p -values for member areas at Cross Sections 1 and 2.....	210
Table 6.14 Summary of Task 3 Results	211
Table 6.15 Load rater demographics	216
Table 6.16 Material strength assumptions	217
Table 6.17 As-built gross and net section area estimates	218
Table 6.18. Summary of as-inspected gross and net section area estimates.....	219
Table 6.19 Summary of Task 1 results (as-built condition).....	221
Table 6.20 Summary of Task 1 results (as-inspected condition).....	222
Table 6.21 Assumed condition and system factors.....	223
Table 6.22 Summary of Task 2 results	224

LIST OF FIGURES

Figure 3.1 Probability of detection training structure with POD specimens	46
Figure 3.2 Hands-on inspection of the (a) girder, (b) overhead mounted riveted plate, and (c) welded cover plate specimens.....	47
Figure 3.3 Photograph of a typical (a) out-of-plane crack, (b) rivet hole crack, and (c) weld toe crack.....	47
Figure 3.4. Standard inspection form for specimen 1GPP1-A with notes from two inspectors...	51
Figure 3.5 Hits and false calls by inspector during the hands-on inspections	55
Figure 3.6 Inspector detection rates by crack type	56
Figure 3.7 Weld toe crack detection rate plotted against out-of-plane crack detection rate.....	57
Figure 3.8. Detection rate plotted against the number of days since the first POD inspection	58
Figure 3.9 Measured crack length versus actual crack length for out-of-plane cracks.....	59
Figure 3.10 Measured crack length versus actual crack length for weld toe cracks.....	60
Figure 3.11 Measured crack length versus actual crack length for rivet hole cracks	60
Figure 3.12 Number of crack detections by crack	63
Figure 3.13 Detection rate by crack length.....	64
Figure 3.14 Out-of-plane crack detection rate by crack length	65
Figure 3.15 Weld toe crack detection rate by crack length.....	65
Figure 3.16 Rivet hole crack detection rate by crack length.....	66
Figure 3.17 Probability of detection curves considering all cracks	68
Figure 3.18 Probability of detection curves considering all cracks, excluding inspectors 13CA-09, 06JD-15, and 18GA-03	69
Figure 3.19 Probability of detection curves considering only the out-of-plane cracks	72
Figure 3.20 Probability of detection curves considering only the weld toe cracks	73
Figure 3.21 Probability of detection curves considering only the rivet hole cracks.....	74
Figure 3.22 Detection rate plotted against average temperature.....	78
Figure 3.23 Detection rate plotted against inspection duration	79
Figure 3.24 Inspection duration plotted against average temperature	80
Figure 3.25 Detection rate plotted against the number of training courses taken by the inspector	81

Figure 3.26 Detection rate plotted against inspection experience	82
Figure 3.27 Weld toe crack detection rate plotted against inspection experience	83
Figure 3.28 Inspection experience plotted against inspection duration.....	84
Figure 3.29 Detection rate plotted against inspector age.....	85
Figure 3.30 Weld toe crack detection rate plotted against inspector age.....	85
Figure 3.31. The number of false calls plotted against (a) average temperature, (b) inspection duration, (c) number of training courses, and (d) inspection experience.....	86
Figure 3.32 (a) Detection rate and inspection duration by girder row and (b) false alarms and inspection duration by girder row	99
Figure 4.1 Prototypical screens from the (a) search and (b) decision exercises	114
Figure 4.2 Select slides from the Inspection Tools and Techniques training module	122
Figure 5.1 UAS-assisted field inspections of the POD specimens	126
Figure 5.2 Vision test charts used during the UAS-assisted field inspections.....	126
Figure 5.3 Hits and false calls by inspector during the UAS-assisted field inspections.....	130
Figure 5.4 Inspector detection rates by crack type	131
Figure 5.5 Estimated crack length versus actual crack length for out-of-plane cracks	132
Figure 5.6 Estimated crack length versus actual crack length for weld toe cracks	133
Figure 5.7 Estimated crack length versus actual crack length for rivet hole cracks.....	133
Figure 5.8 Detection rate by crack length.....	136
Figure 5.9 Detection rate plotted against wind speed	137
Figure 5.10 False positives plotted against wind speed.....	138
Figure 5.11 Time ratio (field time/inspection time) plotted against wind speed.....	138
Figure 5.12 Comparison of average detection rate during the UAS-assisted field and hands-on inspections.....	141
Figure 5.13 Comparison of average number of false calls made during the UAS-assisted field and hands-on inspections	141
Figure 5.14 Hits and false calls by inspector during the UAS-assisted desk inspections.....	151
Figure 5.15 Inspector detection rates by crack type	153
Figure 5.16 Estimated crack length versus actual crack length for out-of-plane cracks	155
Figure 5.17 Estimated crack length versus actual crack length for weld toe cracks	156
Figure 5.18 Estimated crack length versus actual crack length for rivet hole cracks	157

Figure 5.19 Number of crack detections by crack	160
Figure 5.20 Detection rate by crack length	161
Figure 5.21 Detection rate plotted against the number of false calls.....	164
Figure 5.22 Detection rate plotted against the number of routine steel bridge inspections performed in the previous 12 months	165
Figure 5.23 Still image from a UAS video (a) uncorrected, (b) zoomed in on the crack-prone region at the bottom of the transverse stiffener, (c) with brightness adjusted, and (d) zoomed in on the crack-prone region with brightness adjusted	170
Figure 5.24 Effort and focus level during UAS-assisted desk inspection compared to traditional bridge inspection.....	172
Figure 5.25 Quality of a UAS-assisted desk inspection compared to a UAS-assisted field inspection and a hands-on inspection	175
Figure 5.26 Comparison of average detection rate during the UAS-assisted desk and field inspections.....	179
Figure 5.27 Comparison of average number of false calls made during the UAS-assisted desk and field inspections	179
Figure 5.28 Comparison of average detection rate during the UAS-assisted desk and hands-on inspections.....	184
Figure 5.29 Comparison of average number of false calls made during the UAS-assisted desk and hands-on inspections	185
Figure 6.1 Truss chord elevation view (looking from joint L4 to L2).....	194
Figure 6.2 Elevation and section views of the truss chord	195
Figure 6.3 View of typical corrosion damage resulting in section loss and pack rust.....	195
Figure 6.4 Inspectors taking thickness measurements of the truss chord	197
Figure 6.5 Illustration of the method used to calculate remaining member area from field measurements.....	199
Figure 6.6 Critical section locations identified by the inspectors during Task 1.....	200
Figure 6.7 Truss chord near the critical section identified by three inspectors	201
Figure 6.8 Truss chord at (a) Cross Section 1 and (b) Cross Section 2	202
Figure 6.9 Disassembled pieces from Cross Sections 1 and 2.....	203
Figure 6.10 Estimated area plotted against reference area for Cross Section 1 and 2.....	205
Figure 6.11 Cover plates at Cross Section 2	207
Figure 6.12 Limits of inspection for Task 3	210

Figure 6.13 Critical section locations identified by the inspectors during Task 3.....	211
Figure 6.14 Typical photos and metalwork loss measurements included in the mock inspection report.....	214
Figure 6.15 2D SAP model of Span 16.....	215

ABSTRACT

Author: Campbell, Leslie, E. PhD

Institution: Purdue University

Degree Received: May 2019

Title: A Benchmark for Evaluating Performance in Visual Inspection of Steel Bridge Members and Strategies for Improvement

Major Professor: Dr. Robert J. Connor

Visual inspection is the primary means of ensuring the safety and functionality of in-service bridges in the United States and owners spend considerable resources on such inspections. While the Federal Highway Administration (FHWA) and many state departments of transportation have guidelines related to inspector qualification, training, and certification, an inspector's actual capability to identify defects in the field under these guidelines is unknown. This research aimed to address the knowledge gap surrounding visual inspection performance for steel bridges in order to support future advances in inspection and design procedures. Focusing primarily on fatigue crack detection, this research also considered the ability of inspectors to accurately and consistently estimate section loss in steel bridge members.

Inspection performance was evaluated through a series of simulated bridge inspections performed in representative in-situ conditions. First, this research describes the results from 30 hands-on, visual inspections performed on full size bridge specimens with known fatigue cracks. Probability of Detection (POD) curves were fit to the inspection results and the 50% and 90% detection rate crack lengths were determined. The variability in performance was large, and only a small amount of the variance could be explained by individual characteristics or environmental conditions. Based on the results, recommendations for improved training methods, inspection procedures, and equipment were developed. Above all, establishment of a performance based qualification system for bridge inspectors is recommended to confirm that a satisfactory level of performance is consistently achieved in the field.

Long term, managing agencies may eschew traditional hands-on bridge inspection methods in favor of emerging technologies imagined to provide improved results and fewer logistical challenges. This research investigated the potential for unmanned aircraft system (UAS)

assistance during visual inspection of steel bridges. Using the same specimens as in the hands-on inspections, four UAS-assisted field inspections and 19 UAS-assisted desk inspections were performed. A direct comparison was made between performance in the hands-on and UAS-assisted inspections, as well as between performance in the two types of UAS-assisted inspections. Again, significant variability was present in the results suggesting that human factors continue to have a substantial influence on inspection performance, regardless of inspection method.

Finally, to expand the findings from the crack detection inspections, the lower chord from a deck truss was used to investigate variability in the inspection of severely corroded steel tension members. Five inspectors performed a hands-on inspection of the specimen and four engineers calculated the load rating for the same specimen. Significant variability was observed in how inspectors recorded thickness measurements during the inspections and engineers interpreted the inspection reports and applied the code requirements.

1. INTRODUCTION

1.1 Motivation

The National Bridge Inspection Standards (NBIS) were originally conceived in the 1968 Federal-Aid Highway Act. The NBIS currently mandates that all structures longer than 20 feet located on public roads be inspected every two years by a qualified bridge inspector. The results of these inspections are reported for inclusion in the National Bridge Inventory maintained by the Federal Highway Administration [1]. These routine inspections are typically conducted visually from the surface of the bridge deck and from the ground beneath the structure. Following the collapse of the Mianus River Bridge in 1983, additional inspection requirements were invoked for bridges considered vulnerable to collapse due to their lack of structural redundancy. These bridges include steel tension members whose failure would be expected to result in the partial or complete collapse of the bridge and are termed “fracture critical members” [2]. For fracture critical members, the inspection must be completed from within an arm’s length of the member. Therefore, some form of aerial access, either through climbing or the use of specialized equipment (e.g., a snooper truck or man lift), is required to gain access to the surface to be inspected. In addition to increasing the person hours, equipment, and traffic control required to complete the inspection, this increases the risk for both the inspectors and the public. In 2017, approximately three percent of the 615,000 bridges subject to the federal inspection requirements were classified as fracture critical [3].

Hands-on, visual inspections are considered the primary means of ensuring the safety of the 155 million drivers who travel across these fracture critical bridges each day [3]. Although there is little question that these inspections increase public safety, only a small number of studies have attempted to quantify the reliability or accuracy of these inspections [4], [5]. And while 23% of the respondents to a survey conducted in 2005 indicated that a hands-on fracture-critical inspection had identified defects, such as cracks or corrosion, that may have led to catastrophic failure without detection, these findings come with a significant price tag [6]. The same study estimated the cost of a fracture critical inspection to be approximately 2 to 5 times more than that associated with a routine inspection, and for some structures, increases of between 10 and 50 times have been

reported [6]. Thus, a large percentage of an agency's inspection budget may be consumed by a relatively small number of bridges.

In recent years, an increased emphasis has been placed on addressing many of the challenges associated with visual bridge inspection and identifying opportunities for improving safety and reducing cost. However, the development of new inspection strategies requires, among other things, a comprehensive understanding of what defects can and will be reliably detected under the current policy. For example, the introduction of reliability-based inspection intervals in lieu of the current, federally-mandated, uniform inspection interval, necessitates an understanding of the detectable crack size during in-service inspections [7], [8].

This research examines the performance of bridge inspectors during visual inspection of steel bridges, primarily focusing on the ability of inspectors to detect fatigue cracks and estimate section loss in steel bridge members. First, the research assessed inspection performance during traditional, hands-on inspections. Then, the potential for unmanned aircraft system (UAS) assistance during visual inspections was investigated.

This research project expands the understanding of visual inspection accuracy and consistency for steel bridges. Specifically, the probability of detection study is believed to be the first of its kind specific to steel bridges and will provide a quantitative baseline measure of hands-on inspection performance. Additionally, in recent years, UAS technology for civil infrastructure inspections has matured to the point that it is a viable supplement to traditional hands-on inspections of steel bridges. Ultimately, the results of this research may be used to develop improved inspection strategies which can provide increased safety for inspection personnel and the motoring public along with cost-savings for managing agencies.

1.2 Research Objectives

The research objectives for this project are as follows:

- Investigate the ability of bridge inspectors to accurately detect fatigue cracks in steel bridge members during hands-on, visual inspection. Qualitatively and quantitatively measure the

effect of inspector characteristics and inspection environment on visual inspection performance. Identify key factors that influence performance.

- Establish the probability of detection for fatigue cracks in steel bridges using visual inspection.
- Investigate the effectiveness of UAS assistance in visual inspection of steel bridge members. Compare performance during a UAS-assisted bridge inspection to performance during a hands-on inspection.
- Develop recommendations for inspection equipment, training, and procedures to improve performance during visual inspection of steel bridge members.
- Investigate the ability of bridge inspectors to accurately estimate the remaining thickness of steel bridge members during hands-on, visual inspection.
- Investigate the accuracy and consistency of load rating evaluations of corroded steel tension members based on visual inspection findings.

1.3 Development of the National Bridge Inspection Standards

Over the last half century, significant changes to bridge inspection requirements have been implemented at federal, state, and local levels. Prior to the 1967 collapse of the Silver Bridge in Point Pleasant, West Virginia, no federal standards or requirements governed the maintenance of in-service bridge structures [2]. This failure of a non-redundant eyebar in the suspension chain resulted in 46 fatalities and placed increased emphasis on bridge inspection [9]. In 1968, Congress passed the Federal-Aid Highway Act which called for the Secretary of Transportation to develop a national bridge inspection standard alongside a program for training bridge inspectors [10, 73 Stat. 145], and in 1971, the National Bridge Inspection Standards (NBIS) were released. In addition to establishing a national policy for inspection procedure, frequency, and reporting, the NBIS outlined the necessary qualifications for inspection personnel and maintenance requirements.

The 1983 collapse of the Mianus River Bridge in Greenwich, Connecticut revealed some lingering weaknesses in the NBIS, specifically the lack of guidance regarding the inspection of critical elements. The Mianus River Bridge collapse was caused by the failure of the pin-and-hanger assembly. Despite regular visual inspections, the assembly was allowed to deteriorate and eventually a fatigue crack grew to critical size and fractured the hanger [11]. As a result of this

failure, the term “fracture critical member (FCM)” was introduced to describe non-load-path-redundant, steel tension members. A supplementary inspection manual titled *Inspection of Fracture Critical Members* was released in 1986, and in 1988, the NBIS was revised to include special inspection criteria for bridges with FCMs. As of 2005, the NBIS require routine inspections to be performed every 24 months, although this interval may be extended to 48 months on a case-by-case basis as approved by the FHWA. Fracture critical members require an arm’s length inspection at intervals not to exceed 24 months [1], [2].

1.4 Requirements for Bridge Inspection Team Leaders

Currently, the NBIS outline five possible ways to qualify as a bridge inspection team leader based on a combination of training, education, and experience level [1]. Education and experience requirements vary inversely, but in all cases, satisfactory completion of a “comprehensive inspection training course” [1, §650.309] is mandated. Additionally, the NBIS requires “periodic bridge inspection refresher training” [1, §650.313] for bridge inspection team leaders. The FHWA has developed a full complement of training courses to satisfy the demands of the NBIS and offers these through the National Highway Institute (NHI). Successful completion of the 80-hour, instructor-led course titled *Safety Inspection of In-Service Bridges* and its prerequisite satisfies the requirement for comprehensive inspection training. The prerequisite consists of a one-week instructor led course titled *Engineering Concepts for Bridge Inspectors*, a 14-hour web based training course titled *Introduction to Safety Inspection of In-Service Bridges*, or a web-based assessment. Similarly, a 3-day, instructor-led course titled *Bridge Inspection Refresher Training* fulfills the continuing education requirement. The frequency of refresher training is determined by each managing agency. For inspection of bridges with fracture critical members, the FHWA offers a 3.5-day, instructor-led course titled *Fracture Critical Inspection Techniques for Steel Bridges*. While conducted mainly in the classroom, these training courses typically include either a field component or a virtual bridge inspection exercise completed on a computer.

Additionally, state or local agencies have the flexibility to supplement the minimum qualifications stipulated in the NBIS with additional requirements. For instance, both the Oregon Department of Transportation [12] and Minnesota Department of Transportation [13] require inspectors to pass a field proficiency test in addition to satisfying the requirements of the NBIS. Likewise, the

Indiana Department of Transportation [14] requires successful completion of the 2-day instructor-led course developed at Purdue University and titled *Inspecting Steel Bridges for Fatigue*.

1.5 Evaluating Visual Inspection Capability

The probability of detection (POD) metric was developed in the 1960s to quantitatively describe the detection capability of various nondestructive examination (NDE) techniques [15]. The introduction of “damage tolerance” into structure and system designs demanded a probabilistic description of NDE capability and previous evaluation methods became inadequate. A damage tolerant structure is designed with sufficient redundancy so that it can tolerate significant cracking or single member failure to such an extent that this damage would be easily detectable during regularly scheduled maintenance and inspection activities [16]. This requires an understanding of both the critical flaw size and the detectable flaw size, specifically, the largest flaw that may be missed during an inspection. The majority of the existing research on inspection reliability and POD has been performed in the aeronautical and nuclear industries, although the methods can be applied to similar industries that rely heavily on NDE for quality control.

In response to the growing popularity of POD studies and a general concern that the data generated from these studies could be easily misinterpreted, the United States Department of Defense (DoD) released *MILK-HDBK-1823A* [17] with guidance for measuring NDE system reliability with POD statistics. This handbook is widely cited throughout the world and is applicable to any NDE method that produces either a quantitative signal, such as ultrasonic testing, or a binary response, such as visual inspection. It presents a uniform approach for assessing the capability of an NDE method by relating the probability of detection (POD) to target size (a). The ideal inspection would be represented by a step function with $\text{POD}=0$ for all defects smaller than the detectable defect size and $\text{POD}=1$ for all defects larger than the detectable defect size [17].

Prior attempts to quantify detection probability considered only the number of defects detected divided by the total number of defects present, thereby disregarding the influence of defect size. Attempts to rectify this by calculating detection rate per size range reduced the accuracy of the calculations because of the corresponding reduction in sample size [17]. Instead, the DoD

handbook proposes eight possible underlying mathematical relationships between POD and size that can be fit to the test data and used to predict the results of future inspections.

In order to represent binary response data (hit/miss) as a linear function of the defect size, a link function is introduced to transform the binary data (0 or 1) to probability data (between 0 and 1). In this way, size can be related to the probability of a hit or the probability of a miss. The DoD handbook includes four possible link functions to transform the hit/miss data. These are the log-odds function, the inverse normal function, the complementary log-log function, and the log-log function. Table 1.1 shows the definition of each link function and the equation for probability derived from each function. The log-odds, or logistic, function is the most widely used link function, although all four functions should be considered in practice [17], [18].

Although it has been customary to assume that probability of detection is a function of the logarithm of defect size, this is not necessary, or even the best approach [17], [18]. Both a logarithmic and a Cartesian size response should be investigated for each link function. The model which best fits the data can be selected based on deviance, which is defined as two times the maximized log-likelihood ratio. A smaller deviance indicates better agreement between the model and the data.

Table 1.1 Probability of detection link functions

Link Name	Definition	Probability
Log-odds function (logistic)	$f(X) = g(y) = \ln\left(\frac{p}{1-p}\right)$	$p = \frac{\exp(f(X))}{1 + \exp(f(X))}$
Inverse normal function (probit)	$f(X) = g(y) = \Phi^{-1}(p)$	$p = 1 - \Phi(f(X))$
Complementary log-log function (Weibull)	$f(X) = g(y) = \ln(-\ln(1-p))$	$p = 1 - \exp(-\exp(f(X)))$
Log-log Function	$f(X) = g(y) = -\ln(-\ln(p))$	$p = \exp(-\exp(-f(X)))$

Continuing with the logistic function, let $p = POD(a)$ and $f(X) = \beta_0 + \beta_1(a)$ where a is crack length so the probability equation from Table 1.1 can be rewritten as

$$POD(a) = \frac{\exp(\beta_0 + \beta_1(a))}{1 + \exp(\beta_0 + \beta_1(a))} \quad (1.1)$$

The parameters, β_0 and β_1 , have no obvious physical meaning but can be related to a location parameter, μ , and scale parameter, σ [18]. The location parameter corresponds to the defect size at which $\text{POD} = 0.5$ and the scale parameter corresponds to the inverse of the slope of the curve. By reparametrizing $f(X)$ in terms of μ and σ such that

$$(\beta_0 + \beta_1(a)) = \frac{a - \mu}{\sigma} \quad (1.2)$$

it can be determined that $\beta_1 = \frac{1}{\sigma}$ and $\beta_0 = \frac{-\mu}{\sigma}$.

The author of *MIL-HDBK-1823A* developed accompanying software that utilizes the recommended statistical algorithms to produce the probability of detection versus defect size curves. This software, *mh1823POD*, utilizes the R-programming language and is freely available to anyone who requests it [19]. It provides diagnostic plots of each possible model and allows the user to specify the link function and size relationship which best fit the data. The model parameters are automatically estimated by the software using an iteratively reweighted least squares technique that accounts for the non-uniform variance in the error terms. Confidence bounds are computed by varying the parameters away from the values that provide the maximum likelihood in order to generate a family of curves that represent the specified confidence bounds on the maximum likelihood curve. For example, at 95% confidence, the confidence limits are expected to enclose the true POD curve in 95 out of 100 similar inspections. Note that in the case of hit/miss data, confidence bounds are not the same as prediction bounds [18]. Since the outcome of a single future inspection will be either a hit (1) or a miss (0), it is meaningless to consider a prediction that may be between 0 and 1. Ultimately, the following values typically used to describe NDE system capability can be extracted from the POD curves:

- a_{50} : defect size that has a 50% probability of detection (50% confidence)
- a_{90} : defect size that has a 90% probability of detection (50% confidence)
- $a_{90/95}$: defect size that has a 90% probability of detection (95% confidence)

1.6 Human Factors in Visual Inspection

Visual inspection performed by a human inspector is an inherently imperfect and unpredictable process. Individual inspectors have a wide range of physical and mental abilities, differing

responses to environmental stimuli, varied experiences and expectations, and fluctuating levels of focus and motivation. Since the 1950s, many industries whose quality control programs rely heavily on visual inspection, such as aircraft maintenance or airport baggage screening, have dedicated significant funds to research focused on understanding and improving visual inspection performance [20]. Due to the subjective nature of the task, the influence of human factors on inspection performance is the primary focus of much of the available research. A small selection of the human factors that have been investigated in the literature is discussed below. They can be divided into five categories: task factors, individual factors, environmental factors, organizational factors, and social factors [20].

Task factors are related to the physical nature of the inspection task and the structure. A systematic search pattern, area-by-area or defect-by-defect, and an inspector-paced approach have been found to improve inspection accuracy [21], [22]. Similarly, inspection accuracy increases as the defect rate increases and the number of defect types decreases [23]. Finally, detection rate tends to improve with increasing inspection time, although the improvement gradually saturates whereby additional time does not provide additional benefits [21], [24]. Breaks should be taken at regular intervals to avoid the performance decrement which occurs with sustained time on a vigilance task [25]. The rate of false alarms may also increase with time since more indications are likely to be found during a longer inspection and inspectors have been observed relaxing their rejection criteria as the number of indications increases [26].

Individual factors are related to the characteristics of the individual performing the inspection. Across the research, few individual factors have been found to have a consistent influence on inspection performance. A small number of recent studies have found inspection performance to decrease with age [27], [28], although prior research does not entirely support this finding [20]. When present, age effects are thought to be caused by a decline in visual acuity and cognitive processing speed that may or may not be offset by experience. Research has found that inspection accuracy may increase, decrease, or remain constant with experience. In some cases, more experienced inspectors are able to recognize abnormal attributes at a faster speed than novice inspectors [29]. Other studies have found no correlation or negative correlation between experience and inspection accuracy as inspectors may be overly reliant on knowledge gained from

past inspections and may overlook an uncommon defect in an unusual location [30], [31]. Most research measured only a small variation in visual acuity within the inspector population, thereby making visual acuity a poor predictor of performance, but an effective screening tool [21]. General intelligence and personality have not been found to have a significant influence on inspection performance, although specific intelligence tests, such as the Embedded Figures Test, have been found to correlate with inspection performance in some studies [31], [32]. In general, the research suggests that the interaction among the individual factors may be more significant than any single factor in understanding inspection performance.

Environmental factors are related to the atmosphere in which the inspection is being performed. Following the recommendations of the Illuminating Engineering Society, available research recommends lighting levels of 500 to 2000 lux based on the difficulty of the inspection task [24]. Different defect types may require different lighting types, and in addition to providing adequate illumination, it is important to control glare and eliminate hot spots. Although no research was found that specifically focused on the effects of temperature on inspection performance, the literature suggests that extreme temperatures (high or low) will negatively affect vigilance performance [33]. Finally, several studies suggest that inspection performance is significantly worse during overnight shifts. Inspectors working at night were more likely to miss defects and inaccurately estimate defect sizes [34].

Organizational factors are related to the administrative or managerial organization directing and supporting the inspection. Research frequently cites inspector training as the most cost effective and efficient strategy for improving inspection performance. Past studies have shown that training courses should address the physical, procedural, and cognitive aspects of the inspection task [20], [22]. Additionally, both feedforward information and feedback have been found to positively influence performance due to their informational and motivational properties. High quality information, related to inspection performance or strategy, can improve inspection accuracy and efficiency [22], [35].

Social factors are related to the social context in which the inspection is being performed. Research has shown that inspectors may relax their rejection policy due to external pressure [36] and consultation among inspectors has been found to increase performance [37].

1.7 UAS-assisted Bridge Inspection

The Federal Aviation Authority (FAA) defines an unmanned aircraft system (UAS) as “an unmanned aircraft and the equipment necessary for the safe and efficient operation of that aircraft” [38]. In the context of this research, UAS is assumed to refer to the UAS platform and its payload, the pilot, and the inspector. UAS technology is rapidly advancing for both recreational and commercial use, and in recent years, there has been increasing interest in the use of UAS for civil infrastructure applications, including bridge inspection. Advocates suggest that UAS assistance may address many of the logistical challenges confronting traditional bridge inspections, including access and traffic control. Their size, mobility, and ability to carry a wide range of imaging systems are oft-cited benefits, while flight restrictions and the short battery life are well known drawbacks. Despite the frequent endorsements, only a small number of studies have investigated the inspection capability of these systems.

The Minnesota Department of Transportation launched a three phase study in 2015 to demonstrate the effectiveness of UAS assistance for bridge inspections [39], [40], [41]. Conclusions and recommendations were generated based on a series of more than 40 field trials. The first phase was completed with a high-end, multi-purpose UAS platform, Phase II used an inspection specific aircraft, and the Phase II inspections were performed with a corrosion tolerant vehicle. The field trials considered applications beyond traditional bridge inspection, including 3D modeling and site mapping and collecting condition information to support rehabilitation projects. The findings suggest that UAS assistance may offer a cost-effective method for collecting detailed information about the bridge condition that might not be recorded during a typical routine inspection. Additionally, the reports recommend their use for long span bridges with heavy traffic and difficult access conditions. Although the UAS-assisted inspections were performed concurrent with the biennial NBIS inspections, a comparison of the inspection results was not presented.

In partnership with the Florida Department of Transportation, researchers at the Florida Institute of Technology performed a proof-of-concept study for UAS assistance in high mast light pole and bridge inspections [42]. This study included a comprehensive literature review, a detailed explanation of the UAS platform selection criteria, and a series of indoor and field trials using the selected system. Selection criteria for the platform included maneuverability, adaptability, software compatibility, payload, size, and user controls, while the criteria for the imaging system included performance, weight, dimension, price, battery life, ease of use, and software compatibility. Since the potential systems were ranked relative to each other, minimum technical requirements were not established. Findings from the field test suggest that images collected by the UAS were of similar or better quality to those collected by inspectors during a routine inspection. However, no direct assessment of inspection performance was provided.

In partnership with the Michigan Department of Transportation (MDOT), researchers at Michigan Technological University completed a two phase study investigating the use of UAS for a variety of civil infrastructure applications, including bridge inspection [43], [44]. During Phase II, five UAS platforms were considered, and the strengths and weaknesses of each were identified. For instance, the Bergen hexacopter could be equipped with optical and visual sensors to collect high resolution imagery of bridge decks while the DJI Mavic Pro quadcopter was best suited for close-up inspection of the bridge superstructure due to its collision avoidance sensors. Five bridge inspections were conducted during Phase II and the researchers determined that they were able to quickly and accurately collect information on the surface and subsurface condition of the deck that compared well to the findings from traditional methods, such as chain dragging and hammer sounding. This comparison was made visually, but no quantitative analysis was presented. Going forward, the research team and MDOT are interested in developing a methodology to automatically detect deck defects and assign element level condition ratings without closing lanes.

Researchers at Utah State University launched a study specifically focused on fatigue crack detection during UAS-assisted bridge inspections with the support of the Idaho Department of Transportation [45]. The researchers started with a series of controlled experiments in the laboratory to determine the maximum distance at which a fatigue crack could be detected under varying light conditions. This was not a “blind” test as the researchers knew both that a fatigue

crack was present and where it was located within the specimen. Still, they determined that a 1/2-inch crack could be detected using the built-in camera on the DJI Mavic Pro at a standoff distance of three feet and under normal lighting conditions. The researchers confirmed that this standoff distance could be achieved both in the laboratory and in a confined space underneath a bridge. One field inspection was performed at a steel bridge with known fatigue cracks over the Fall River in Idaho. The research team has difficulty flying the UAS in the GPS-denied environment under the bridge and over the river because the downward facing visual positioning system could not identify any feature points in the moving water. The inspection team was unable to identify the fatigue cracks because they were obscured by previous inspection markings, however they did identify corrosion, efflorescence, and concrete cracks. A final field demonstration was performed at the Purdue University's S-BRITE Center. The researcher, acting as the inspector, was able to identify fatigue cracks in the specimens, but also struggled to differentiate between actual fatigue cracks and other surface imperfections. The researchers concluded that fatigue crack detection with a UAS was possible, but many challenges exist.

Researchers from South Dakota State University worked with the United State Department of Agriculture and the South Dakota Department of Transportation (SDDOT) to perform a UAS-assisted inspection of a three span timber bridge near Keystone, South Dakota [46]. Before performing the inspection, the researchers considered 13 different UAS platforms, evaluating them based on set of metrics including flight time, cost, camera and video resolution, and camera location. Then, the team proceeded through the following five-stage bridge inspection methodology: (1) Bridge Information Review, (2) Site Risk Assessment, (3) Drone Pre-Flight Setup, (4) Drone-enabled Bridge Inspection, and (5) Damage Identification. Following the inspection, the results from the UAS-assisted inspection report were compared to the findings from the most recent SDDOT inspection report and the researchers determined that the same damage was reported in each. Additionally, researchers found that the 3D model developed using photogrammetry software provided a more complete view of the damage as compared to the video or still images. In terms of the UAS platform, the researchers discovered that the DJI Phantom 4 was able to provide a high quality inspection of the underside of the deck even though the camera was mounted underneath the body of the drone.

The Oregon Department of Transportation (ODOT) and researchers from Oregon State University performed six bridge inspections and three communication tower inspections with UAS assistance [47]. These inspections were performed with three different UAS platforms, two multi-purpose and one inspection specific. During the inspections, the team investigated various flight parameters including flight mode, flying speed, camera angle and field of view, aircraft standoff distance, and weather conditions. The researchers recommended using sensor-assisted or waypoint-assisted flight modes, in lieu of manual control, when possible. The recommended standoff distance varied based on weather conditions, camera capability, and flight mission, but in general, the authors recommended a standoff distance of at least 10 feet. For close range inspection missions, the researchers determined that the flying speed should be less than 2 mph to collect high-resolution imagery without blur. For overhead mapping flights with longer standoff distances, flying speeds of approximately 10 mph were found to be acceptable. Instead of identifying a single maximum operating wind speed, the researchers recommended identifying a wind-speed threshold specific to the aircraft and inspection conditions. Finally, the researchers concluded that UAS could most benefit routine and initial inspections and determined that UAS assistance reduces field time by 20% while increasing office time by 30%, on average. Although ODOT inspectors were involved in the field trials, the results from the UAS-assisted inspections were not compared to the results from routine inspections.

In 2015, two inspection missions were carried out at the behest of the National Forest Service on a timber pedestrian bridge located on the Kenai Peninsula in Alaska [48], [49]. The aircraft used in this inspection was a custom built hexacopter with an adjustable camera mount. First, a close range inspection of the exterior face of each timber truss was performed. This was intended to replicate a traditional routine inspection. Then, a series of 22 unique flights was completed to collect a sufficient number of overlapping digital images to build a 3D model of the structure. The 3D model developed from the UAS imagery was compared to a 3D model developed from LIDAR data, and the UAS imagery was determined to yield a more complete and detailed model. A quantitative assessment of the accuracy of the close range inspection or the 3D scene reconstruction was not provided, and the minimum imaging requirements were not discussed. The development of replicable UAS inspection methods and protocols was listed as a topic for future research.

2. PRIOR RESEARCH

A comprehensive review of previous research related to visual inspection performance and reliability was performed as part of the current research. Limited research has been performed specific to visual inspection of steel bridges, however, other industries that rely heavily on visual inspection for quality control have dedicated significant resources to understanding and improving inspection performance, and many of their findings may be relevant to the bridge industry. Therefore, visual inspection research from other industries, such as nuclear, aerospace, marine, and medical, was reviewed.

The following chapter contains a brief description of the five studies thought to be most significant to the current work. When necessary, references have been provided throughout this document to additional work reviewed. The first summary includes previous research from Purdue University that provided the foundation for the current research. These researchers designed and fabricated the Probability of Detection (POD) specimens, developed the POD inspection course, and conducted the initial set of 11 inspections. The second summary includes research that the Federal Highway Administration (FHWA) performed evaluating the reliability of routine and in-depth inspections of highway bridges. Next, a study from the Federal Aviation Administration (FAA) to benchmark visual inspection performance of aircraft inspectors is summarized. This is followed by a discussion of the best practices developed by the FAA to lessen the variability in inspection results caused by human factors. The final study includes research performed under NCHRP 12-28(7) to develop practical guidelines for inspectors and engineers tasked with evaluating the effects of corrosion on steel bridges members. Each of these studies offered great value to the current research.

2.1 Whitehead (2015) and Snyder (2015)

The Probability of Detection study developed at Purdue University aimed to assess the ability of a bridge inspector to locate and size fatigue cracks using current inspection procedures and techniques [50], [51]. Much of the current research continues the work of Snyder and Whitehead who initiated the POD study.

A training structure was designed and constructed at Purdue University's Steel Bridge Research, Inspection, Training, and Engineering (S-BRITE) Center to mimic a two span, three girder highway bridge. The structure supports 183 full size bridge specimens with known fatigue cracks. Inspectors were asked to complete a typical hands-on inspection of a subset of the specimens, and were evaluated based on their ability to correctly identify fatigue cracks. In addition to performance results, background information about each inspector and the environmental conditions during the inspection were recorded throughout the research.

This study included three types of common fatigue cracks: out-of-plane distortion induced cracks, weld toe cracks, and cracks emanating from rivet holes. The out-of-plane cracks and the weld toe cracks were introduced through cyclic fatigue loading in a controlled laboratory setting. The out-of-plane cracks were located in girder specimens (wide flange or plate girders) with welded web attachments and the weld toe cracks were located at welded cover plate terminations. The rivet hole cracks were simulated cracks, cut into plates using the electrical discharge machining (EDM) process. These plates were intended to represent any riveted tension member, such as a cover plate or truss chord. In all cases, the exact size, location, and orientation of the cracks was carefully measured and recorded to serve as the "answer key" for the inspectors participating in the study. All specimens were considered "full size" in that their proportions and geometry were representative of a typical highway bridge. For instance, the plate girders had a 3/8-inch by 36-inch web and two 1-inch by 12-inch flanges and the W-shapes were W36x135 rolled sections. To represent a greater number of in-service bridges, some of the specimens were painted with a typical bridge coating system (excluding the polyurethane top coat) while others were uncoated and exposed to a series of wet-dry cycles with a salt water solution until the surface corrosion resembled the protective rust patina characteristic of weathering steel. Based on the suggestion of officials at the Indiana Department of Transportation (INDOT), all of the welded cover plate and riveted plate specimens were painted since these details are not commonly used in weathering steel bridges. To imitate in-service bridges, all specimens (flawed and unflawed) were subjected to an accelerated weathering process. A combination of salt water, metal shavings, and scratches were applied to simulate typical damage and deterioration of the coating.

The specimens were suspended 25 feet above the ground. The painted specimens were installed on one span of the training structure and the uncoated specimens on the other span. Each span measured approximately 40 feet in length and the three girder lines were spaced approximately 8 feet apart. The plate girder and wide flange specimens were suspended from the frame beams, and the welded cover plates were attached to the bottom flanges of the girder specimens. The riveted plates were attached to bottom flange of the girder specimens and to the frame columns. A timber deck and timber framing were installed to mimic actual access and lighting conditions beneath a bridge.

Due to time constraints, only a subset of the specimens installed on the training structure were included in the inspection scenario. In total, the inspection course included 70 cracks in 147 painted specimens and 17 cracks in 16 weathering steel specimens. Of these, 45 were out-of-plane cracks, 23 were weld toe cracks, and 19 were rivet hole cracks. The ratio of unflawed to flawed specimens was 163:87, although it is important to note that many of the specimens included numerous possible crack sites making the actual noise ratio much higher. The crack lengths ranged from 1/2 to 5-3/8 inches with 37 cracks between 1-1/12 and 3-1/2 inches in length.

On the day of the inspection, the inspector reported to Bowen Laboratory at 8 AM. Two visual acuity tests were administered and the inspector was asked to sign a confidentiality agreement. Additionally, the inspector was given information on the inspection procedures, a list of assumptions about the specimens, and a blank inspection form for each specimen. At the test site, the inspector was first instructed to perform a routine inspection from the ground beneath the specimens. Then, the inspector proceeded with the hands-on inspection, first of the weathered specimens and then the painted specimens. The inspector was instructed to record the length and location of all cracks on the corresponding inspection form. If no cracks were found, the inspection form was to be completed indicating such. Throughout the inspection, the manlift was operated by a member of the research team to ensure that the specimens were inspected in the same predetermined order. At the conclusion of the hands-on inspection, the inspector completed an exit survey providing information about their background and training history.

Eleven (11) inspectors from the INDOT participated in the initial round of POD inspections. The results from these inspectors are combined with the results from 19 additional inspectors and presented in Chapter 3. They will not be discussed in detail here. However, even with a small sample size, the inspection results were highly variable with detection rates on the painted specimens ranging from 31% to 77%. The average 50% detection rate crack length (a_{50}) was 1-1/4 inches and the average 90% detection rate crack length (a_{90}) was 8 inches. No single variable correlated significantly with detection rate. Moreover, the detection rates on the uncoated steel specimens were so low (0% to 39%) that the research team deemed them invalid and they were not included in the subsequent inspections. Recommendations for improving the reliability of visual inspection included implementing performance testing for inspector qualification, requiring the use of a standard set of tools during hands-on inspections, and developing detail-specific inspection procedures.

2.2 Moore, Phares, Graybeal, Rolander, & Washer (2001)

In 1998, the FHWA launched a comprehensive study focused on the reliability of routine and in-depth inspections of highway bridges [5]. The objectives of this study were to provide an overall measure of the accuracy and reliability achieved with current inspection standards, understand the influence of key factors expected to affect inspection results, and identify differences in inspection procedures among the participating state agencies. This study included 49 inspectors from 25 state agencies participating in 10 separate inspection tasks.

Most relevant to the current research project were the findings from Tasks F and H which each included the in-depth inspection of a bridge superstructure. Task F required a partial inspection of a 90-foot single span, riveted plate girder bridge. Overall, the bridge was in fair condition with general deterioration of the coating system and localized corrosion of the girders, secondary members, and fasteners. In addition to the global defects, five local defects were identified in the bridge structure. Two (2) of these defects, a missing rivet head and a crack indication at a tack weld, were implanted by the research team. The other defects consisted of localized member distortion due to impact in two locations and abnormal rotation at a bearing. The research team expected that these defects would be noted in an in-depth inspection report.

Based on proctor observations and self-reports from the inspectors, the majority of the inspectors were focused, not rushed, and comfortable during this inspection task. Despite this, few of the inspectors completed a thorough inspection with no more than 90% of the inspectors inspecting any single feature. Additionally, only 48% of the inspectors used a flashlight, despite measured lighting levels far below the recommended values, 36% of the inspectors used a tape measure, and 5% of the inspectors used a magnifying glass. Notably, when asked to describe the geometric characteristics of the bridge, only 10% of the inspectors noted the bridge skew. This is an important feature as bridge skew may lead to distortion induced cracking which should be a primary focus during an in-depth inspection.

While the majority of the inspectors noted the general paint system failure and corrosion, only 45% of the inspectors noted the severe fastener corrosion and section loss. Additionally, only 50% of the inspectors observed the bearing abnormality, 17% noted the impact damage at one or both locations, 7% located the crack indication, and 5% identified the missing rivet head. In this task, defect detection was found to correlate with inspection duration, mental focus, flashlight use, and inspector-reported structure complexity. Inspectors that worked more slowly, were more focused, used a flashlight, and rated the bridge as more complex tended to note one or more of the local defects. Conversely, inspectors that worked quickly, were less focused, did not use a flashlight, and rated the structure as less complex tended to indicate that there were no defects.

Task H required the inspection of a single bay in one span of a four span, welded plate girder bridge. The superstructure contained connection details typical of major highway bridges, including longitudinal and transverse stiffeners, bolted and welded flange transitions, and a lateral bracing system attached to the girders with gusset plates welded to the web near the bottom flange. Overall, the bridge was in good condition, although there were crack indications at the weld toes of some of the connection details.

Once again, despite being given adequate time to complete the task, the majority of the inspectors did not perform a thorough inspection. For instance, although the lateral gusset plate connection includes a Category E fatigue detail and is a likely location for fatigue crack formation, only 56% of the inspectors inspected at least 76% of these connections. Similarly, only 47% of the inspectors

inspected all locations where the drain pipe was welded to the girder web with a short attachment. Despite the presence of poor fatigue details, only 57% of the inspectors indicated that fatigue cracking was a possible issue in this type of structure. Additionally, when asked to describe the bridge geometry, only 52% of the inspectors noted that it was continuous.

In total, there were 174 likely crack locations and 7 weld crack indications within the inspection limits. Four (4) of the indications were located at vertical stiffener to girder connections and three of the indications were located at the lateral gusset plate to girder connection. Additionally, the drain pipe to girder connection and utility bracket to girder connections were classified as likely crack locations, although no indications were present. Out of 304 observations made at the locations with crack indications, only 12 observations correctly identified a crack indication (3.9%). Additionally, out of 7,234 observations made at locations without crack indications, 43 indications were identified (0.6%). Finally, only 16% of the inspectors correctly identified at least one weld crack indication and only 41% indicated the presence of any weld crack indications within the inspection limits.

In this task, crack detection was found to correlate with inspection duration, inspector- and proctor-reported comfort level, inspector-reported structure complexity and accessibility, observed variation in viewing angle and distance to weld, flashlight use, and number of annual bridge inspections. Inspectors that worked more slowly, were more comfortable working at heights, rated the bridge as more complex and less accessible, varied their viewing angle, viewed the weld from a shorter distance, used a flashlight, and inspected fewer bridges annually tended to note one or more of the weld crack indications. Conversely, inspectors that worked more quickly, were less comfortable in the lift, rated the bridge as less complex and more accessible, viewed the weld from a constant angle and greater distance, did not use a flashlight, and inspected more bridges annually tended to indicate that the bridge contained no weld crack indications. Additionally, 86% of the inspectors that indicated before the inspection that fatigue cracking was a possible issue correctly located at least one weld crack indication. Similarly, although not directly related to the likelihood of fatigue crack formation, 71% of the inspectors that noted that the bridge spans were continuous on the pre-task questionnaire correctly identified at least one weld crack indication.

In summary, the research team concluded that in-depth inspections may not provide additional information beyond that noted in a routine inspection. Additionally, inspectors that found small, detailed defects in one task tended to perform similarly on the other tasks, indicating that performance does not depend on bridge type. Final recommendations included improved training on bridge behavior to increase the knowledge base of the inspectors and considering the factors identified as affecting performance during inspector selection and training.

2.3 Spencer (1996)

Recognizing the role of visual inspection in ensuring the safety of the nation's civil air fleet, the FAA tasked its Aging Aircraft Nondestructive Inspection Validation Center (AANC) with developing a program to evaluate the reliability of these inspections [52]. The objective was to determine the benchmark capability of visual inspection performance in major airline maintenance facilities. Twelve (12) inspectors from four major airlines performed 10 controlled inspection tasks to establish a benchmark measure of visual inspection capability. Test specimens included a retired Boeing 737 aircraft and manufactured components from the AANC Specimen Library. A baseline inspection of the Boeing 737 was performed by AANC staff to catalog the locations of the existing defects, including cracks, corrosion, and general wear-and-tear. The presence and size of the cracks was verified with eddy-current inspection. Tasks 501 through 510 required inspection of specific regions of the Boeing 737 while Task 701 was completed using fabricated lap splice specimens. The full complement of tasks was expected to represent the conditions an inspector would encounter during routine maintenance inspections. Separate job cards describing each task were provided to the inspector. Information about each inspector and the inspection environment were recorded throughout the study. Additionally, each inspection was video recorded.

Inspection results from Task 701 were used to develop POD curves relating inspection performance to defect size. The lap splice specimens included 117 cracks at 382 possible locations. The 50% detection rate crack lengths (a_{50}) for the 12 inspectors varied between 0.09 inches and 0.32 inches while the 90% detection rate crack lengths (a_{90}) ranged from 0.16 inches to 0.91 inches. The average a_{50} was 0.13 inches and the average a_{90} was 0.26 inches. For this task, the research team attempted to categorize each error as either a search error or a decision error based on a

review of the video tape. On the video, the transition from the search function to decision function was indicated by a lengthy pause during which the inspector used a flashlight or repeatedly adjusted their viewing angle. Performance on the search task was consistently poor, with detection rates between 44% and 69%, while performance on the decision task was more variable. Some inspectors were perfect, making correct decisions each time they entered the decision step, while others made incorrect decisions more than 50% of the time. From this, the research team concluded that all inspectors needed supplemental training on how to perform the search task, while only a few inspectors required additional guidance on performing the decision task.

A separate probability of detection analysis was carried out on 48 identified cracks within the Boeing 737 aircraft. These cracks were expected to be located during Tasks 501 through 510. The relationship between probability of detection and length was far less obvious among this set of cracks. The average a_{50} for this sample of cracks was 4-1/2 inches with 95% confidence bounds of 2 and 72 inches. This far exceeds the average 50% detection rate crack length calculated from Task 701. The research team was not surprised by the difference since Task 701 focused on a single defect type in a single detail while the other tasks required inspectors to consider multiple defect types in a range of details. Considering this, the research team recommended that POD curves should only be fit to specific tasks and inspections conditions.

Other relevant findings from Tasks 501 through 510 include:

- The use of a light shaping diffuser did not have a statistically significant effect on crack detection. The diffuser was intended to provide more uniform illumination, thereby reducing glare and hot spots.
- Intimate familiarity with the aircraft model improved performance. Inspectors with extensive previous experience with this model of aircraft were more likely to locate cracks in known problem areas. In a few cases, inspectors did exhibit “tunnel vision” due to their expectations, and one inspector missed a clearly visible defect in an adjacent area because he was overly focused on an area where he anticipated cracking.
- More experienced inspectors, as measured by aviation experience not inspection experience, performed better.
- Increased inspection time did not improve crack detection performance.

- Performance on one task was not a reliable predictor of performance on another task.

2.4 Drury and Watson (2002)

Tasked by the FAA with developing human factors guidance to improve the reliability of visual inspection, Drury and Watson used the hierarchical task analysis technique to provide best practices for each of the five identified sub-tasks within a visual inspection: initiate, access, search, decision, and response [21]. The initiate, access, and response functions are primarily manual. They require a physical action by the inspector to gather the appropriate information and document the findings. Conversely, the search and decision functions are primarily cognitive. These sub-tasks require the inspector to interpret, analyze, and categorize visual cues and are both the most critical steps and the most error prone. To be successful in the inspection process, the inspector must be proficient in both the search and the decide functions, therefore it is often useful to determine if an underperforming inspector requires intervention in the search or the decision task. Misses can result from errors in the decision process or the search process, while false positives result from flawed judgements. A summary of the inspection process and best practices for visual inspection are provided in the following paragraphs.

The search function involves examining the inspection surface for possible defects. This step is performed iteratively with the decide step as each possible indication is either accepted as a flaw or rejected. The search process should be designed to improve the visibility of flaws. Providing good access and utilizing the appropriate tools will improve the likelihood of detecting a defect. Additionally, inspector comfort can improve task motivation and adequate illumination may make flaws more distinguishable from the background. To ensure a thorough inspection, a search strategy, informed by the objective of the inspection, is required. In research, inspectors trained to use a systematic search strategy consistently perform better than inspectors that employed a random search strategy. A systematic search pattern not only ensures that the entire inspection surface is inspected, but it helps to eliminate repeat inspections of the same area.

The decision function involves determining if an indication is a recordable flaw. This step is performed iteratively with the search step and search resumes if the indication does not require documentation. Often, the search and decide tasks occur nearly simultaneously and are

indistinguishable to an observer; however, at times, there is a clear transition from “searching” to “deciding”. The use of tools or alternate senses may indicate that the inspector has detected an indication and is seeking more information to support the decision. Similar to the search function, the decision function should be guided by a decision strategy. Unlike more advanced NDE techniques where a quantitative measure of the signal is used to make decisions, in visual inspection each inspector employs their own response criterion to separate noise from signal-plus-noise. This strategy should be informed by the payoff matrix which considers the relative cost of each decision. The payoff matrix assumes that there are four possible outcomes for each decision and each outcome has an associated probability and an associated cost. The optimal response strategy balances probability and cost considering the acceptable level of risk and available resources. Final decisions are made based on an inspector’s experience, training, and expectations. Expectations must be based in reality and experience and training should complement each other. Knowledge based reasoning should be used when experience or training are insufficient. Once again, tools should be used to improve the discriminability between flaws and background noise.

In visual inspection, there is a well-documented speed versus accuracy trade-off. As speed increases, accuracy decreases. This relationship exists in both the search and decision steps, although improvements in decision making saturate quickly with decreasing speed. A thorough understanding of the structure or element being inspected can improve accuracy without slowing the inspection. Similarly, focusing on critical areas will reduce inspection time without sacrificing accuracy. Another well documented inspection phenomenon is the decline in performance with increasing time on that task, typically referred to as the vigilance decrement. In research settings, detection performance decreased steadily over the first 20 to 30 minutes of the task and then remained at the lower level for the remainder of the task. Although the degree of the decline varied across inspection types, it is most severe when defects are infrequent and difficult to detect. The vigilance decrement may affect the search and decision components differently. A change in sensitivity, identified by a decrease in detection rate without a corresponding reduction in false alarms, results in a true decline in performance due to an increasing inability to distinguish between true cracks and cracklike surface defects. A change in bias, identified by a decrease in both the detection rate and the number of false alarms, indicates an increasing reluctance to locate or record

any indication, regardless of its true nature. To overcome this drop in performance, breaks should be taken at regular intervals.

2.5 Kulicki, Prucz, Sorgenfrei, & Mertz (1990)

In 1990, the National Cooperative Highway Research Program (NCHRP) released Report 333, *Guidelines for Evaluating Corrosion Effects in Existing Steel Bridges*, to provide practical guidelines for inspectors and engineers tasked with evaluating the effects of corrosion on steel bridges [53]. The report is divided into four parts, with the first part providing field inspection guidelines and the second offering office evaluation guidelines. Two levels of inspection and evaluation are defined. A Level I evaluation is performed by current inspection staff and is focused on quantifying the degree of damage while a Level II evaluation includes corrosion experts and objects to identify the corrosion process and methods for controlling or mitigating the damage.

Relevant to this research is the discussion of crevice corrosion in built-up tension members. Crevice corrosion is defined as “a form of localized corrosion occurring at confined locations where easy access to the outside environment is prevented” [45, p. 8]. It typically occurs in the gaps between mating surfaces and is caused by the variation in environment inside and outside the crevice. The gaps between adjoining plates or shapes in built-up members are especially vulnerable to the development of crevice corrosion. Beyond section loss, crevice corrosion can cause significant member distortion since the products of the chemical reaction have a greater volume than the reactants. The volume change depends on the exact nature of the reaction, but the volume of rust is typically two to four times greater than the steel consumed in the process. When this reaction occurs in confined spaces, the build-up of corrosion product can generate pressures of up to 10,000 psi which may cause plastic deformation in members and fasteners.

The field inspection guidelines intend to clarify and standardize corrosion inspection procedures. These guidelines cover the various types of corrosion, corrosion susceptible bridge details, and opportunities to mitigate corrosion damage. A Level I evaluation may be supported by either a cursory or general inspection. A cursory inspection does not require access to the inspection surface or measurements of the section loss. A general inspection requires hands-on access and a few spot measurements of section loss representative of the general and worst case conditions. A

combination of measurement and estimation is used to quantify the extent of corrosion damage. A Level II evaluation requires a detailed inspection with hands-on access to all members and precise measurements of remaining thickness. In both the general and detailed inspections, measurements are generally made with calipers or an ultrasonic thickness gauge.

The office evaluation guidelines intend to provide comprehensive and practical advice for evaluating the effects of corrosion on the structural capacity of steel bridges. These guidelines include a description of the effects of corrosion on structural capacity, various analysis methods, and relevant performance criteria. Specifically, the guidelines consider the effects of corrosion on a member's strength, stiffness, stability, and resistance to fatigue. Focusing on tension members, the reduction in cross sectional area due to uniform corrosion will increase the stress in the member and decrease its stiffness. Therefore, the member should be evaluated for yielding on the gross section, fracture on the net section, and displacement. Additionally, the slenderness ratio of the member should be calculated with the reduced section and compared to the limiting ratio defined in Section 6.8.4 of the *AASHTO LRFD Bridge Design Specifications* [54]. This limit was established to prevent excessive sag, vibration, and lateral movement of tension members. The guidelines recommend that in areas of localized corrosion, a tension member should be evaluated for yielding on the net section of the remaining member. In built-up members that rely on lacing bars or batten plates for load sharing, deterioration of these secondary elements might reduce their ability to redistribute load and increase the slenderness of individual segments. In truss bridges, section loss may cause load redistribution away from the corroded member. For section losses exceeding 20%, the guidelines recommend neglecting the effects of load redistribution on the corroded member, but considering the effects on adjacent members.

The negative effects of corrosion on fatigue resistance are due to both the loss of section and the introduction of stress concentrations at localized corrosion sites, such as pits. For corroded members that have been blast cleaned and repainted such that no active corrosion sites are present, the effects of pits can be accounted for by calculating the stress in the member with the reduced section and applying a pitting factor to the resulting Category A fatigue life. The stress concentrations inherent in the lower fatigue categories are assumed to be greater than any generated by corrosion. In structures with active corrosion, subsequent research recognized that

fatigue cracks were rarely detected at corrosion notches before significant section loss occurred, and determined that the ongoing corrosion process actually blunts potential fatigue initiation sites [55]. Therefore, fatigue resistance in these members can be determined based on the remaining section thickness.

The report states that the effects of unintended pressure and distortion due to corrosion build-up can be neglected in the strength evaluation of tension members, except in the case of severe deterioration. Although the distortion may alter the section properties of the member and introduce eccentricities, the overall strength of the member will not be reduced since the performance criteria allow for areas of localized yielding and tension members are self-stabilizing.

3. HANDS-ON VISUAL INSPECTION

3.1 Introduction

Using the probability of detection specimens and inspection procedures developed by Snyder [50] and Whitehead [51], 30 bridge inspectors were invited to complete a typical hands-on inspection of each specimen. The intention of this study was not to evaluate the performance of any individual inspector, but rather to assess the performance of the group of inspectors in order to understand the ability of the average bridge inspector following the current inspection procedures.

The inspectors were evaluated based on their ability to correctly identify fatigue cracks and distinguish cracks from less severe surface defects (corrosion, debris, scratches, etc.). In addition to performance results, environmental conditions and inspector attributes were recorded during the inspections. During each inspector's hands-on inspection, data were collected from the completed inspection forms, proctor observations, and pre- and post-inspection evaluations.

The inspection set-up and results are discussed in detail in the following sections. The methods used to evaluate the data will be explained along with findings, recommendations, and conclusions regarding visual inspection of steel bridge members.

3.2 Inspection Scenario

The test specimens and inspection course are described in detail in Snyder [50] and Whitehead [51], and only a brief summary will be provided herein.

The inspection procedures were established to simulate an actual field inspection in a controlled testing environment. Only minor changes to the inspection procedures were made throughout the study to allow for a direct comparison of the results. The first 11 inspectors that participated in the study completed three separate inspection tasks: a routine inspection from the ground beneath the specimens, a hands-on inspection of the weathered steel specimens, and a hands-on inspection of the painted steel specimens. However, the first two tasks were ultimately eliminated to ensure

adequate time for the primary focus of the study, the hands-on inspection of the painted steel specimens.

The training structure, shown in Figure 3.1, was designed and constructed at Purdue University's Steel Bridge Research, Inspection, Training, and Engineering (S-BRITE) Center to model a two span, three girder highway bridge. The structure supports 183 full size bridge specimens with known fatigue cracks. Considering the typical spacing of these details on an actual bridge, the test frame represents a bridge of approximately 450 to 500 feet. The specimens are suspended 25 feet above the ground and a timber deck and timber framing were installed to mimic typical access and lighting conditions. Due to time constraints, only a subset of the specimens installed on the training structure were included in the inspection scenario. After eliminating the weathering steel specimens, the inspection included 72 girder specimens, 16 riveted plate specimens (8 mounted overhead and 8 mounted vertically), and 59 welded cover plate specimens. Each inspector inspected the same specimens in the same order.

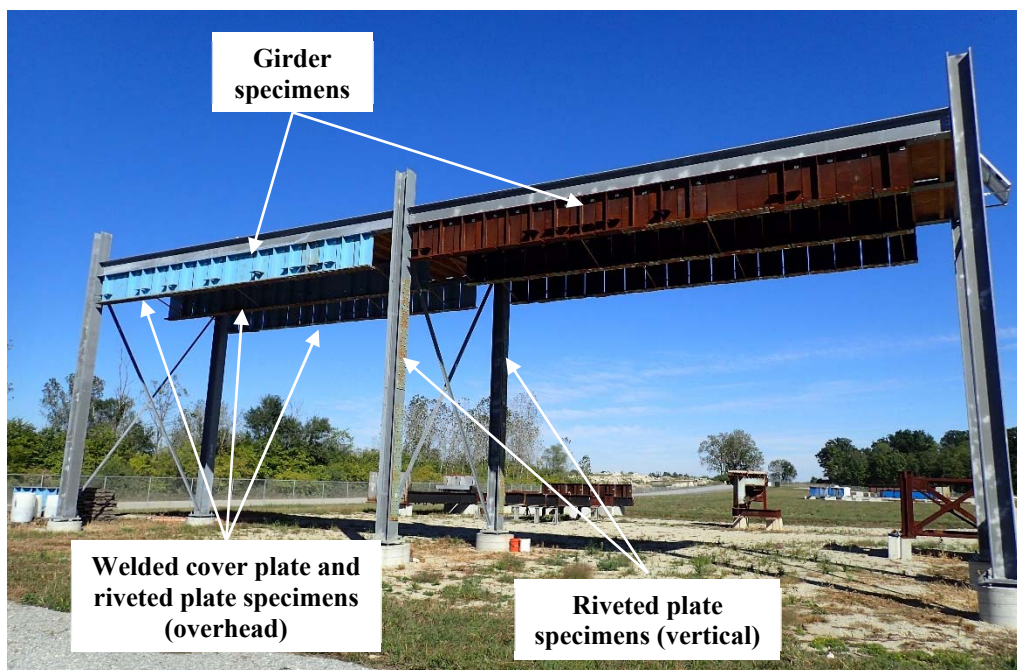


Figure 3.1 Probability of detection training structure with POD specimens

The test articles for this inspection consisted of 147 bridge specimens fabricated in a controlled laboratory setting to ensure that the exact length and location of all 70 fatigue cracks were known. As shown in Figure 3.2 and Figure 3.3, three unique types of specimens and cracks were used:

plate girder or wide flange sections with out-of-plane distortion induced cracks, welded cover plate terminations with weld toe cracks, and riveted plates with rivet hole cracks. The painted specimens were coated with a typical bridge paint system, excluding the polyurethane top coat. To imitate in-service bridges, all specimens (flawed and unflawed) were subjected to an accelerated weathering process. A combination of salt water, metal shavings, and scratches were applied to simulate typical damage and deterioration. Each specimen had one crack, multiple cracks, or no cracks. Although the specimens were fabricated in a controlled laboratory setting, all cracks were sized, oriented, and located as they typically are in the field. A detailed discussion of the preparation of the specimens is provided in [50].

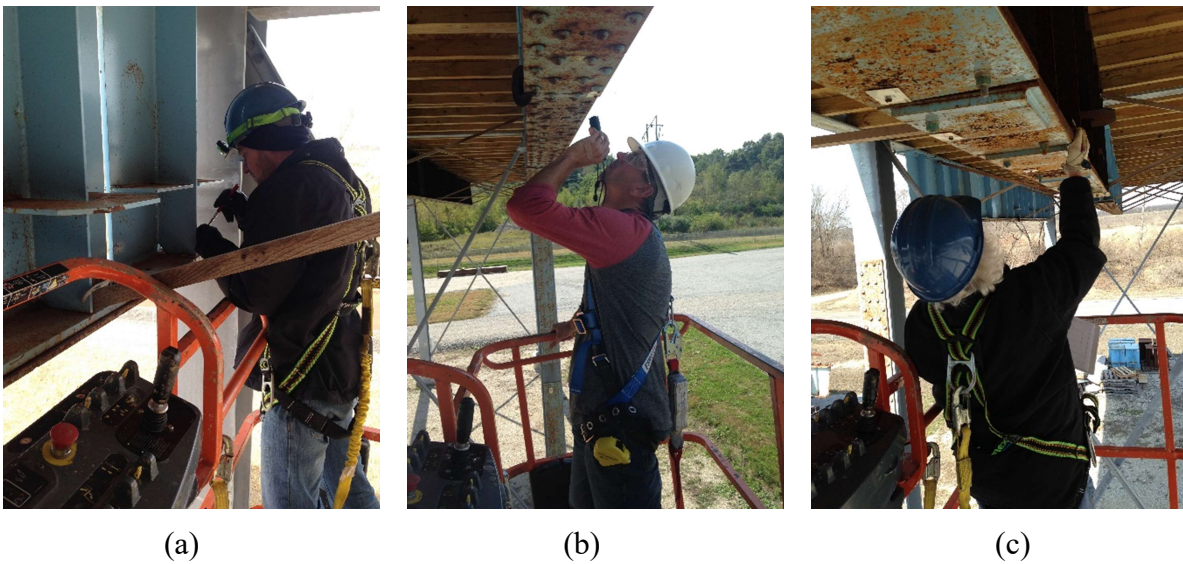


Figure 3.2 Hands-on inspection of the (a) girder, (b) overhead mounted riveted plate, and (c) welded cover plate specimens

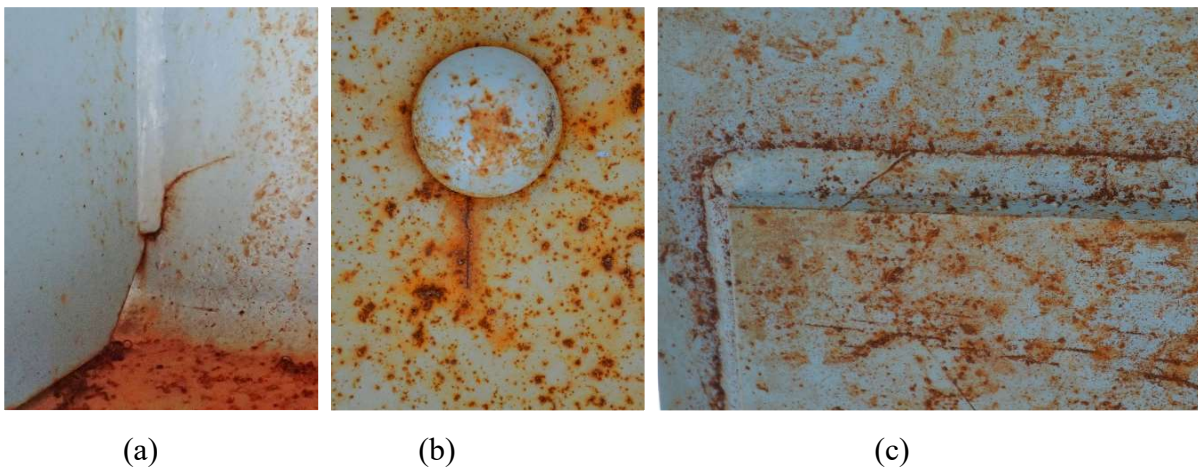


Figure 3.3 Photograph of a typical (a) out-of-plane crack, (b) rivet hole crack, and (c) weld toe crack

On the day of the inspection, the inspector reported to the test site at approximately 8 AM. Two visual acuity tests, a near visual acuity test and a contrast sensitivity test, were administered and the inspector was asked to sign a confidentiality agreement. The near visual acuity test was administered with the Jaeger chart and the contrast sensitivity test was performed with the Pelli-Robson chart. Additionally, the inspector was given the inspection procedures, a list of assumptions about the specimens, and a blank inspection form for each specimen. Each form included a scaled drawing of the specimen on which the inspector could record the location of any detected crack(s). If no cracks were found, the inspection form was still to be completed indicating such. The inspection procedures and instructions, confidentiality agreement, and sample inspection forms are provided in Appendix A.

Each inspection was completed individually over approximately 6 hours, depending on the inspector's pace. Throughout the inspection, a manlift was operated by Purdue research staff to ensure that the specimens were inspected in the same predetermined order. During the inspection, the research staff also recorded general observations about the inspection, including notes about the weather, the inspection start and end times, tools used, and general impressions of the inspector and the inspection strategy. Temperature and wind speed statistics were recorded by the KLAF weather station at the Purdue University Airport, located adjacent to the test site [56]. At the conclusion of the exercise, the inspector completed an exit survey providing information about their background and training history. This document is included in Appendix A.

Inspections were conducted outdoors on the campus of Purdue University in West Lafayette, Indiana at the S-BRITE Center between October 2014 and December 2016. Inspections were typically performed between the hours of 8 AM and 4 PM and they were performed in any weather conditions, with the exception of heavy rain and lightning.

No time restrictions were placed on the inspectors and they were encouraged to use any tools that they would normally use during a hands-on inspection. In order to not bias future inspections, inspectors were asked to not use wire brushes or grinders that would leave lasting marks on the specimens, but they could lightly clean an area if needed. Since none of the cracks were entirely obscured by dirt or debris, this was not expected to have a significant effect on crack detection.

3.3 Inspector Demographics

Thirty (30) inspectors (27 males and 3 females) with experience in bridge inspection and/or bridge design participated. In order to accurately represent the current bridge inspector population, inspectors from a wide range of backgrounds were invited to participate in the study. This project included 16 inspectors from state departments of transportation (DOT), 12 inspectors from private engineering and/or inspection firms, and two inspectors from federal agencies. Participating inspectors were based in four states, Indiana (22), Illinois (6), Virginia (1) and California (1), although most of the private consultants perform bridge inspections around the country. Twenty-six (26) of the 30 inspectors had completed the two-week FHWA/NHI *Safety Inspection of In-Service Bridges* training course prior to their participation in the study. Of the remaining four participants, one had 30 years of experience inspecting railroad bridges, one was a Certified Weld Inspector and NDT technician with over 25 years of experience, one was a licensed professional engineer with a PhD in civil engineering and nearly 20 years of design and research experience specific to steel bridges, and one was a junior engineer in the process of completing the bridge inspector training and certification requirements. The average experience of the participating inspectors was 10.6 years and the inspectors had completed an average of 10.6 hands-on inspections in the 12 months prior to their participation. Twenty-three (23) of the 30 inspectors possessed an engineering license; six were engineering interns (EI), 15 were professional engineers (PE), and two were licensed structural engineers (SE). All 30 inspectors could read the smallest paragraph of text on the test card for the Jaeger vision test and 26 of the inspectors recorded a Log Contrast Sensitivity of 1.95 on the Pelli-Robson vision test. Select inspector demographics are compiled in Table 3.1.

Table 3.1 Inspector demographics for hands-on inspections

Inspector ID	Employer	Age (years)	Inspection Experience (years)	Professional Licensure	Number of Hands-on (Routine) Inspections	Log Contrast Sensitivity
10EH-10	State DOT	56	30	None	0 (120)	1.95
11CO-02	State DOT	46	11	None	40 (50)	1.95
18GA-03	Private Consultant	63	29	PE	4 (20)	1.95
13CA-09	State DOT	58	20	None	6 (200)	1.95
09SD-08	Private Consultant	55	28	None	20 (0)	1.95
20MD-19	State DOT	45	20	PE	6 (100)	1.95
04MY-41	Private Consultant	40	15	PE	9 (300)	1.95
09ME-03	Private Consultant	28	6	PE	5 (5)	1.95
08GS-32	State DOT	30	1.25	EI	5 (35)	1.95
01VM-02	Private Consultant	30	6	PE	23 (600)	1.95
12AE-04	State DOT	41	1	PE	1 (40)	1.95
11LB-22	State DOT	55	8	None	4 (145)	1.65
10JW-16	State DOT	40	10	PE	9 (200)	1.95
21RI-01	Private Consultant	33	10	PE	5 (112)	1.95
08LK-23	Federal Agency	38	0	PE	0 (0)	1.95
06MA-03	State DOT	56	10	PE	3 (6)	1.8
21GZ-14	Federal Agency	55	30	None	0 (100)	1.95
09VK-18	State DOT	25	2.5	None	8 (25)	1.95
18RT-25	State DOT	53	13	PE	10 (240)	1.95
06JD-15	State DOT	33	8	EI	20 (163)	2.10
06TI-56	State DOT	24	0.75	EI	0 (175)	1.95
27GH-57	Private Consultant	29	3	EI	18 (35)	1.95
01DS-23	State DOT	28	4	PE	30 (450)	1.95
24BR-25	State DOT	39	1	PE	25 (45)	1.65
27PC-37	Private Consultant	24	1.5	EI	0 (20)	1.95
12LA-04	Private Consultant	45	20	PE	25 (700)	1.95
11NH-05	Private Consultant	32	10	EI	13 (250)	1.95
10CA-07	State DOT	42	14	PE	20 (70)	1.95
26RO-49	Private Consultant	30	5	PE/SE	0 (4)	1.95
25HQ-08	Private Consultant	33	1	PE/SE	10 (50)	1.95

3.4 Inspector Response Evaluation

Upon the completion of the course, each inspector submitted their binder to Purdue research staff. The binder submitted by the inspector was compared to the answer key binder which included an AutoCAD drawing of each of specimen that stated “NO DEFECT” for the unflawed specimens or

were considered hits if they were shown in the vicinity of a known fatigue crack, and were otherwise considered false calls.

- Miss: A miss was assigned for any crack that was not indicated as a crack or possible crack by the inspector, i.e. false negative.
- False Call: A false call was defined as a crack reported in a region of the specimen without a known defect, i.e. false positive. This definition was applied to all indications labelled as a “possible” or “probable” crack and any notes requesting follow-up inspection or testing. Although typically less consequential than a miss, a false call requires resources to be expended inefficiently and undermines the veracity of the process.
- Detection Rate: The detection rate was defined as the number of hits divided by the total number of cracks.
- False Call Rate: The false call rate was defined as the number of false calls divided by the total number of possible crack locations. Although most research evaluates performance based on the false call rate or the probability of a false call, it was not practical in this study because the total number of possible crack locations could not be easily quantified.
- Hit/Call Ratio: The hit/call, or hit-to-call, ratio indicates how often an inspector’s recording of a crack was actually a crack. The hit/call ratio was calculated for each inspector by dividing the number of hits by the total number of calls (hits plus false calls). A higher, or better, hit/call ratio indicates that the inspector required fewer calls to record a hit as compared to an inspector with a lower, or worse, hit/call ratio.

3.5 Inspection Results

Inspector evaluation considered 70 visually detectable cracks with lengths between 1/2 and 5-3/8 inches in 147 painted specimens. Two of the overhead mounted riveted plates were removed before the final two participants completed their inspections. For these inspectors (08GS-32 and 06TI-56), the total number of visually detectable cracks was 68. The inspectors were awarded a ‘1’ for each hit and a ‘0’ for each missed. From this, the detection rate could be easily calculated from the total number of hits divided by 70 (or 68). The total number of false positives was tallied and recorded for each specimen and then summed over the 147 specimens to determine the total number of false positives. Additionally, the inspection instructions directed the inspectors to record the length for each identified crack. For the hits, the recorded length was compared to the

actual length to assess the inspector's accuracy in measuring or estimating crack length in the field. Inspector performance was evaluated for the entire specimen inventory and for each type of specimen (girder, welded cover plate, riveted plate) separately.

It is important to note that the findings of this phase likely represent the upper bound on visual inspection performance. The POD inspections were performed in isolation, free from many of the distractions (such as looking for other damage or traffic) and time pressures present during a typical inspection. Additionally, unobstructed access to the inspection surface was provided and cracks were present at a higher than average rate based on the author's experience. Finally, although inspectors were encouraged to perform the inspection as they would a typical hands-on inspection, they may have been motivated to excel by the presence of the proctor.

3.5.1 Crack Detection and False Calls

Thirty (30) inspectors completed a hands-on inspection of the painted specimens. A summary of the results is provided in Table 3.2 and the individual result for each inspector is shown in Table 3.3. The most successful inspector detected 60 of the 70 possible cracks (86%). The least successful inspector detected 22 of the 70 possible cracks (31%). On average, the inspectors detected 46 cracks (65%), meaning 24 cracks were not located. The standard deviation of detection rate was 14% indicating that the majority of the inspectors detected between 36 and 55 of the cracks. The number of false calls made during the inspections ranged from 14 to 268. The average number of false calls was 90 with a standard deviation of 67. The average time to complete the inspection was 248 minutes and the standard deviation was 76 minutes. The fastest inspector completed the inspection in under two hours (116 minutes) and the longest inspection lasted over seven hours (457 minutes).

Table 3.2 Summary of results from hands-on inspections

Specimen Type/Crack Type	Detection Rate			Number of False Calls		
	Mean (%)	Standard Deviation (%)	Minimum/Maximum (%)	Mean	Standard Deviation	Minimum/Maximum
All Specimens	65	14	31/86	90	67	14/268
Girder/Out-of-Plane	62	14	18/82	65	54	5/205
Welded Cover Plate/Weld Toe	74	27	4/96	16	14	0/59
Riveted Plate/Rivet Hole	59	17	21/95	9	9	0/34

Table 3.3 Hands-on inspection results by inspector

Inspector ID	Hits	Possible Cracks	Detection Rate	False Calls	Hit/Call Ratio	Total Time (minutes)
10EH-10	26	70	37%	156	14%	175
11CO-02	50	70	71%	60	45%	279
18GA-03	33	70	47%	56	37%	309
13CA-09	22	70	31%	14	61%	224
09SD-08	49	70	70%	51	49%	457
20MD-19	45	70	64%	23	66%	236
04MY-41	41	70	59%	85	33%	327
09ME-03	53	70	76%	62	46%	226
08GS-32	47	68	69%	105	31%	269
01VM-02	55	70	79%	31	64%	296
12AE-04	46	70	66%	42	52%	179
11LB-22	47	70	67%	103	31%	224
10JW-16	50	70	71%	174	22%	320
21RI-01	58	70	83%	187	24%	368
08LK-23	59	70	84%	68	46%	279
06MA-03	53	70	76%	195	21%	241
21GZ-14	35	70	50%	18	66%	206
09VK-18	46	70	66%	89	34%	138
18RT-25	42	70	60%	146	22%	202
06JD-15	30	70	43%	47	39%	262
06TI-56	43	68	63%	57	43%	211
27GH-57	49	70	70%	76	39%	217
01DS-23	54	70	77%	37	59%	228
24BR-25	46	70	66%	15	75%	160
27PC-37	33	70	47%	217	13%	116
12LA-04	55	70	79%	162	25%	347
11NH-05	60	70	86%	77	44%	345
10CA-07	42	70	60%	20	68%	192
26RO-49	52	70	74%	67	44%	213
25HQ-08	47	70	67%	268	15%	192

The number of false positives were compared to the number of hits for each inspector and the hit/call ratio was calculated. The highest, or best, hit/call ratio was 75% and the lowest, or worst, hit/call ratio was 13%. The average hit/call ratio was 41%, meaning that one hit was recorded for every 2.5 calls. As illustrated in Figure 3.5 where the inspectors are plotted in order of increasing detection rate from left to right, the number of false calls (negative y-axis) is not correlated with

the number of hits (positive y-axis). Therefore, inspectors that made more calls did not necessarily locate more cracks.

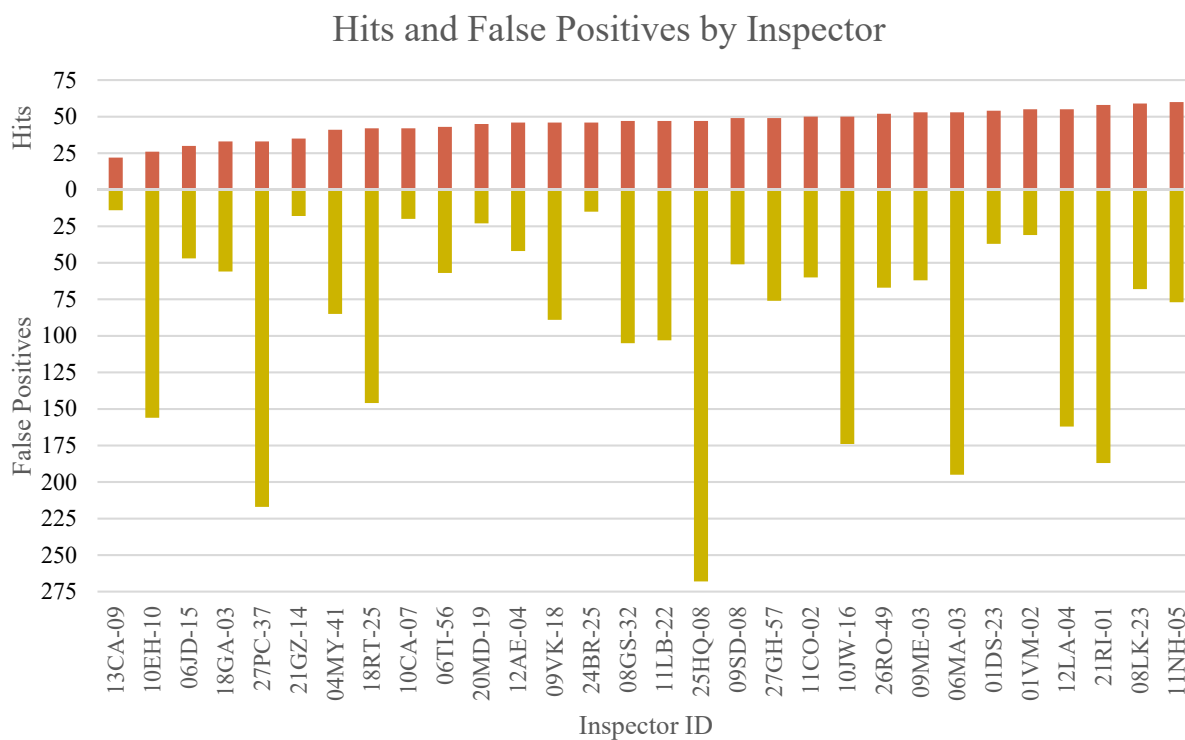


Figure 3.5 Hits and false calls by inspector during the hands-on inspections

The detection rates and number of false calls by specimen and crack type were also examined. The course included 28 out-of-plane cracks in the girder specimens, 23 weld toe cracks in the welded cover plate specimens, and 19 rivet hole cracks in the riveted plate specimens. The weld toe cracks had the highest average detection rate while the average number of false calls was the highest on the girder specimens. The detection rate by crack type is shown for each inspector in Figure 3.6.

The average detection rates for each specimen type for each inspector were also studied. Twenty-three (23) of the 30 inspectors recorded their highest detection rate on the weld toe cracks. Four inspectors recorded their highest detection rate on the rivet hole cracks and three inspectors recorded their highest detection rate on the out-of-plane cracks.

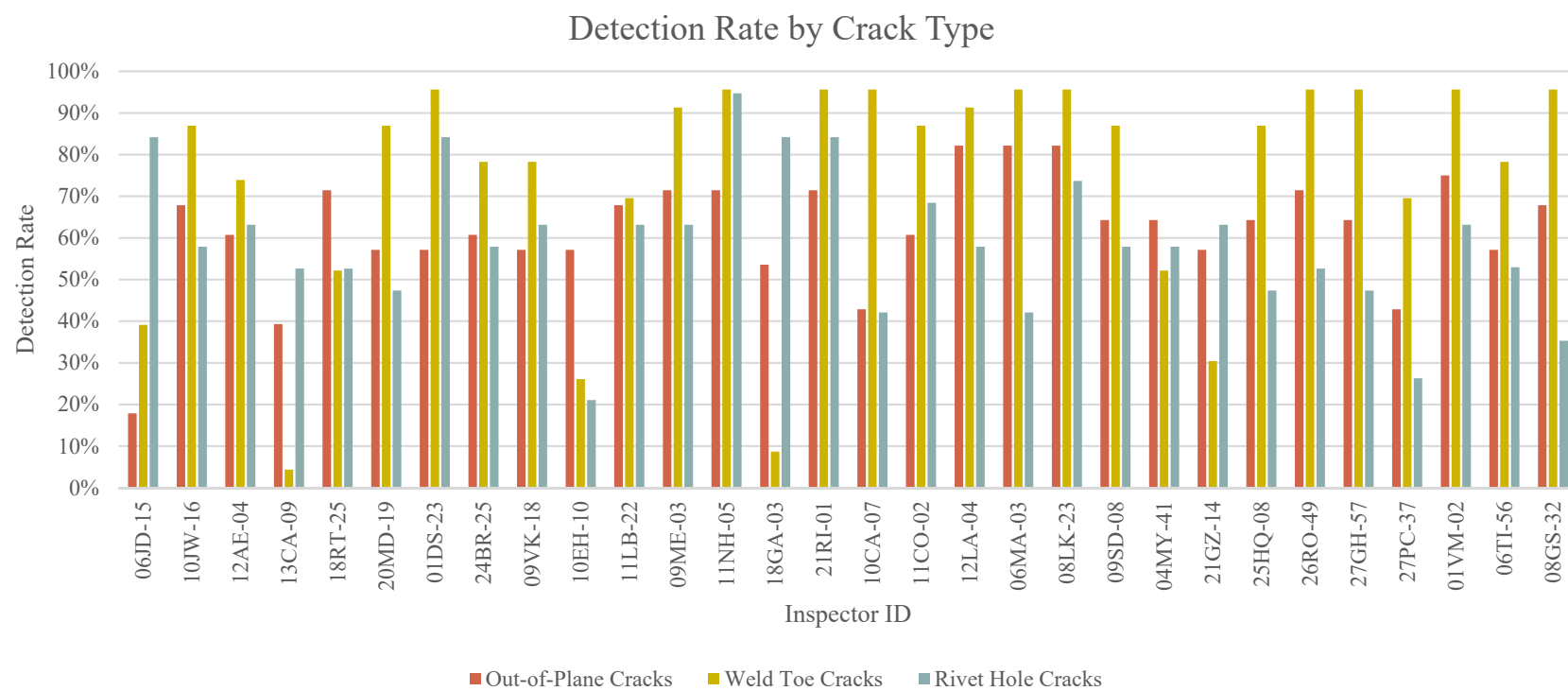


Figure 3.6 Inspector detection rates by crack type

Finally, there was a positive correlation between detection rates of the out-of-plane cracks and the weld toe cracks, meaning that performance on one subset of the inspection could be used to predict performance on another subset. However, there was no correlation between detection rates of the out-of-plane cracks and the rivet hole cracks or between the detection rates of the rivet hole cracks and the weld toe cracks. The relationship between weld toe crack detection rate and out-of-plane crack detection rate is shown in Figure 3.7.

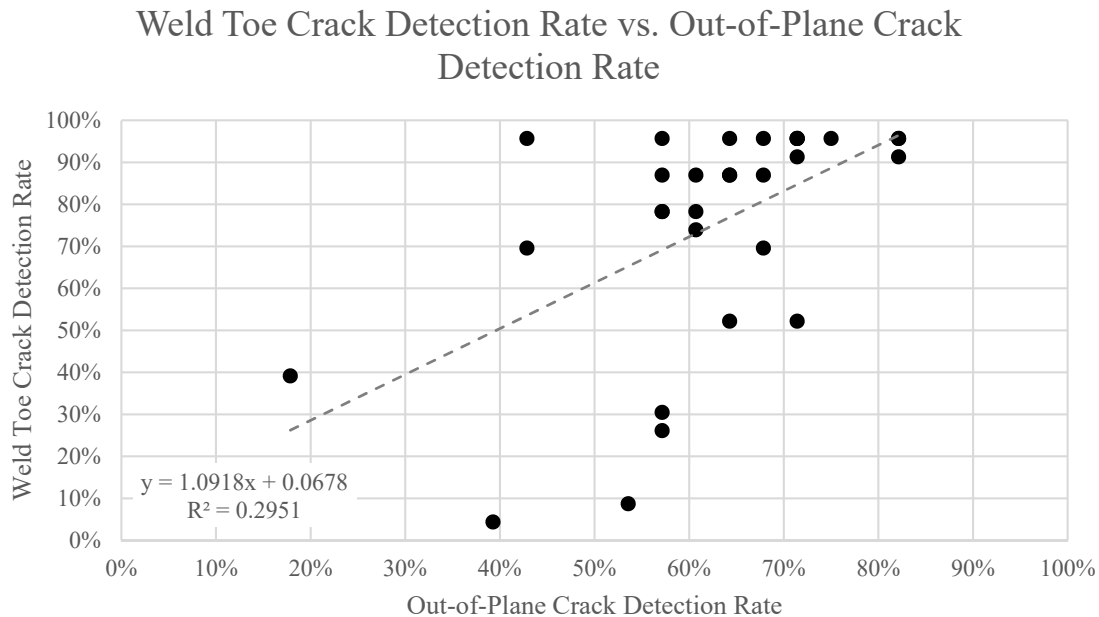


Figure 3.7 Weld toe crack detection rate plotted against out-of-plane crack detection rate

One of many challenges in conducting a realistic POD study is ongoing maintenance of the specimens [17]. Since the specimens in this study are not regularly exposed to cyclic loading from vehicular traffic, the appearance of the cracks may degrade over time. Rust staining and debris accumulation may gradually obscure some of the cracks. Figure 3.8 shows the detection rate for each inspector versus the number of days since the first inspection. The total length of time between the first and last inspection was 762 days. Although the visibility of some of the cracks may have changed over time, the overall difficulty of the inspection seems to have remained relatively constant. Large differences in performance were recorded in inspections that occurred on consecutive days, indicating that the variability in results is likely due more to inspector characteristics than changes in the conditions of the of the specimens. Additionally, following the

long gap between the 11th and 12th inspections, the conditions of all the specimens were revalidated by Purdue research staff.

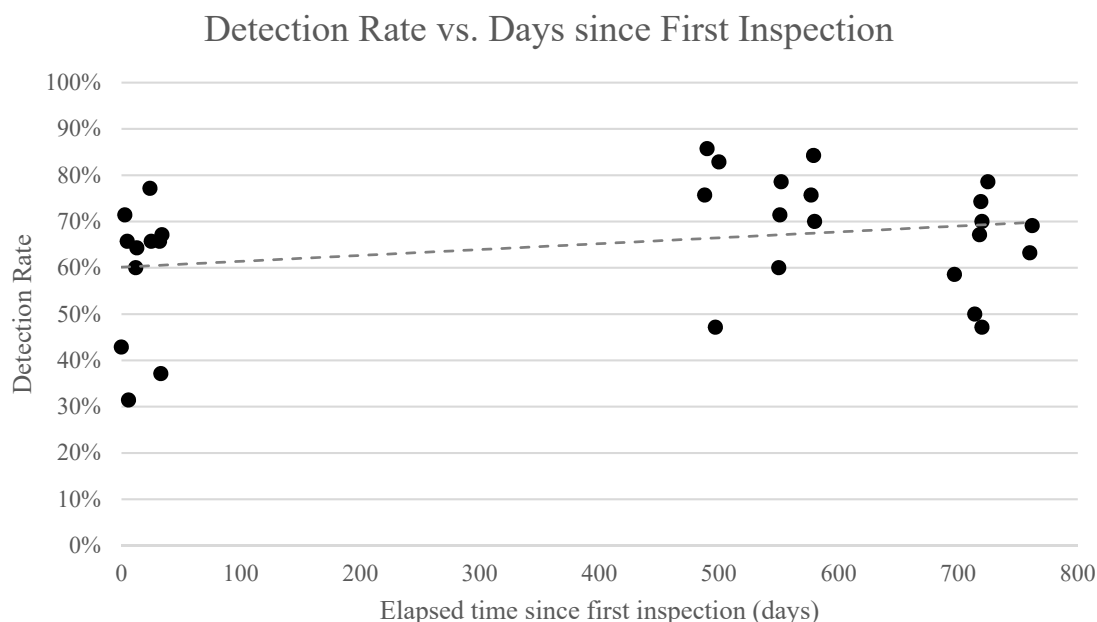


Figure 3.8. Detection rate plotted against the number of days since the first POD inspection

3.5.2 Crack Sizing

In addition to recording the location of detected cracks, inspectors were asked to record the length of the cracks on their inspection forms. Some inspectors meticulously measured each crack, others eyeballed the length, and most used a combination of the methods. Due to concerns about time, inclement weather, or possibly fatigue, many inspectors gradually stopped recording crack lengths as their inspections progressed. Although this study was focused more on crack detection than crack sizing, the relationship between measured crack length and actual crack length was investigated.

Figure 3.9 shows the crack length data for the out-of-plane cracks in the girder specimens. The actual length of the crack is shown on the horizontal axis and the vertical axis displays the measured length reported by the inspector. The diagonal 1:1 reference line represents exact agreement between the actual length and the measured length. For the majority of the cracks in the girder specimens, the average of the measured lengths plots below the 1:1 line indicating that

the inspectors tended to underestimate the crack length. Since most of these cracks were wrapped around the ends of the stiffeners, it was difficult to accurately measure the length and most inspectors just approximated the length of each leg. The average measurement error was -0.38 inch and the average absolute error was 0.92 inch. The average absolute error increased with crack length and the percent absolute error remained constant with crack length. Table 3.4 shows the error analysis for the reported length measurement for all the cracks in the girder specimens.

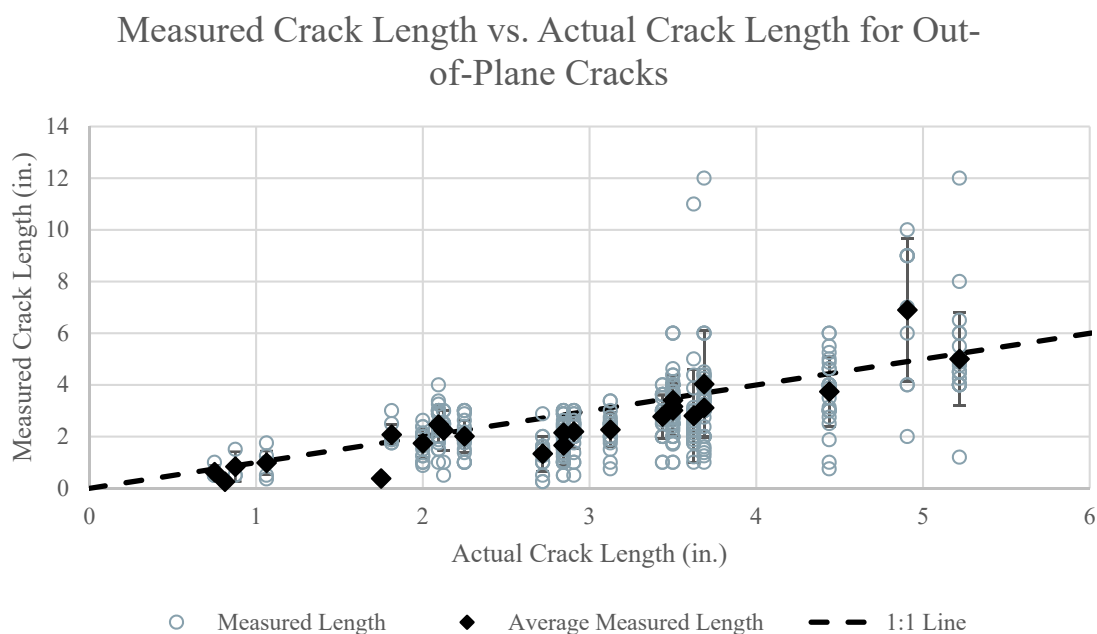


Figure 3.9 Measured crack length versus actual crack length for out-of-plane cracks

Figure 3.10 and Figure 3.11 present the crack length data for the welded cover plate and riveted plate specimens, respectively. In contrast to the girder specimens, the average of the measured lengths of these cracks is generally above the 1:1 line indicating that the inspectors had a tendency to overestimate the length. For the weld toe cracks, the average measurement effort was 0.51 inch and the average absolute error was 0.76 inch. For the rivet hole cracks, the average measurement error was 0.14 inch and the average absolute error was 0.24 inch. For both crack types, the average absolute error increased with crack length and the percent absolute error decreased with crack length. Table 3.4 shows the error analysis for the reported length measurement for all the cracks in the welded cover plate and riveted plate specimens.

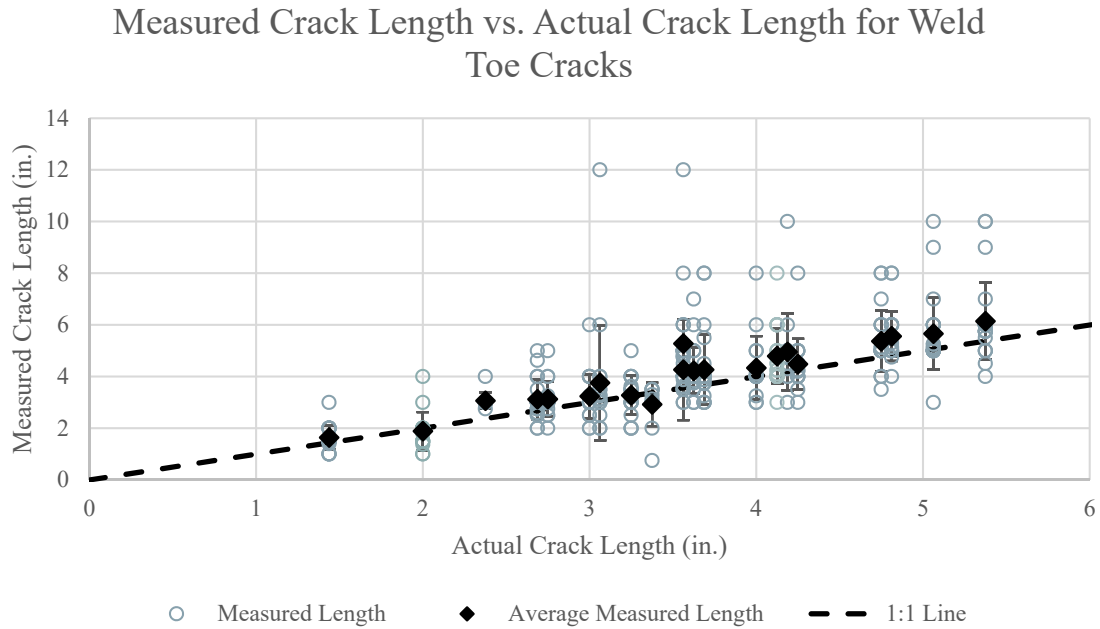


Figure 3.10 Measured crack length versus actual crack length for weld toe cracks

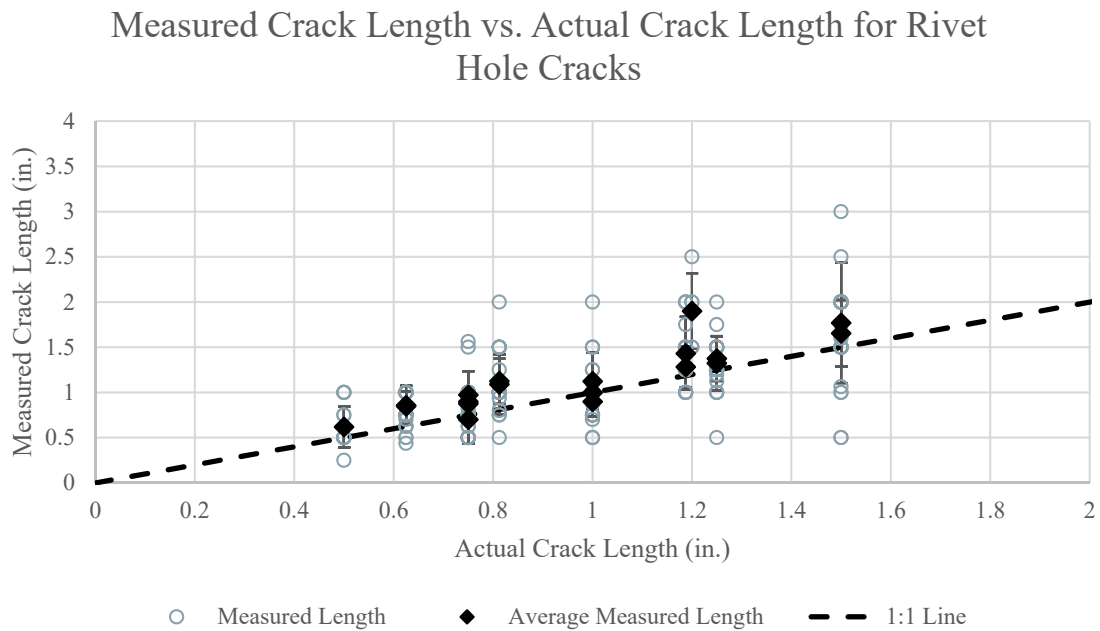


Figure 3.11 Measured crack length versus actual crack length for rivet hole cracks

Table 3.4 Error analysis for crack length measurements from hands-on inspections

Statistic	Out-of-Plane Cracks (28 cracks)		Weld Toe Cracks (23 cracks)		Rivet Hole Cracks (19 cracks)	
	Error (actual length – meas. length)	Absolute Error	Error (actual length – meas. length)	Absolute Error	Error (actual length – meas. length)	Absolute Error
Min./Max. (in.)	-4.02/8.31	-	-2.63/8.94	-	-1/1.5	-
Average (in.)	-0.38	0.92	0.51	0.76	0.14	0.24
St. Dev. (in.)	1.26	0.94	1.20	1.06	0.32	0.25
Average %	-13	29	14	22	16	26

3.5.3 Other Defects

Within the specimen inventory there were two non-crack defects – a “Hoan-like” detail with insufficient web gap between the gusset plate and the transverse stiffener and a missing rivet on a vertically mounted riveted plate. This study was not intended to evaluate an inspector’s ability to locate these two defects, but many inspectors did note their presence. Nine inspectors (30%) recorded the “Hoan-like” detail on the inspection form or verbally mentioned it to the proctor. These inspectors typically indicated that it was a problematic detail and difficult to inspect. The inspectors who located this defect did not perform significantly better than inspectors who did not locate this defect. Only two inspectors (6%) noted the missing rivet on their inspection forms, although more inspectors may have noticed this defect and not recorded it. The inspectors with the highest detection rates (11NH-05 and 08LK-23) were the two inspectors who documented the missing rivet. Notably, this is about the same detection rate as was observed during the FHWA study of bridge inspection [5].

3.6 Crack Length Analysis

Absent from the hit/miss results presented previously is information about the length of the cracks that can be reliably detected. Two different approaches were taken to investigate the relationship between crack length and crack detection. First, a general sense of the relationship was obtained by dividing the number of cracks detected by the total number of cracks inspected within discrete crack length ranges. Then, a probabilistic approach was used to determine the probability of detection as a function of crack length.

3.6.1 Crack Detection by Crack Length

Each crack had the potential to be detected 30 times, once by each participant. One crack was detected by all 30 inspectors, while three cracks were not detected by any participants. The largest undetected crack was 3-1/4 inches. This undetected crack was located along the weld toe of a tapered cover plate specimen. The other two undetected cracks were out-of-plane cracks in the girder specimens. These cracks were both located in the web gap between a longitudinal stiffener and a transverse stiffener and they were 3/4 inch and 1-11/32 inches in length. The smallest crack, measuring 1/2 inch, was detected by 23 of the 30 inspectors. This crack was located at a rivet hole. Figure 3.12 shows each crack, sorted by length, and the number of times it was correctly identified by an inspector. Visually, it appears that there was a correlation between crack size and number of detections. Generally, the longer cracks were detected more often than the shorter cracks.

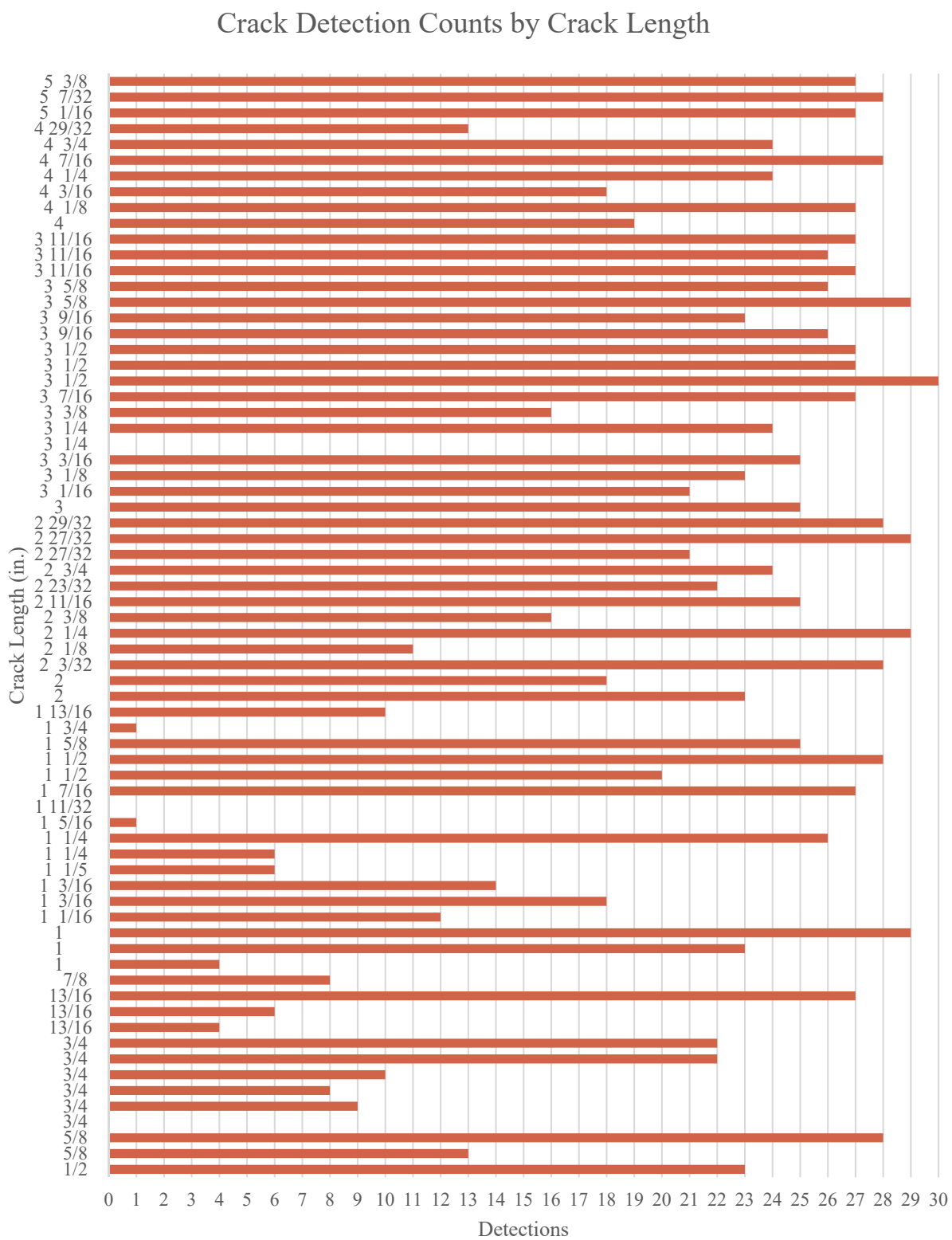


Figure 3.12 Number of crack detections by crack

The cracks were grouped into 1/2-inch length increments, beginning at 1/2 inch, and the detection rate for bin was computed. The detection rate was determined by summing the number of hits for each crack in the increment and dividing by the total number of detection attempts made. The smallest crack size ranges, 1/2 to 1 inch and 1 to 1-1/2 inches, had the lowest detection rate, 46%. The largest crack size range, 5 to 5-1/2 inches, had the highest detection rate, 91%. The detection rates for the crack size ranges in between the extremes varied from 56% to 89%. Figure 3.13 shows the detection rates for each length increment. The number of cracks in each bin are shown on the right hand vertical axis.

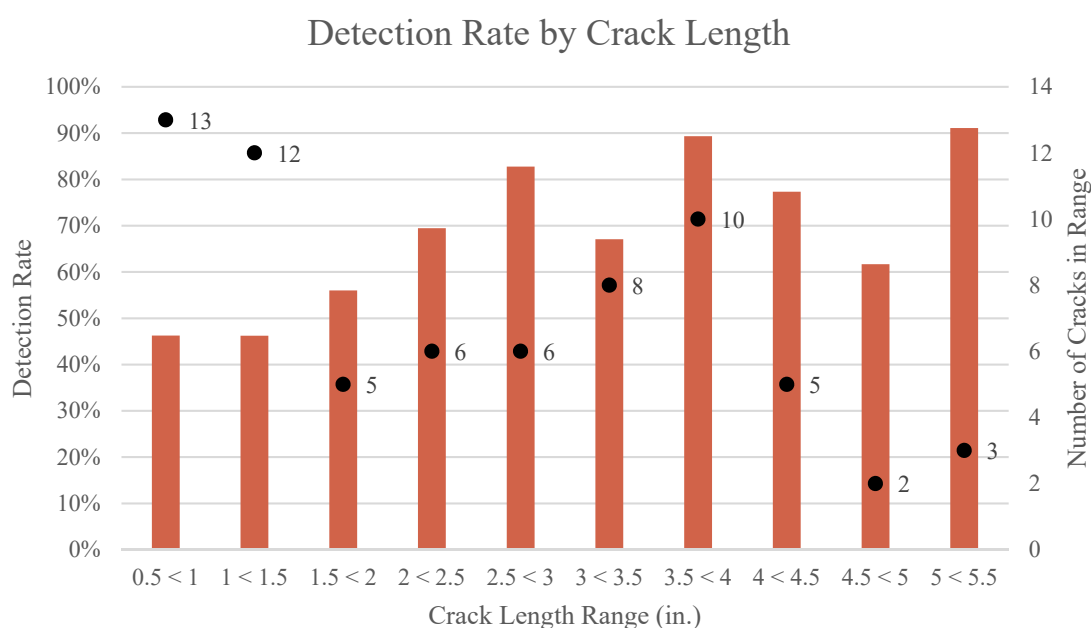


Figure 3.13 Detection rate by crack length

The detection results for each crack type were separated from the total data and examined for a relationship between crack length and detection. As shown in Figure 3.14, the detection rate increased with increasing crack length for the out-of-plane cracks. Although the increase is not constant, there is a clear improvement in detection for cracks longer than 2 inches. Figure 3.9 and Figure 3.10 show the relationship between detection rate and crack length for the weld toe cracks and the rivet hole cracks, respectively. Detection rate does not appear to be heavily influenced by crack length for these crack types, although the more limited range in crack lengths may have restricted the relationship.

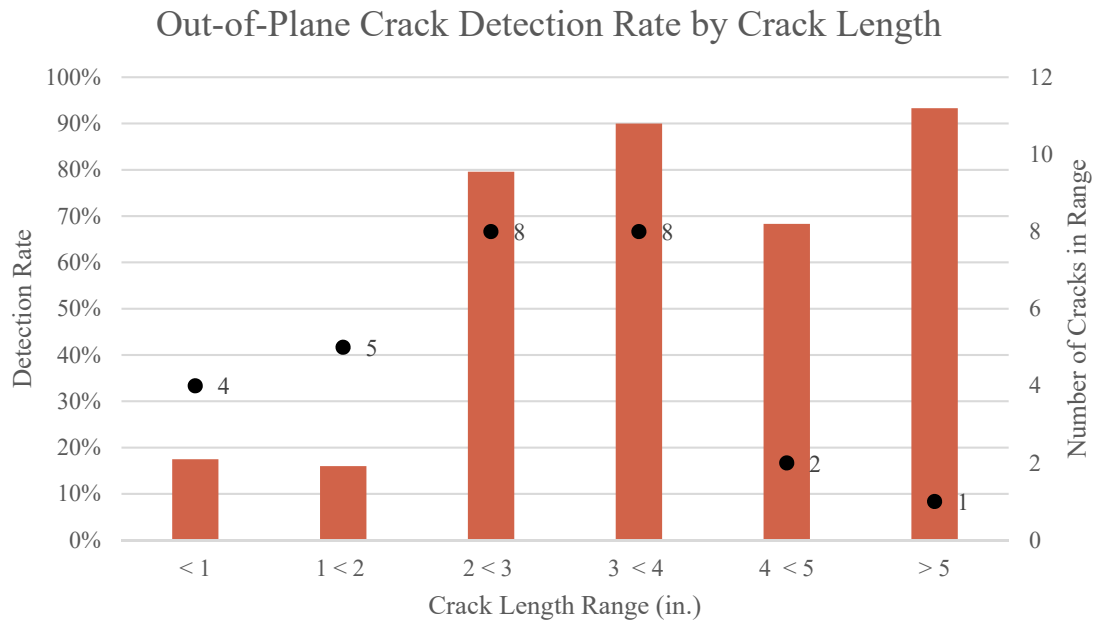


Figure 3.14 Out-of-plane crack detection rate by crack length

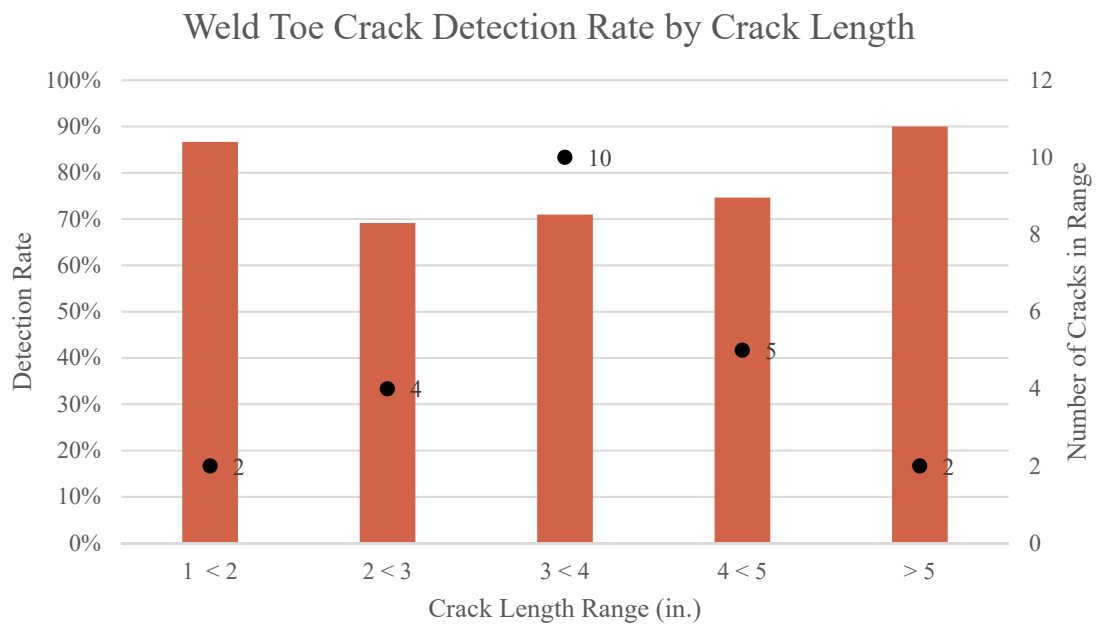


Figure 3.15 Weld toe crack detection rate by crack length

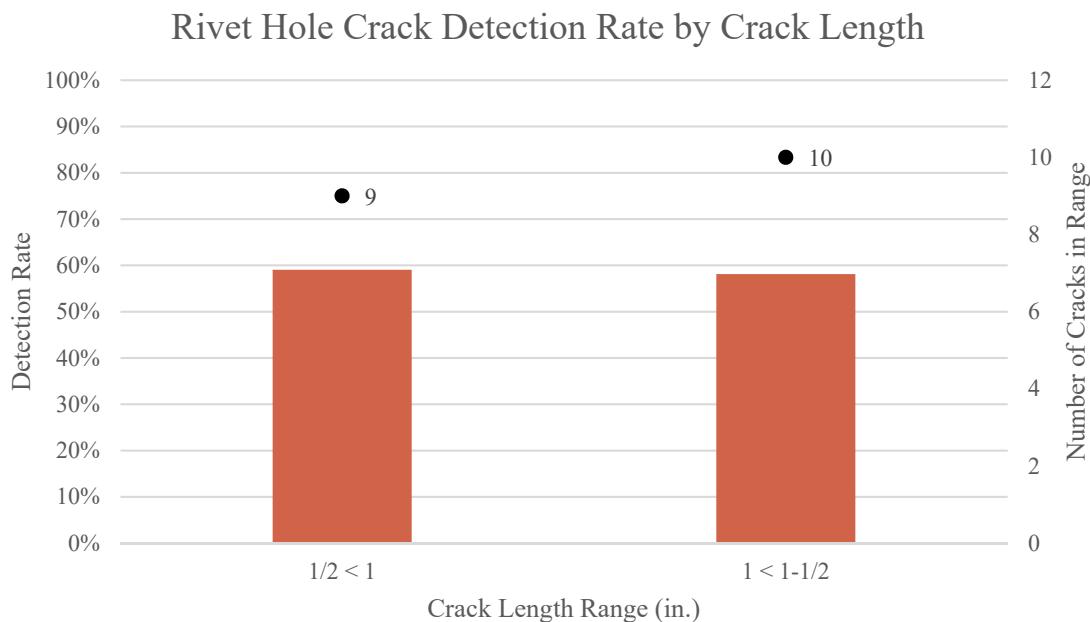


Figure 3.16 Rivet hole crack detection rate by crack length

3.6.2 Probability of Detection Curves

The hit/miss results and crack size information were used to generate probability of detection (POD) versus crack length (a) curves in accordance with Military Handbook 1823a, Nondestructive Evaluation (NDE) System Reliability Assessment [52]. The accompanying software, *mh1823POD*, was used to complete the statistical analysis [19]. POD curves were developed for the full specimen inventory and for each specimen type separately. The best fit was provided by the logistic function and the length response was found to be Cartesian, not logarithmic, meaning that POD was correlated with a , not $\log(a)$. It must be noted that the POD curves developed for the individual specimen types violate two of the recommended guidelines for producing valid POD curves due to the reduced number of defects included in each [17], [57]. First, the recommended minimum number of cracks for hit/miss modelling is 60 and second, the crack length range should provide coverage from $\text{POD} = 3\%$ to $\text{POD} = 97\%$. While the precision of the $\text{POD}(a)$ function may be reduced by these deficiencies, the variability and shapes of the POD curves still offer useful insight into inspector performance. As such, these curves are presented below but with limited discussion specific to the estimated detection rate crack lengths.

The POD versus crack length plot for the full specimen inventory is shown in Figure 3.17 and the 50% and 90% detection rate crack lengths for each inspector are shown in Table 3.5. The hits are plotted at $POD = 1$ and the misses are plotted at $POD = 0$. In total, 2100 observations (30 inspectors x 70 cracks) were considered. Individual POD curves were developed for each inspector, shown in gray in the figure, and a single curve was generated based on the results from all 30 inspectors, shown in black. For all inspectors and all specimens, the average 50% detection rate crack length (a_{50}) was 1 inch and the average 90% detection rate crack length (a_{90}) was 5-1/2 inches. Due to the variability in inspector performance, the 95% confidence limits could not be applied to a_{90} . Inspectors 13CA-09, 06JD-15, and 18GA-03 showed a negative relationship between probability of detection and crack length, meaning they were more likely to find shorter cracks than longer cracks. It is unlikely that these inspectors are actually better equipped to find shorter cracks, but instead environmental or human factors contributed to them detecting shorter cracks or missing longer cracks. Excluding these three inspectors, the shortest 90% detection rate crack length was approximately 2-3/8 inches and the longest 90% detection rate crack length was approximately 22-1/2 inches.

Table 3.5 Probability of detection crack sizes considering all cracks

Inspector ID	50% Crack Length (in.)	90% Crack Length (in.)	90% Crack Length w/ 95% confidence (in.)	Inspector ID	50% Crack Length (in.)	90% Crack Length (in.)	90% Crack Length w/ 95% confidence (in.)
10EH-10	3.52	7.67	N/A	06MA-03	0.96	2.81	4.19
11CO-02	0.94	3.44	5.63	21GZ-14	2.42	22.46	N/A
18GA-03 ¹	2.16	-2.07	N/A	09VK-18	0.95	5.43	N/A
13CA-09 ¹	-1.09	-10.75	N/A	18RT-25	0.58	10.31	N/A
09SD-08	1.03	3.57	5.85	06JD-15 ¹	0.32	-15.62	N/A
20MD-19	1.56	3.77	N/A	06TI-56	1.43	4.99	N/A
04MY-41	1.02	9.66	N/A	27GH-57	1.30	3.10	4.42
09ME-03	1.00	2.71	3.96	01DS-23	-0.39	4.27	N/A
08GS-32	1.43	3.02	4.16	24BR-25	1.17	4.78	N/A
01VM-02	0.78	2.63	4.01	27PC-37	2.56	5.11	N/A
12AE-04	0.99	5.30	N/A	12LA-04	0.44	3.10	5.62
11LB-22	0.62	5.73	N/A	11NH-05	-0.11	2.36	4.53
10JW-16	0.76	4.50	N/A	10CA-07	1.82	4.80	N/A
21RI-01	-0.10	2.84	N/A	26RO-49	1.15	2.54	3.56
08LK-23	0.17	2.41	4.23	25HQ-08	1.26	4.04	N/A

¹ Probability of detection decreasing with increasing crack length

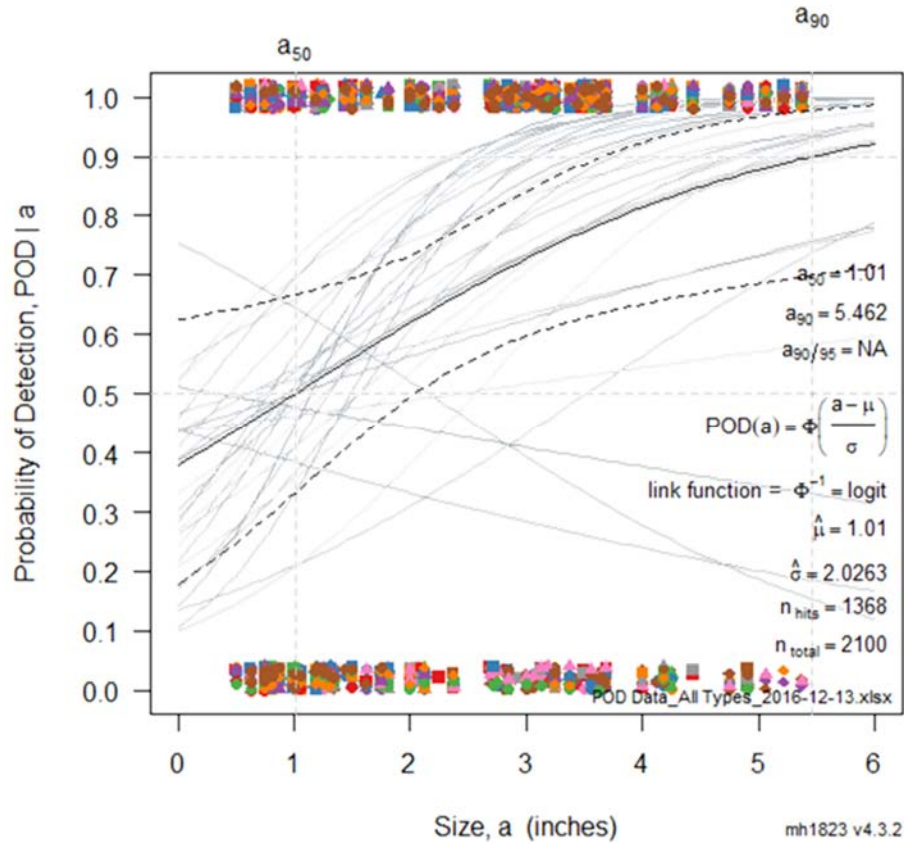


Figure 3.17 Probability of detection curves considering all cracks

After removing the results from the three inspectors whose performance did not conform to the assumptions of the POD model (POD increasing with increasing crack length), the average probability of detection curve was regenerated and is shown in Figure 3.18. For the 27 remaining inspectors and all the specimens, a_{50} remained 1 inch, but a_{90} was reduced to 4-1/2 inches. Due to variability, the 95% confidence limits could still not be applied to a_{90} . The average POD curve for all specimens does not appear asymptotic with $\text{POD} = 1.0$, even for large crack sizes. This indicates that there is a residual probability of missing cracks, even when the length is very long. Additionally, the relatively shallow slope of the curve as it transitions from $\text{POD} = 0$ to $\text{POD} = 1.0$ is due to the large variability in performance and suggests that variables other than crack length are affecting the probability of detection.

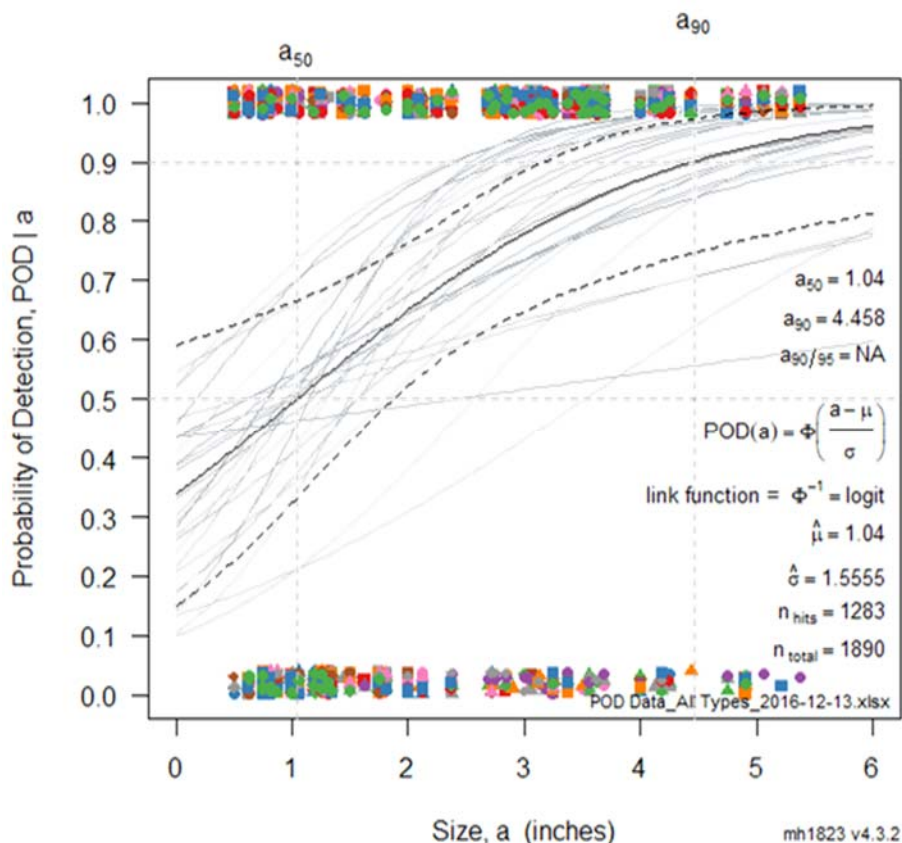


Figure 3.18 Probability of detection curves considering all cracks, excluding inspectors 13CA-09, 06JD-15, and 18GA-03

An average POD curve, as created for this sample of inspectors, may disguise the strengths and weaknesses of each inspector and is likely not a reliable predictor of the performance of a future inspector, who could be a “good” or a “bad” inspector. Instead, the results for each inspector should be considered independently to identify effective inspection techniques, promising inspector characteristics, and favorable environmental conditions. The longest 50% detection rate crack length, 3-1/2 inches, was calculated for an inspector with 30 years of inspection experience that had performed no hands-on inspections in the previous 12 months. The average temperature on the day of the inspection was 33°F and the proctor recorded that inspector was “very aware he was being tested” and explained that he was “only there because he was told to be there”. Conversely, the shortest 90% detection rate crack length, 2-3/8 inches, was achieved by an inspector with 10 years of inspection experience that had performed 10 to 15 hands-on inspections in the previous 12 months. The average temperature on the day of the inspection was 47°F and the

proctor noted that the inspector was “methodical and friendly” and “made a few comments about the length of the course”.

The probability of detection at 1-inch increments was calculated for each inspector. Table 3.6 shows the average probability of detection at each length increment along with the minimum and maximum probabilities of detection for the 27 inspectors whose performance can be described as a function of crack size.

Table 3.6 Probability of detection by crack length (all cracks)

Crack Size (in.)	Average	Standard Deviation	Minimum	Maximum
0.5	40%	13%	14%	63%
1	49%	13%	21%	73%
2	66%	14%	31%	87%
3	79%	14%	43%	95%
4	87%	13%	54%	99%
5-3/8	92%	10%	58%	99.8%

The POD versus crack length plot for the out-of-place cracks in the girder specimens is shown in Figure 3.19 and the 50% and 90% detection rate crack lengths for each inspector for the girder specimens are shown in Table 3.7. For this subset, 840 observations (30 inspectors x 28 cracks) were considered. For all inspectors and the girder specimens, the average 50% detection rate crack length was 2 inches and the average 90% detection rate crack length was 4 inches. Due to the variability in inspector performance, the 95% confidence limits could not be applied to a_{90} . The shape of the average POD curve for the girder specimens indicates that both the residual probability of a miss at large crack sizes and the variability in performance is reduced as compared to the full specimen inventory. Although this model considers only half of the recommended number of cracks, the S-shape of the curve and slenderness of the confidence bounds indicate that the model is able to predict the POD(a) function with reasonable accuracy. The shortest 90% detection rate crack length was approximately 1-3/4 inches and the longest 90% detection rate crack length was approximately 14-3/4 inches.

Table 3.7 Probability of detection crack sizes considering only the out-of-plane cracks

Inspector ID	50% Crack Length (in.)	90% Crack Length (in.)	90% Crack Length w/ 95% confidence (in.)	Inspector ID	50% Crack Length (in.)	90% Crack Length (in.)	90% Crack Length w/ 95% confidence (in.)
10EH-10	2.27	3.68	5.76	06MA-03	0.72	2.75	N/A
11CO-02	2.16	2.78	3.69	21GZ-14	2.25	3.88	N/A
18GA-03	1.74	14.81	N/A	09VK-18	2.12	5.18	N/A
13CA-09	3.42	7.10	N/A	18RT-25	1.06	4.24	N/A
09SD-08	2.00	2.73	3.72	06JD-15	4.69	7.10	N/A
20MD-19	2.27	3.68	5.58	06TI-56	2.20	4.40	N/A
04MY-41	1.91	3.50	5.95	27GH-57	1.98	3.03	4.37
09ME-03	1.78	1.79	1.86	01DS-23	2.21	4.30	N/A
08GS-32	1.78	3.13	4.98	24BR-25	2.08	3.71	N/A
01VM-02	1.48	2.53	3.93	27PC-37	2.93	4.68	N/A
12AE-04	2.12	3.38	5.09	12LA-04	0.85	2.62	N/A
11LB-22	1.78	3.13	4.98	11NH-05	1.70	2.30	3.22
10JW-16	1.34	4.50	N/A	10CA-07	3.08	6.29	N/A
21RI-01	1.59	3.01	5.02	26RO-49	1.78	1.79	1.86
08LK-23	1.12	2.19	3.75	25HQ-08	1.90	3.58	N/A

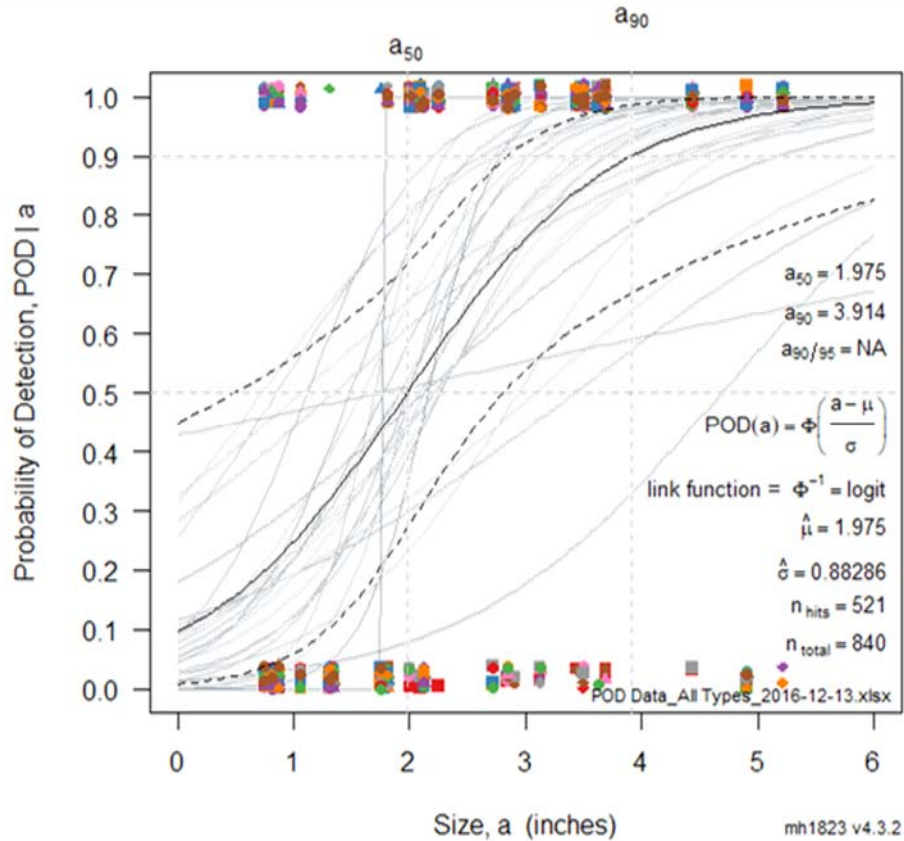


Figure 3.19 Probability of detection curves considering only the out-of-plane cracks

The POD versus crack length plot for the weld toe cracks in the welded cover plate specimens is shown in Figure 3.20. For this subset, 690 observations (30 inspectors x 23 cracks) were considered. According to this model, even the shortest weld toe crack, 1-7/16 inches long, had a 50% probability of being detected, but for a 90% chance of being detected, the crack would need to be over 10 inches long. The range of crack lengths and the number of cracks included in the welded cover plate specimens is likely not wide enough to accurately predict the true POD curve. This is supported by the bar graph presented in Section 3.6.1 which also showed a weak relationship between detection rate and crack length within the range of crack lengths considered in this study. The width of the confidence bounds reflects the variability in results and indicates the imprecision in the average POD curve. Although the exact a_{90} and a_{50} values may not be accurate, the individual curves show that a few inspectors really struggled with these specimens while other inspectors were able to identify nearly every crack in the inventory. Ten of the inspectors detected 22 out of the 23 possible cover plate weld toe cracks and therefore have a high

probability of detection for all lengths of cracks of this type. Inspectors 13CA-09, 18RT-25, 04MY-41, and 25HQ-08 showed a negative relationship between probability of detection and crack length. This was likely due to flawed decision making rather than failure to locate the crack. Based on previous experience, some of the inspectors believed the weld toe cracks to be cracks in the paint, not fatigue cracks in the base metal. This belief was either mentioned to the proctor or noted on the inspection form. In these cases, personal biases or expectations may have had more influence on performance than defect size.

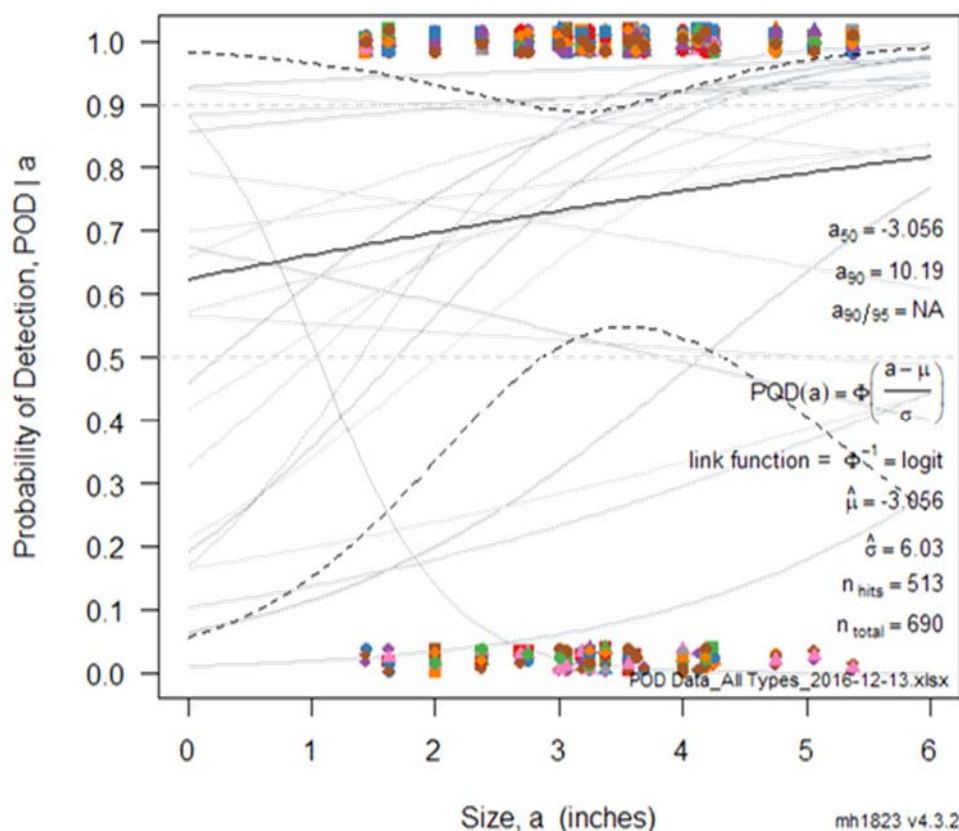


Figure 3.20 Probability of detection curves considering only the weld toe cracks

The POD versus crack length plot for the rivet hole cracks in the riveted plate specimens is shown in Figure 3.21. For this subset, 570 observations (30 inspectors x 19 cracks) were considered. According to this model, even the shortest rivet hole crack, 1/2 inch long, had a 50% probability of being detected, but for a 90% chance of being detected, the crack would need to be over 84 inches long. As shown in the detection rate versus crack length bar graph presented in Section 3.6.1, the range of crack lengths was not wide enough to develop the full POD curve. Within the

1-inch range provided, the probability of detection is nearly constant at approximately $POD = 0.6$. Seventeen of the inspectors had a negative relationship between probability of detection and crack length for the riveted hole cracks. The variability in performance over a very small range of crack lengths suggests that performance is being influenced by factors beyond crack length.

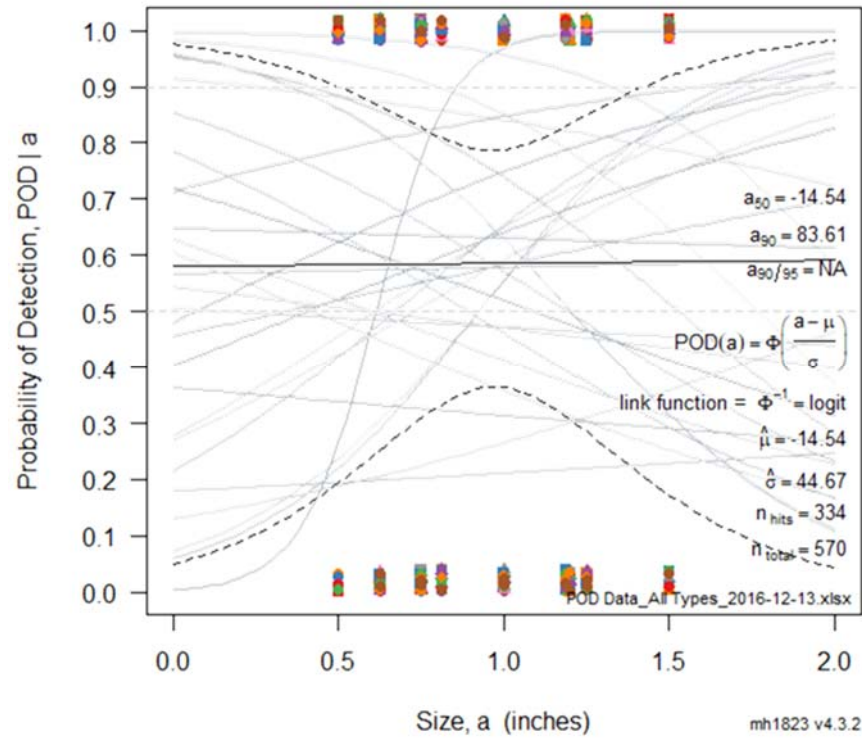


Figure 3.21 Probability of detection curves considering only the rivet hole cracks

3.7 Human and Environmental Factors

A series of complementary statistical models was applied to the inspection data in order to identify key factors that influenced inspection performance. Performance was described by detection rate and the number of false calls. Environmental conditions, specimen characteristics, and inspector attributes were considered in the statistical models. Descriptive statistics for the independent and dependent variables are shown in Table 3.8.

Table 3.8 Descriptive statistics for dependent and independent variables

Variable Description	Min./Max. Values	Mean of Observations	Standard Deviation of Observations
Performance Measures (dependent variables)			
Detection rate (all specimens)	0.31/0.86	0.65	0.14
False calls (all specimens)	14/268	90	67
Specimen Characteristics (independent variables)			
Crack length	0.5/5.375	2.42	1.37
Out-of-plane crack (1 if the crack is an out-of-plane crack, 0 otherwise)	-	0.4	1
Weld toe crack (1 if crack is a weld toe crack, 0 otherwise)	-	0.329	1
Rivet hole crack (1 if crack is at a rivet hole, 0 otherwise)	-	0.271	1
Environmental Conditions (independent variables)			
Average temperature on the day of the inspection (°F)	23/73	45.7	16.1
Average wind speed on the day of the inspection (mph)	4/16	8.53	2.52
Maximum wind speed on the day of the inspection (mph)	12/29	18.1	4.53
Flashlight (1 if used, 0 otherwise)	-	0.667	-
Tape measure (1 if used, 0 otherwise)	-	0.6	-
Inspection mirror (1 if used, 0 otherwise)	-	0.067	-
Magnifier (1 if used, 0 otherwise)	-	0.13	-
Inspector Attributes (independent variables)			
Inspection experience (yrs.)	0/30	10.6	9.33
Age	24/63	40.2	11.7
Inspection duration (min.)	116/457	247	75.2
No. of routine inspections performed in the last 12 months	0/700	142	174
No. of hands-on inspections performed in last 12 months	0/40	10.6	10.5
No. of training courses attended (out of 8 listed on exit survey)	0/7	2.93	1.66
Elapsed time since first inspection (days)	0/762	403	304
Professional licensure (1 if licensed PE or SE, 0 otherwise)	-	0.567	-
Employer (1 if employed by a private consultant, 0 otherwise)	-	0.4	-
Gender (1 if male, 0 otherwise)	-	0.9	-
Safety Inspection of In-Service Bridges (1 if inspector had taken the course, 0 otherwise)	-	0.867	-
Fracture Critical Inspection Techniques (1 if inspector had taken the course, 0 otherwise)	-	0.8	-
Underwater Bridge Inspection (1 if inspector had taken the course, 0 otherwise)		0.2	
Introduction to Element Level Bridge Inspection (1 if inspector had taken the course, 0 otherwise)		0.267	

Initially, a simple univariate regression analysis was used to identify statistically significant correlations between inspection performance and the independent variables and the independent two-samples t -test was used to determine which independent variables were significant in discriminating between higher and lower performing inspectors. Then, a multivariate regression analysis was used to predict the influence that multiple variables, considered simultaneously, had on inspection performance. Finally, a binary logit model was used to determine which factors, beyond crack length, affected the likelihood of detection for an individual crack. Unless otherwise noted, statistical significance was determined using a two-tailed test with a 95% confidence level and goodness of fit was evaluated using the adjusted R-squared (R^2) or rho-squared (ρ^2) statistics.

3.7.1 Univariate Regression and Two Samples t -Test

To determine the influence of the individual and environmental factors on the performance measures, a simple linear regression analysis with a single variable was performed. The univariate linear regression model is assumed to take the form

$$y_i = \beta_0 + \beta_1(x_{1i}) + \varepsilon_i \quad (3.1)$$

where the dependent variable y_i is a function of the intercept parameter β_0 , the slope parameter β_1 times the independent variable x_{1i} , and a disturbance term ε_i [58]. The subscript i corresponds to the inspector ($i = 1, \dots, 30$). Ordinary least squares estimation is used to predict the parameter values. To understand the significance of the model, a test statistic (t -statistic) is used to determine the probability that the slope parameter, β_1 , is equal to zero and therefore the independent variable has no effect on the value of the dependent variable. For linear regression, the t -statistic is defined as the estimate of the coefficient of the independent variable divided by the standard error of the estimated coefficient. It is assumed to have a t distribution with degrees of freedom equal to the number of observations minus the number of model parameters (2). The probability value (p -value) expressing the probability that the coefficient is equal to zero can be calculated from the t distribution table. A low p -value indicates that the coefficient has a low likelihood of being equal to zero and therefore, the independent variable is more likely to be a meaningful predictor of the dependent variable. Table 3.9 shows the p -values for all the combinations of performance measures and independent variables. The p -values less than 5% are shown in bold since this was the threshold used to indicate significance in this study.

Table 3.9 p -values for the univariate linear regression analyses

Performance Measure	Independent Variable								
	Insp. Duration	Avg. Temp.	Age	Insp. Exp.	No. of Training Courses	Max. Wind Speed	Avg. Wind Speed	No. of Routine Insp.	No. of Hands-On Insp.
All Specimens									
Detection Rate	0.032	0.047	0.028	0.013	0.018	0.492	0.865	0.361	0.191
False Calls	0.938	0.563	0.714	0.462	0.367	0.180	0.935	0.923	0.196
Girder Specimens									
Detection Rate	0.022	0.073	0.955	0.500	0.021	0.222	0.834	0.345	0.674
False Calls	0.804	0.710	0.615	0.311	0.529	0.082	0.759	0.895	0.261
Welded Cover Plate Specimens									
Detection Rate	0.379	0.058	0.001	0.001	0.052	0.183	0.901	0.802	0.135
False Calls	0.637	0.405	0.353	0.430	0.092	0.942	0.350	0.458	0.239
Riveted Plate Specimens									
Detection Rate	0.004	0.894	0.706	0.845	0.657	0.037	0.263	0.257	0.126
False Calls	0.350	0.464	0.240	0.497	0.671	0.866	0.351	0.278	0.337

To complement the univariate regression analysis, the independent two-samples t -test was used to determine which independent variables were significant in discriminating between higher and lower performing inspectors. It compares the mean and variance of two independent samples to determine the likelihood that they were drawn from the same population. The test assumes normal population distributions and is appropriate for sample sizes less than 25 [58]. For the two samples t -test, the test statistic, t , and degrees of freedom, df , are calculated from Equations 3.2 and 3.3 where n_1 and n_2 are the number of observations (inspectors) within each sample, s_1 and s_2 are the sample variances, and \bar{x}_1 and \bar{x}_2 are the sample means. Again, the test statistic is assumed to have a t distribution. The p -value expressing the probability that the two samples were drawn from the same population can be calculated from the t distribution table. A low p -value indicates that there is a low likelihood that the difference between the sample means is due to chance alone and the samples were likely drawn from different populations.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (3.2)$$

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1} \left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1} \left(\frac{s_2^2}{n_2}\right)^2} \quad (3.3)$$

As shown in Figure 3.22, detection rate increased with increasing temperature, indicating that inspector comfort may influence detection rate. The average daily temperature during the inspections ranged from 23°F to 73°F, with an overall average of 46°F. The *t*-test revealed that inspectors working on days with an average temperature above 51°F ($M = 0.726$, $SD = 0.077$) detected significantly more cracks than inspectors who worked on days with an average temperature below 51°F ($M = 0.610$, $SD = 0.148$), $t(27) = 2.81$, $p = 0.009$. It is likely that an upper bound limit on temperature exists, but it was not noticed in this study because the average daily temperature did not exceed 73°F during any of the inspections.

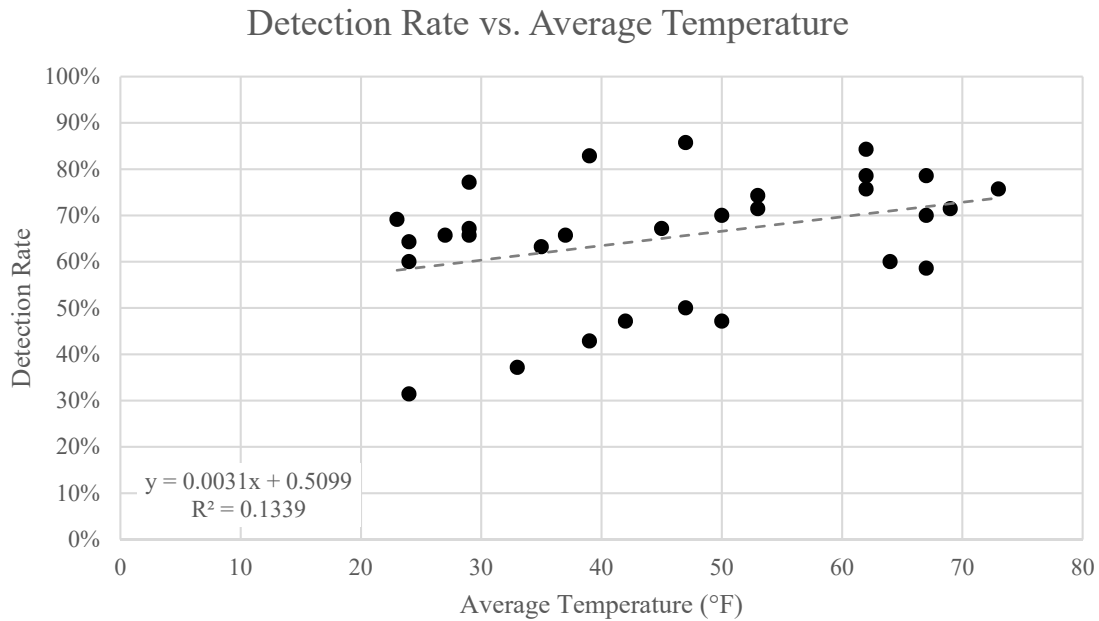


Figure 3.22 Detection rate plotted against average temperature

As shown in Figure 3.23, detection rate increased with increasing duration. This trend was observed in the total detection rate, as well as the detection of out-of-plane and rivet hole cracks. The average time to complete the inspection was 248 minutes, with the fastest inspector completing the inspection in under two hours and the longest inspection lasting over seven hours.

The t -test suggests that inspectors should spend at least 90 seconds inspecting each specimen, or detail, since inspectors who spent more than 3.75 hours completing the inspection ($M = 0.708$, $SD = 0.124$) performed significantly better than inspectors who spent less than 3.75 hours ($M = 0.589$, $SD = 0.127$), $t(27) = 2.60$, $p = 0.015$.

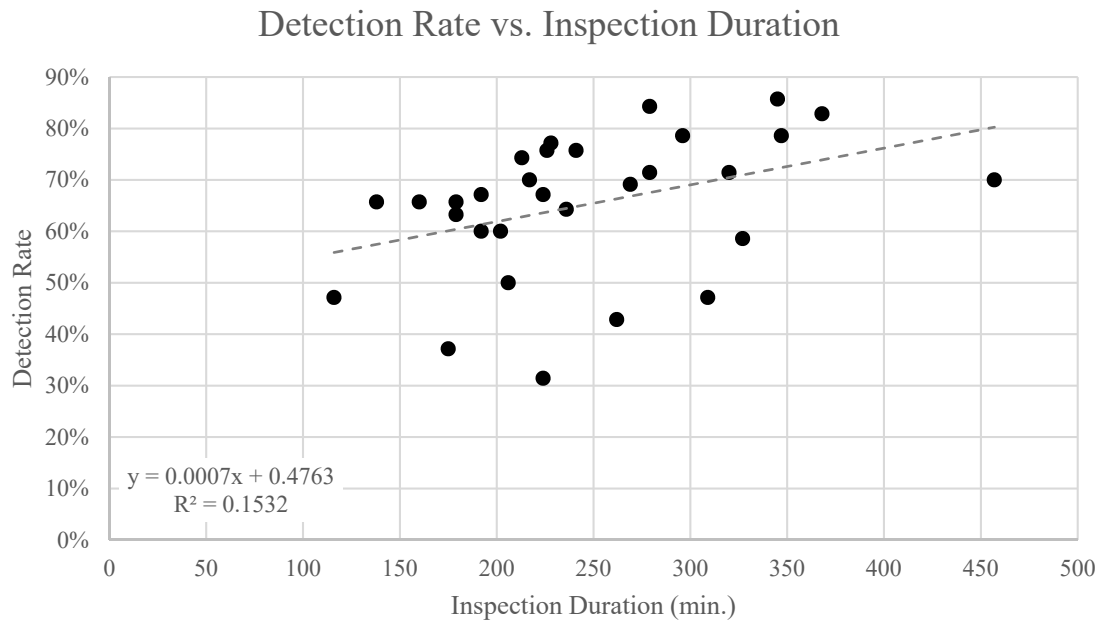


Figure 3.23 Detection rate plotted against inspection duration

Covariance between air temperature and inspection time was investigated to determine whether inspection speed varied significantly with weather conditions. A positive relationship was observed; the inspectors tested on the warmest days tended to spend more time completing the inspection. This suggests that inclement or uncomfortable weather may encourage inspectors to increase inspection speed. The plot of inspection time against air temperature can be found in Figure 3.24.

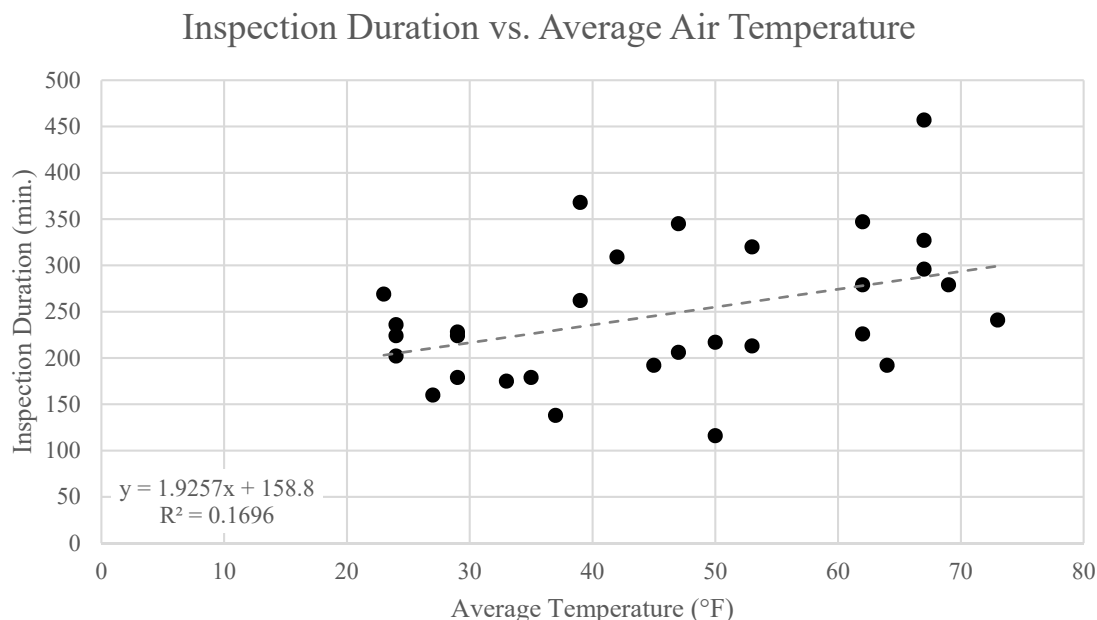


Figure 3.24 Inspection duration plotted against average temperature

On the exit survey, inspectors were asked to list the inspection training that they had completed. To improve the consistency of the results, the exit survey was revised partway through the study to include a list of eight common training courses and asked the inspector to indicate which ones they had completed. The eight courses listed on the revised exit survey were *Safety Inspection of In-Service Bridges*, *Bridge Inspection Refresher Training*, *Engineering Concepts for Bridge Inspectors*, *Underwater Bridge Inspection*, *Fracture Critical Inspection Techniques for Steel Bridges*, *Inspection and Maintenance of Ancillary Highway Structures*, and *Introduction to Element Level Bridge Inspection* from the FHWA/NHI, and *Inspecting Steel Bridges for Fatigue* from the S-BRITE Center/Purdue University. Although some of the inspectors indicated that they completed one or more of these courses multiple times throughout their career, only the first attendance was considered in this study. The average number of courses was three. None of the inspectors had taken all eight classes and three inspectors had not completed any of the training courses. Considering only the eight training courses listed on the exit survey, detection rate increased with increasing attendance at these courses as shown in Figure 3.25. Detection of the out-of-plane cracks was also positively correlated with training. Twenty-three (23) of the 30 inspectors had completed both the FHWA/NHI *Safety Inspection of In-Service Bridges* and *Fracture Critical Inspection Techniques for Steel Bridges* courses prior to their participation in

this study, so the improvement in detection rate is likely due to the completion of more focused or specialized training. In fact, inspectors who had attended at least three training courses ($M = 0.692$, $SD = 0.123$) performed significantly better than inspectors who had attended less than three of the listed training courses ($M = 0.573$, $SD = 0.135$), $t(16) = 2.35$, $p = 0.032$. In addition to the in-depth discussion provided in these courses, they may offer a useful review of the basic principles of inspection. By federal law, inspectors are required to complete refresher training periodically [1, §650.313]. The majority of the inspectors participating in the study had completed one of these trainings within the previous three years, so time since training was not helpful in explaining the variability in detection rate.

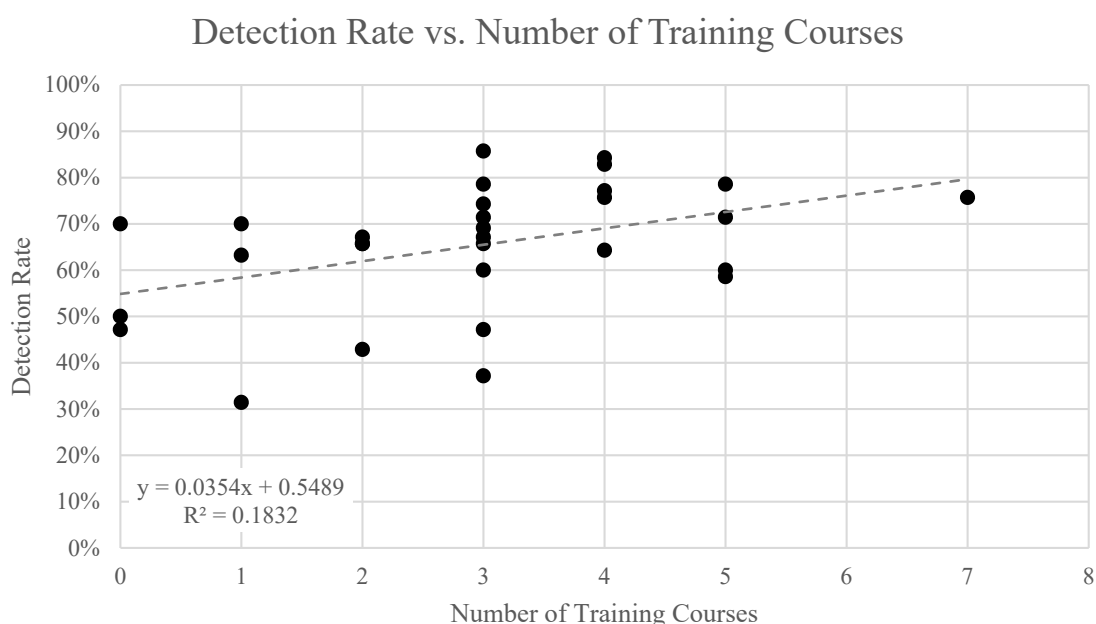


Figure 3.25 Detection rate plotted against the number of training courses taken by the inspector

As shown in Figure 3.26, the relationship between experience and detection rate was negative indicating that inspectors with more experience actually performed worse on this inspection. The average experience of the participants was 10.6 years, with a minimum of 0 years and a maximum of 30 years. There was significant variability in performance, even among inspectors with the same length of experience. For instance, there was a 48% difference in detection rate between two inspectors that both had 20 years of inspection experience. The negative relationship was most pronounced in the welded cover plate specimens as shown in Figure 3.27.

Although common sense suggests that inspection performance should increase with experience, previous research presents a varied relationship between performance and experience [29]–[31]. It is possible that unmeasured factors such as physical health, expectations, or motivation are neutralizing the influence of inspection experience. For instance, an inspector with 29 years of experience explained that even though welded cover plate terminations are fatigue prone details, they rarely find fatigue cracks in these locations and instead supposed that the cracklike indications were paint flaws. As a result, this inspector detected only 9% of the fatigue cracks in these specimens. Another similarly experienced inspector made their displeasure in participating known to the proctor, spent less than 3 hours on the inspection, and only detected 37% of the cracks. The *t*-test suggests that inspectors with between 2 years and 14 years of experience ($M = 0.706$, $SD = 0.106$) performed significantly better than inspectors with less than 2 years or more than 14 years of experience ($M = 0.60$, $SD = 0.147$), $t(25) = 2.26$, $p = 0.033$. This group of inspectors may have enough experience to identify likely crack locations and recognize fatigue cracks, but not so much experience that they are overly influenced by attitude or personal biases.

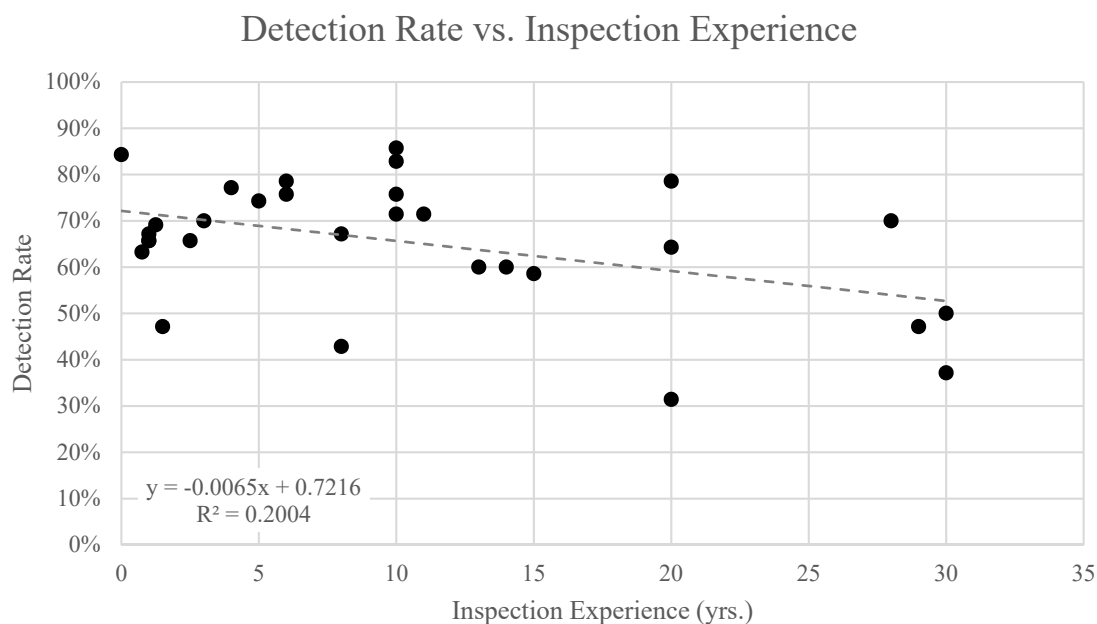


Figure 3.26 Detection rate plotted against inspection experience

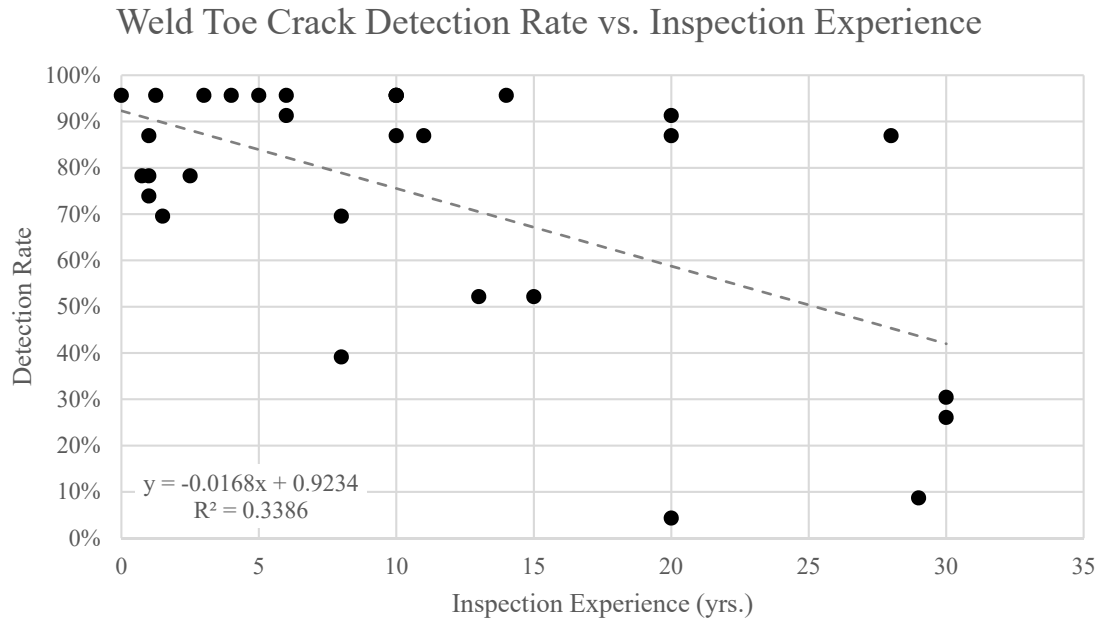


Figure 3.27 Weld toe crack detection rate plotted against inspection experience

Covariance between inspection duration and experience was also investigated to determine if more experienced inspectors inspect more quickly or more slowly than less experienced inspectors. A positive relationship between inspection experience and inspection duration was observed; more experienced inspectors tended to spend more time completing the inspection. Notably, the inspectors with more than 25 years of experience demonstrated the greatest variability in inspection duration with inspection times ranging from 175 minutes to 457 minutes. The plot of inspection experience against inspection duration can be found in Figure 3.28.

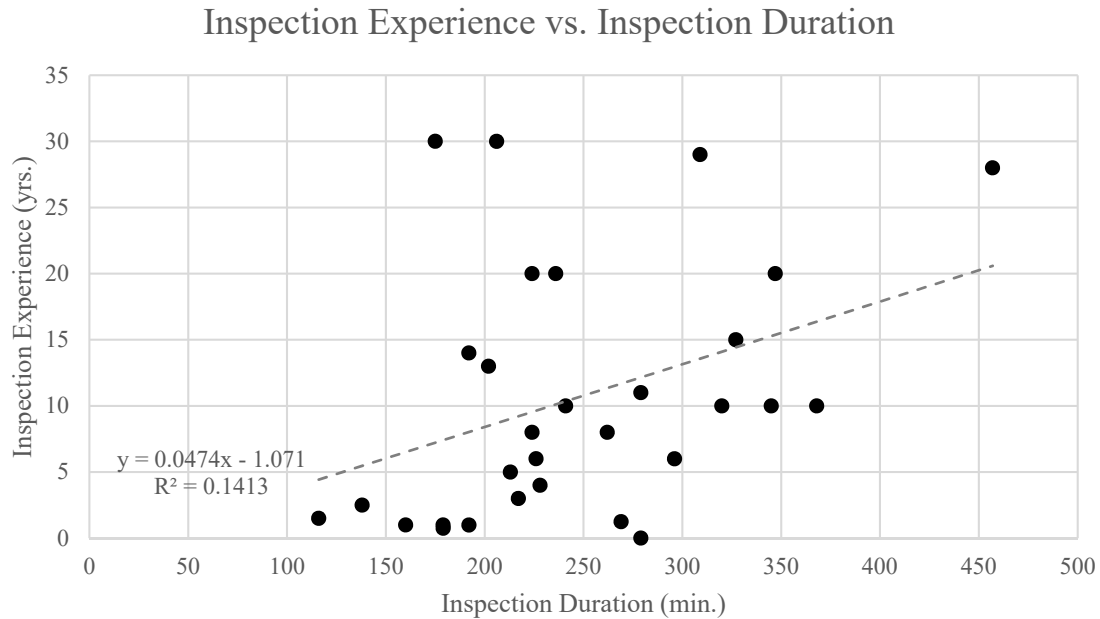


Figure 3.28 Inspection experience plotted against inspection duration

Similar to experience, there was a slight negative trend between the age of the inspector and detection rate as shown in Figure 3.29. The participants ranged in age from 24 to 63, with an average of 40 years. Once again, this trend is strongest for the welded cover plate specimens. As shown in Figure 3.30, four of the oldest participants had the lowest detection rates on these specimens. In at least three of these cases, the inspectors saw the majority of the cracks, but decided that they were only cracks in the paint, not in the base metal. These decisions were primarily based on previous experience with this detail and they were noted on the inspection forms and/or discussed with the proctor.



Figure 3.29 Detection rate plotted against inspector age

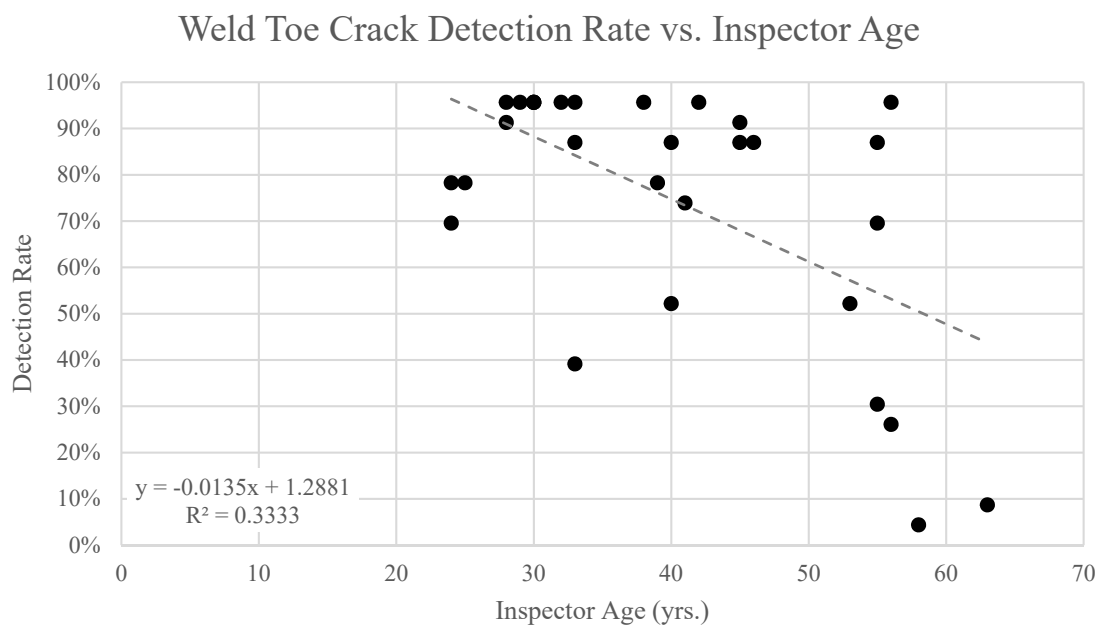


Figure 3.30 Weld toe crack detection rate plotted against inspector age

No single variable showed a statistically significant correlation with the number of false positives. As shown in Table 3.9 and Figure 3.31, neither temperature, experience, training, nor duration showed more than a very slight trend with the number of false calls made by the inspector.

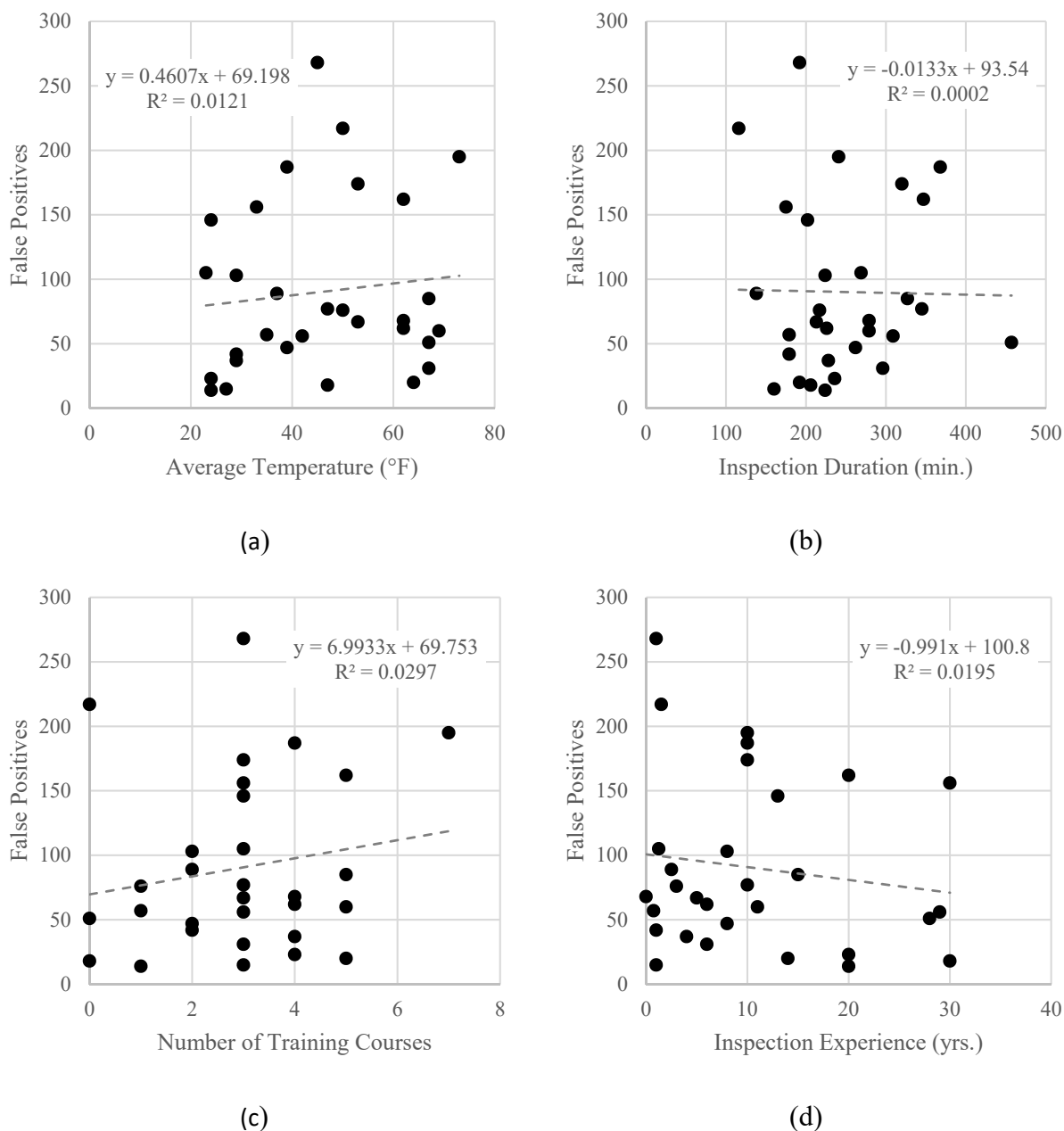


Figure 3.31. The number of false calls plotted against (a) average temperature, (b) inspection duration, (c) number of training courses, and (d) inspection experience

The student t -test indicated that the only variable that could be considered significant in discriminating between inspectors that made more and less false calls was the number of hands-

on inspections performed in the previous 12 months. Inspectors who had performed more than 12 hands-on inspections in the previous 12 months ($M = 57.6$, $SD = 42.3$) made significantly fewer false calls than inspectors who completed 12 or fewer hands-on inspections in the previous 12 months ($M = 106$, $SD = 72.4$), $t(27) = 2.33$, $p = 0.027$.

3.7.2 Multivariate Regression

To assess the interaction among the human and environmental factors, two multivariate linear regression analyses were performed to predict detection rate and the number of false positives. These analyses were completed in NLOGIT 6 by Econometric Software, Inc. In a multivariate regression model, the dependent variable Y_i is a function of many independent variables and so it is most easily expressed in matrix form

$$Y_{nx1} = X_{n \times p} \beta_{p \times 1} + \epsilon_{nx1} \quad (3.4)$$

where the dependent variable Y_{nx1} is a function of the parameters $\beta_{p \times 1}$ times the independent variables $X_{n \times p}$ and disturbance term ϵ_{nx1} [58]. The subscripts indicate the size of the matrices where n is the number of inspectors and p is the number of variables measured during each inspection. Again, ordinary least squares estimation was used to predict the parameter values for the multivariate regression model. Unlike the single variable models, the multivariate models can include both categorical and numeric independent variables. To properly account for categorical variables, an indicator, or dummy, variable that takes the value of either 0 or 1 for each entry is introduced. For example, the categorical variable for inspector training was divided into individual dummy variables for each training course. The inclusion of each independent variable in the model can be justified using the t -statistic and p -value discussed in Section 3.7.1 for univariate linear regression.

The adjusted R^2 value was used to assess the goodness of fit of the competing models. This statistic reflects the proportion of the total variance that is explained by the independent variables [58]. The adjusted R^2 value, as compared to the conventional R^2 value, considers the number of variables in the model and expresses the relative contribution of each variable. While models with more variables will have larger R^2 values, they will not necessarily have larger adjusted R^2 values. The R^2 and adjusted R^2 values are calculated using Equations 3.5 and 3.6 when Y_i is the actual value

of the dependent variable, \hat{Y}_i is the predicted value of the dependent variable, \bar{Y} is the average of the actual values of the dependent variable over all the observations, n is the number of observations, and k is the number of parameters included in the model [58]. The subscript i corresponds to the inspector ($i = 1, \dots, 30$).

$$R^2 = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} \quad (3.5)$$

$$adjusted R^2 = 1 - (1 - R^2) \frac{n - 1}{n - k} \quad (3.6)$$

For a perfect model, the variation of the fitted line around the observations is zero and R^2 will be one. This indicates that the model is able to explain all the variance in the data.

The multivariate analysis predicting detection rate provided similar results to the univariate analysis with detection rate best described by a function considering inspection duration, inspection experience, and training. The results from this analysis are shown in Table 3.10. Once again, the relationships between detection rate and duration and training were positive, while the relationship between detection rate and experience was negative. This model predicts that the completion of the FHWA/NHI *Underwater Bridge Inspection* course will produce an 11% increase in detection rate. As discussed above, the review of inspection principles, even though not directly related to fracture critical inspections, may improve crack detection. Alternatively, higher performing inspectors may be selected to attend this training, and so this indicator variable might reflect the performance assessment of supervisors or managers. The exclusion of temperature from this model suggest that the relationship between performance and temperature found in the univariate analysis may be an indirect relationship through duration. In other words, favorable inspection conditions encouraged inspectors to spend more time performing the inspection and the increased time resulted in improved performance. The adjusted R^2 statistic for this model was 0.622 indicating that the four parameters included in the model were able to explain 62% of the variability in the data. The R^2 value for the multivariate model is greater than any of the univariate models indicating that the interaction among the variables is better able to predict detection rate than any single variable.

Table 3.10 Results from multivariate regression analysis predicting detection rate

Variable Description	Estimated Parameter	Standard Error	t-statistic	P(> t)
Constant	0.460	0.054	8.56	0.0000
Inspection Duration (min.)	0.001	2.3E-4	4.93	0.0002
Inspection Experience (yrs.)	-0.010	0.002	-5.50	0.0000
Underwater Bridge Inspection (1 if inspector had taken the course, 0 otherwise)	0.105	0.039	2.71	0.012
Number of Observations	30			
Adjusted R ²	0.622			

As noted previously, no single variable showed more than a very slight relationship with the number of false calls. However, the multivariate linear regression analysis suggests that the number of false calls is best described by a function considering the inspector's employment sector, wind speed, tools used, and training. As shown in Table 3.11, the model predicts that inspectors that work for a private consultant, have completed the FHWA/NHI *Introduction to Element Level Bridge Inspection* course, and are exposed to higher maximum wind speeds, and do not use a tape measure will make more false calls. Conversely, inspectors that work for a public agency, have not attended the FHWA/NHI *Introduction to Element Level Bridge Inspection* course, experience lower maximum wind speeds, and use a tape measure will make fewer false calls. The adjusted R² statistic for this model is 0.373 indicating that 37% of the variability in the data is explained by the five parameters in the model. The reduction in the R² statistic implies that the number of false positives is still more difficult to predict than detection rate even with a multivariate model.

Table 3.11 Results from the multivariate regression analysis predicting false calls

Variable Description	Estimated Parameter	Standard Error	t-statistic	P(> t)
Constant	-16.2	48.2	-0.34	0.740
Employed by Private Engineering/ Inspection Firm (1 if yes, 0 otherwise)	77.2	22.5	3.44	0.002
Used a Tape Measure (1 if yes, 0 otherwise)	-65.4	22.1	-2.96	0.007
Max. Wind Speed (mph)	5.50	2.31	2.38	0.025
Introduction to Element Level Bridge Inspection (1 if inspector had taken the course, 0 otherwise)	57.4	22.7	2.53	0.018
Number of Observations	30			
Adjusted R ²	0.373			

The relationship between employer and false calls indicates that private consultants and public employees (state or federal) may apply different response criteria despite inspecting the exact same structure under the exact same instructions. Inspectors working for a public agency, state or federal, averaged 76 false calls while inspectors working for a private engineering or inspection firm made 112 false calls on average. Although the difference is not statistically significant, it does indicate that the two groups may be applying different response criteria. Inspectors that worked for a public agency were either more willing to make a decision at the risk of missing a crack or these inspectors used a more rigorous standard for recording defects.

The correlation between maximum wind speed and the number of false calls suggests that environmental conditions also influence the decision making process. Higher wind speeds may distract the inspector, challenging their ability to make accurate decisions.

Somewhat unexpectedly, the model predicts that completion of the *Introduction to Element Level Bridge Inspection* training course will increase the number of false positives. The reason for this is unclear. It may be an artifact of the study because the inspections began only a month after element level inspection became mandatory for all bridges on the National Highway System, and so only eight of the 30 inspectors had received the training. Alternatively, it could be that the inspection approach utilized in an element level inspection is more detailed and precise, leading inspectors to include every indication on their sketch.

The negative relationship between the number of false calls and the use of a tape measure also merits further investigation since the tape measure is obviously not a tool expected to improve visual acuity or the discriminability of flaws. Instead, this variable might actually be measuring an inspector's tendency to view the specimen more closely and the willingness to touch the inspection surface. These inspection techniques have been found to improve inspection performance in previous studies [5], [21], [52]. In this study, the proctor noted that the inspector that made the greatest number of false calls was "not very hands-on" and also that the inspector did not use a measuring tape. It is not known if using a tape measure forces an inspector to touch the inspection surface, or if inspectors that have predisposition to do this are also more likely to use a measuring tape.

3.7.3 Binary Logit Model

The variability in the POD curves presented in Section 3.6.2 suggests that factors beyond crack length affect the probability of detecting an individual crack. A binary logit model was developed to predict the probability of detection for each crack based on the hit/miss results for each inspector, the crack characteristics, and the human and environmental factors.

Instead of predicting the performance of an individual inspector, the intent of this analysis is to predict the likelihood that an individual crack will be detected by a specific inspector. This model is based on 2096 observations (28 inspectors x 70 cracks, 2 inspectors x 68 cracks). Each observation is recorded as either a hit (1) or miss (0). Because two discrete outcomes are considered, a binary logit model is appropriate [58]. These models are derived by defining a linear utility function, U_{in} , that describes the outcome i for observation n such that

$$U_{in} = \beta_i X_{in} + \varepsilon_{in} \quad (3.7)$$

where β_i is a vector of estimable coefficients for discrete outcome i , X_{in} is a vector of observable characteristics that determine the discrete outcome for observation n , and ε_{in} is a random disturbance term. For this single choice situation, the probability that observation n is a hit can be written as

$$P(hit) = Prob(U_{hit,n} > U_{miss,n}) \quad (3.8)$$

Substituting Equation 3.6 into Equation 3.8 and rearranging, the probability of a hit becomes

$$P_n(hit) = Prob(\varepsilon_{hit,n} - \varepsilon_{miss,n} < \beta_{hit} X_{hit,n} - \beta_{miss} X_{miss,n}) \quad (3.9)$$

and

$$P_n(hit) = F(\beta_{hit} X_{hit,n} - \beta_{miss} X_{miss,n}) \quad (3.10)$$

where $F(\bullet)$ is the cumulative density function of the difference of the disturbance terms. Setting the utility function describing a miss to zero and assuming the disturbance terms are extreme value Type I distributed, the difference will have a logistic distribution and the binary logit model results:

$$P_n(hit) = \frac{EXP(\beta_{hit}X_{hit,n})}{1 + EXP(\beta_{hit}X_{hit,n})} \quad (3.11)$$

However, a couple complications with this set of observations require minor adjustments to the probability formulation shown in Equation 3.11. The first complication is due to the nature of the data collected in this study. In this study, 30 inspectors provided 70 observations on the presence or absence of a crack. Since each inspector generated multiple observations, it is inevitable that these observations will share unobserved effects, thus violating the assumption of serial independence of the disturbance terms. The unobserved effects will cause the standard errors of the coefficients to be underestimated, thereby inflating the t -statistics. This will increase the likelihood of mistakenly rejecting a true null hypothesis [58]. This complication can be addressed with a random effects model which includes an individual specific disturbance term, μ_k , in addition to the overall disturbance term. Modifying Equation 3.6, the linear function describing the utility function for outcome i becomes

$$U_{kt}^i = \beta^i X_{kt}^i + \varepsilon_{kt}^i + \mu_k \quad (3.12)$$

where the observation, n has been replaced with the inspector number, k ($k = 1, \dots, 30$) and the crack number, t ($t = 1, \dots, 70$). All other terms are as previously defined. Accounting for random effects, the probability that crack t is correctly located by inspector k can be written as

$$P_{kt}(hit) = \frac{EXP(\beta^{hit}X_{kt}^{hit} + \mu_k)}{1 + EXP(\beta^{hit}X_{kt}^{hit} + \mu_k)} \quad (3.13)$$

The second complication arises from the complexity of visual inspection. As discussed previously, visual inspection is a complex and imperfect process involving a variety of human factors. Since many of these human factors cannot be quantified and explicitly captured in statistical models, results generated from these models may be biased or erroneous. Random parameter models were developed to consider these unaccounted for factors (typically referred to as unobserved heterogeneity) [58]. While random effects consider unobserved effects within an individual inspector's set of observations, random parameters account for the unobserved factors across the observations which might influence the likelihood of detection [59]. In this model, random

parameters allow the influence of certain explanatory variables to vary across the sample of inspectors. For example, while a standard binary logit model can account for the effect of experience on detection, only a random parameters model would consider the influence of underlying inspector characteristics, such as patience, confidence, or bias. These variables are not included in the model, but could affect the relationship between experience and probability of detection. Statistically, this possibility is accounted for with the introduction of a mixing distribution into the probability equation such that the probability that crack t is correctly located by inspector k is

$$P_{kt}(hit) = \int \frac{EXP(\boldsymbol{\beta}^{hit} \mathbf{X}_{kt}^{hit} + \mu_k)}{1 + EXP(\boldsymbol{\beta}^{hit} \mathbf{X}_{kt}^{hit} + \mu_k)} f(\boldsymbol{\beta}|\boldsymbol{\varphi}) d\boldsymbol{\beta} \quad (3.14)$$

where $f(\boldsymbol{\beta}|\boldsymbol{\varphi})$ is the density function of $\boldsymbol{\beta}$ with $\boldsymbol{\varphi}$ referring to a vector of the mean and variance of that density function. In this mixed logit model, $\boldsymbol{\beta}$ can now account for inspector specific variations of the effect of \mathbf{X} on crack detection probability with the density function, $f(\boldsymbol{\beta}|\boldsymbol{\varphi})$, used to determine $\boldsymbol{\beta}$. Typically, a normal distribution is assumed for the density function, although other distributions are feasible. Due to the complexity of estimating a different parameter for each observation, a simulation based maximum likelihood approach is used. As recommended by Washington, Karlaftis, and Mannering [58], a Halton sequence with 200 draws was used. In this way, values of $\boldsymbol{\beta}$ are drawn from $f(\boldsymbol{\beta}|\boldsymbol{\varphi})$, the probabilities are computed, and estimated parameters are selected to maximize the likelihood function. Model estimation was completed in NLOGIT 6 by Econometric Software, Inc.

The likelihood ratio test was used to confirm that both random effects and random parameters should be included in the final model. This test compares two competing models and determines how many times more likely the results are under one model as compared to the other. In this case, the proposed (unrestricted) model was compared to a reduced (restricted) [58]. The restricted model can be a model in which all parameters, except the constants, are set equal to zero or a model that considers only random effects or only random parameters. The likelihood ratio is calculated using Equation 3.15 where $LL(\beta_{RC})$ is the log-likelihood at convergence of the restricted model and $LL(\beta_U)$ is the log-likelihood at convergence of the unrestricted model.

$$X^2 = -2(LL(\beta_{RC}) - LL(\beta_U)) \quad (3.15)$$

The X^2 statistic is assumed to have a X^2 distribution with the degrees of freedom equal to the difference in the number of parameters in the two models. The p -value expressing the probability that the unrestricted model provides a superior fit to the restricted model by chance alone can be calculated from the X^2 distribution tables. A low p -value indicates that there is a low likelihood that the coefficients of the independent variables are equal to zero and so these variables are likely to be meaningful predictors of the outcome of the observation.

Since this model uses maximum likelihood estimation instead of least squares estimation, the ρ^2 -statistic is used to assess goodness of fit of the proposed models [58]. The adjusted ρ^2 -statistic considers the contribution of each explanatory variable and is calculated using Equation 3.16 where $LL(\beta_{RC})$ is the log-likelihood at convergence of the restricted model, $LL(\beta_U)$ is the log-likelihood at convergence of the unrestricted model, and k is the difference in the number of parameters estimated in the two models.

$$\rho^2 = 1 - \frac{LL(\beta_U) - k}{LL(\beta_{RC})} \quad (3.16)$$

For a perfect model, the likelihood function is equal to one (all predictions are exactly correct) and the log-likelihood function is therefore zero. Thus, a perfect model would have a ρ^2 value of one indicating that all the variance in the data is explained by the model.

Many combinations of factors were analyzed to determine which traits significantly impacted likelihood of detection. For this population of inspectors, the probability of detection for each crack is best described by a function considering crack length, crack type, inspector experience, inspection duration, and the elapsed time since the first inspection. Crack length and experience were found to vary across the inspector population, as indicated by the significance of their standard deviations. Specimen type, inspection duration, and time since the first inspection were considered fixed across the population.

This model includes both random effects due to the repeated observations from a single inspector and random parameters to account for the variability in the influence of the independent variables across the observations. Random parameters were included for estimated parameters with a standard deviation that was significantly different from zero. Random effects were accounted for by allowing the constant term to vary across the inspectors. Since this study consists of panel data (repeated observations from the same observer), random effects were included even though the standard deviation of the constant term was not significantly different from zero. The results from this model are shown in Table 3.12.

Table 3.12 Results from the binary logit model estimating probability of detection

Variable	Estimated Parameter (St. Dev.)	Standard Error Estimate (St. Dev.)	t-statistic (St. Dev.)	P(> t) (St. Dev.)
Constant	-0.902 (0.029)	0.170 (0.059)	-5.32 (0.50)	0.000 (0.618)
Crack Length (in.)	0.645 (0.371)	0.054 (0.029)	11.93 (12.91)	0.000 (0.000)
Out-of-Plane Crack (1 if out-of-plane crack, 0 otherwise)	-0.782	0.085	-9.22	0.000
Weld Toe Crack (1 if weld toe crack, 0 otherwise)	-0.775	0.111	-7.02	0.000
Inspection Duration (min.)	0.004	0.001	6.09	0.000
Inspection experience (yrs.)	-0.009 (0.021)	0.005 (0.004)	-1.89 (4.90)	0.059 (0.000)
Elapsed time since first inspection (days)	-0.0006	0.0002	-4.22	0.000
Number of Observations	2096			
Log-likelihood of constant	-1310			
Log-likelihood with random effects only	-1171			
Log-likelihood at convergence	-1136			
Adjusted ρ^2	0.126			

Marginal effects can be calculated to express how each parameter affects the probability of detection. The marginal effect gives the change in probability of detection for a unit change in the independent variable. The marginal effects for the model parameters can be seen in Table 3.13. A larger marginal effect indicates a greater influence on the likelihood of detecting the crack while a smaller marginal effect indicates a lesser influence. For the dummy variables, marginal effects are computed as the difference in the estimated probabilities when the variable is changed from zero to one, while all the other variables are set equal to their means. For continuous variables, the marginal effects are computed from the partial derivative of the probability equation. In both cases, the reported marginal effect represents the average over all the observations.

Table 3.13 Marginal effects for the parameters in the binary logit model

Variable Description	Avg. Marginal Effect (Std. Dev.)
Crack length (in.)	0.156 (0.087)
Out-of-Plane Crack (1 if out-of-plane crack, 0 otherwise)	-0.159 (0.024)
Weld Toe Crack (1 if weld toe crack, 0 otherwise)	-0.153 (0.031)
Inspection Duration (min.)	6.41E-4 (2.73E-4)
Inspection experience (yrs.)	-0.002 (0.002)
Elapsed time since first inspection (days)	-1.11E-4(4.72E-5)

In addition to the relationships between performance and environmental conditions and inspector attributes, this model provides insight into how the characteristics of the crack may influence the probability of detection. On average, a one-inch increase in crack length resulted in a 15.6% increase in probability of detection. For 96% of the observations, an increase in crack length resulted in a higher probability of detection for that crack. For the remaining 4% of the observations, the probability of detection decreased with increasing crack length. A similar trend was reflected in the probability of detection curves discussed in Section 3.6.2. In these cases, the failure to detect the crack may have been caused by an error in the decision task, as opposed to the search task, as was observed on the welded cover plate specimens.

Supporting the previous finding of a positive correlation between detection rate on the girder specimens and the welded cover plate specimens, the binary logit model indicated that both the out-of-plane cracks and weld toe cracks were approximately 16% less likely to be detected as compared to the rivet hole cracks of the same length.

Cracks were more likely to be located during inspections that lasted longer. This variable was fixed across the population and a one-minute increase in inspection time increased the likelihood of detecting a crack, regardless of length, by 0.06%.

Similar to the regression models, the binary logit model also revealed a negative relationship between probability of detection and inspection experience. However, this model was able to capture the variation in influence across the sample of inspectors. For 67% of the observations, the probability of detecting a crack decreased with experience while, for the remaining 33% of the

observations, the likelihood of finding a crack increased with experience. On average, a one-year increase in experience reduced the probability of detecting the crack by 0.2%.

Individual cracks were more likely to be found during the earlier inspections and less likely to be found during the later inspections. This variable is fixed across the population, and the likelihood of detecting a specific crack decreased approximately 8% from the first inspection on Day 1 of the study to the last inspection on Day 762. This is not unexpected as specimen maintenance is a well-documented challenge to POD studies [17]. This trend was not present in the regression analyses indicating that for an individual inspector, performance is more heavily influenced by human factors than by crack characteristics. As shown in Figure 3.8, on seven occasions there was more than a 10% difference in detection rates between inspections that happened on consecutive days. Conversely, there was a 454-day gap between the 11th and 12th inspectors, but the difference in detection rates was only 9%. The effect of human factors on detection rate significantly outweighs the influence from the change in the crack appearance, and so it remains valid to directly compare the results from all 30 inspectors over the 2+ year study.

For this model, the restricted log-likelihood of constants only ($LL(\beta_{RC})$) is -1310, the unrestricted log-likelihood of the proposed model ($LL(\beta_U)$) is -1136, the number of degrees of freedom (ν) is 8, and the X^2 statistic is 346. This exceeds the critical value of 31.8 and a confidence level of over 99.99% is achieved. In other words, the probability that the unrestricted model provides a superior fit to the restricted model by chance alone is very small (less than 0.01%). Similarly, for a model that includes only random effects, the restricted log-likelihood ($LL(\beta_{RC})$) is -1171, the unrestricted log-likelihood function ($LL(\beta_U)$) is -1136, the number of degrees of freedom (ν) is 2, and the X^2 statistic is 69. This exceeds the critical value of 18.4 and a confidence level of over 99.99% is achieved. Therefore, it is appropriate to include random parameters in the model.

3.7.4 Vigilance

As noted previously, one of the common findings in visual inspection research is the presence of a vigilance decrement [25], [21], and so the relationship between performance and time on the task was investigated in this study. Only performance on the girder specimens was considered since all 72 specimens were typically inspected without a break. Additionally, these specimens are

arranged in four even rows of 18 specimens making it straightforward to compare performance on a row-by-row basis using the independent two-sample *t*-test.

Figure 3.32 shows the average detection rate, number of false calls, and inspection duration for each girder row. The performance measure is plotted on the left vertical axis and represented by the bars while inspection duration is plotted against the right vertical axis and represented by the line with markers. The average detection rate was higher on the first 36 girder specimens ($M = 0.653$, $SD = 0.148$) as compared to the second 36 girder specimens ($M = 0.582$, $SD = 0.151$), $t(57) = 1.85$, $p = 0.070$. Twenty-three (23) of the 30 inspectors achieved higher detection rates on the first two rows as compared to the last two row. Similarly, a slight “warm-up effect” is noticeable in the inspection results as performance on the first row is lower than performance on the second or third rows for 24 of the inspectors. There was no noticeable improvement in the detection rate of inspectors that took a break during the girder inspection and those that did not. One possible explanation is that the break schedule was controlled by the inspector throughout the inspection rather than being predetermined. Inspectors may have already been suffering the ill effects of prolonged vigilance before they took a break.

Unlike detection rate, there was no standard or predictable change in the number of false calls made with increasing time on the task. The average number of false calls per specimen row remained relatively similar at 13, 19, 18, and 15. Individual inspectors may have adjusted their response criterion during the inspection though. For instance, one inspector gradually increased the number of false positives from 2 on the first row of specimens to 18 on the fourth row of specimens, while another inspector reduced the number of false calls from 16 to 2 on the same girder rows. For 13 of the inspectors, the difference in the number of false calls between the first two rows and the last two rows was greater than nine.

In addition to changes in detection rate and the number of false calls, inspection rate also increased with increasing time on the task. The inspection time for the first two girder rows ($M = 86$, $SD = 37$) was significantly longer than the inspection time the last two girder rows ($M = 59$, $SD = 22$), $t(46) = 3.45$, $p = 0.001$. A study considering temporal effects on a security inspection task noticed

a similar trend and determined that the inspectors gradually decreased their search time, but did not adjust their non-search time, as time on the task increased [60].

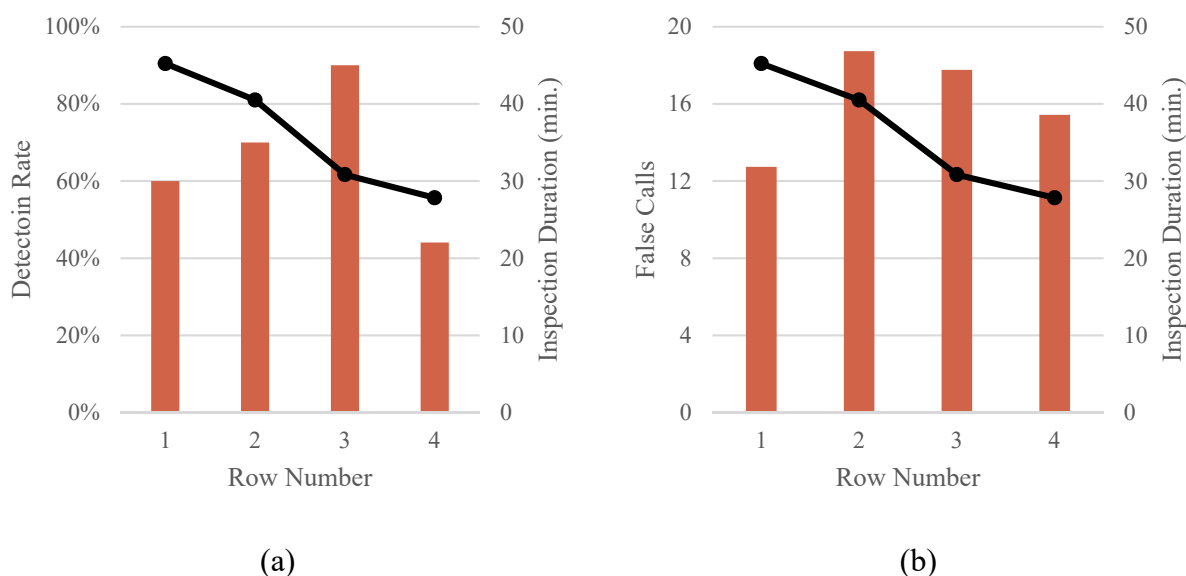


Figure 3.32 (a) Detection rate and inspection duration by girder row and (b) false alarms and inspection duration by girder row

3.7.5 Tool Use

No significant correlation was found between crack detection and any single tool that was used. This is likely due to the large variety in tools used during the inspections and the lack of information collected about these tools and their use.

Although this study did not find significant correlation between flashlight use and inspection performance, this should not be interpreted to mean that flashlights are not necessary or ambient lighting under a bridge is adequate for visual inspection. Additional information that was not collected during this study, such as the ambient lighting under the bridge during the inspection and the specific specimens on which a flashlight was used, is necessary to determine minimum lighting levels and if flashlights improve the likelihood of detecting a fatigue crack. Absent this information, a flashlight with a minimum light output of 100 lumens and an adjustable focus is recommended for use during all hands-on bridge inspections.

3.7.6 Discussion

Several of the human factors findings enhance current understanding of visual inspection of steel bridges and are supported by previous research on the visual inspection process from other industries.

First, there is incredible variability in visual inspection results and most of the variables expected to correlate with performance are actually weak predictors of performance. Slight trends were found between the performance measures and a few of the individual or environmental factors, however no single variable was able to explain much of the variance in the data. Even when considering the interaction among the factors, the predictive power of the models is low. For every trend, there seems to be an exception. Even though detection rate generally decreased with increasing experience, there was an inspector (Inspector A) with 28 years of experience and a 70% detection rate. And while the number of false positives tended to decrease with an increasing number of hands-on inspections, there was an inspector (Inspector B) who made 162 false calls despite performing 25 hands-on inspections in the previous 12 months. While these inconsistencies seem to suggest that human factors cannot be used to explain inspection performance, it is more likely that these quantitative variables simply cannot tell the full story. For instance, the model does not know that Inspector A is a Certified Weld Inspector and Level III NDT technician who mainly inspects offshore drill rigs and ships. Therefore, it is possible that he/she has not developed the same expectations or biases as a full time bridge inspector. Similarly, the model cannot account for the unexpected rain storm that delayed Inspector B and may have caused him/her to rush through the second half of the inspection.

Second, the current results suggest that different interventions may be necessary to improve the search and the decision components of visual inspection. As discussed previously, missed defects can result from search or decision errors while false alarms are the result of decision errors. Therefore, by comparing the factors that affect detection rate and false calls, it may be possible to determine if these factors are influencing the search function or the decision function. For instance, the absence of a clear relationship between the number of false calls and inspection duration suggests that the additional time is more influential in the search task than the decision task. In other words, the additional time may allow an inspector to complete a more thorough examination

of the inspection surface, but may not improve the accuracy of the decision making process. This is supported by previous visual inspection research that has found improvements in decision making to saturate quickly with increasing time [21].

Third, the significant variability in false calls and the lack of correlation with individual or environmental factors imply that decision making may be the more unpredictable aspect of the visual inspection process. Although the large number of false calls may be an artifact of the study, they still suggest a high degree of uncertainty within this inspector population. As noted previously, inspectors were not allowed to clean the surface, remove paint, or use enhanced NDT methods to confirm their findings. To compensate, a few inspectors labelled indications as “possible” or “probable” cracks and noted the need for NDT on the inspection forms. Notably, inspectors rarely equivocated in their decisions on actual fatigue cracks, but showed more uncertainty about cracklike surface flaws. Reviewing the results for two inspectors that did this most consistently, it is interesting to note that actual fatigue cracks were rarely labelled as “possible” cracks, while paint scratches or other surface imperfections were often labelled this way. It seems that these inspectors recognized a true fatigue crack when they saw it, but lacked confidence in their decision making. After re-evaluating the results on the girder specimens for these two inspectors considering only the indications marked clearly as cracks, the first inspector recorded 3 fewer hits but 45 fewer false positives and the inspector’s hit/call ratio increased from 31% to 77%. Similarly, the second inspector recorded 2 fewer hits but 23 fewer false positives and the hit/call ratio increased from 25% to 56%. Once again, it is unknown why so many false positives were recorded, but this type of performance feedback may result in measureable improvements in performance.

Fourth, the probability of detection curves generated through this study exhibit significant inspector-to-inspector variation. Either individual inspectors respond differently to crack length, or crack length explains very little of the variation in performance. For three of the inspectors, the probability of detection actually decreased with increasing crack length, indicating that the influence of crack size is largely overshadowed by inspector or environmental factors. And for other inspectors, the curves are not asymptotic to $POD = 1$, leaving a nonzero probability of missing a crack regardless of length. Therefore, while other aspects of fracture prevention (material toughness, design requirements, fabrication tolerances, etc.) can be used to increase the

critical crack size to a length that is easily visible to the naked eye, it may be unrealistic to assume that even a very long crack will always be found under the current inspection procedures.

3.8 Recommendations for Visual Inspection

Current findings confirm that the visual inspection of a steel bridge is a difficult task and human inspectors struggle with this highly subjective assignment. Potential solutions typically focus on one of three areas: inspection equipment and environment, inspector training, and inspection procedures. Recommendations from this study will be similarly categorized.

3.8.1 Equipment

A standard set of equipment should be provided to each inspector to improve consistency among inspectors. Standard tools should include, but are not limited to, a flashlight with light output of at least 100 lumens and adjustable focus, a wire brush, scraper, and hammer, 5x power and lighted 10x power magnifying glasses, a telescoping inspection mirror, and an easy to read measuring device. Inspectors should be provided with basic training outlining the proper use for each tool. Before beginning the inspection, inspectors should complete an equipment checklist indicating that the prescribed tools were available and in working order.

3.8.2 Training

Visual inspection research frequently cites inspector training as the most cost effective and efficient strategy for improving inspection performance [20]. However, this study revealed that the current training program, which relies heavily on classroom lectures, produces highly variable results. Improvements to the training program should be made so that it addresses all aspects of the inspection task – procedural, physical, and cognitive. Proven training systems employ a modular approach, include immediate feedback during training, encourage active participation and engagement from attendees, allow for self-discovery, clearly define acceptability standards, discuss common errors, and address a wide variety of possible defects [20]. A new training course and training module developed in accordance with these recommendation is discussed in detail in Chapter 4.

While federal law mandates periodic bridge inspection refresher training, no specifics are provided on the topic or frequency [1, §650.313]. Typically, all inspectors in the state receive identical training at a set interval. While this is helpful in improving consistency across the inspection program, it does not address the needs of an individual inspector. For inspectors requiring remedial training, it may be helpful to determine in what function(s) they are deficient so that the appropriate training can be provided. Since the human factors influencing the search and decision functions may differ, it follows that that individual inspectors may require different interventions to improve inspection performance. Although it is not possible to identify the exact cause of each error, the type and total number of errors provides insight into the relative strengths and weaknesses of the inspectors. For instance, an inspector that misses a large number of cracks, but also makes a low number of false calls, likely requires additional training in the search task, while an inspector that detects most of the cracks, but also makes a large number of false calls, may need training focused on decision making. Within this group, 18 inspectors were identified as needing additional instruction in performing the search task, one inspector should receive training in the decision component, and nine inspectors require interventions in both functions. This was determined by comparing an individual's performance to the overall performance of the group. Inspectors with a detection rate less than one standard deviation above the average (79%) were classified as needing search interventions and inspectors with more than the average number of false calls (90) were assumed to need decision interventions. The larger number of inspectors requiring search intervention makes sense since a higher standard was applied to detection rate than the number of false positives. This is reasonable given the importance of the search component and the relative cost of a miss as compared to a false positive. This result also agrees with the FAA study on visual inspection reliability which found that all 12 participating inspectors needed supplemental training on how to perform the search task, while only a few inspectors required additional guidance on performing the decision task [52]. Performance testing, as discussed in Sections 3.8.3 and 3.8.4, can be used to establish performance thresholds for remedial training and to determine what type of training an inspector requires.

3.8.3 Inspection Procedures

Detailed procedures for hands-on visual inspection of steel bridge members should be developed. These procedures should use the same terminology that is used in inspection training courses and

include likely crack locations and the recommended tools and inspection techniques for typical bridge details. Much of the relevant information for steel bridges is compiled in the *Fatigue and Fracture Library for the Inspection, Evaluation, and Repair of Vehicular Steel Bridges* [61]. The procedures should be organized into job cards that include a clear description of the task along with the information needed to correctly complete the task. Well designed and appropriately used job cards were found to improve performance during the FAA study on inspection reliability [52]. The most effective job cards included photos or diagrams of the inspection area that matched the orientation in the field and clearly showed the likely defect locations and inspection boundaries.

Additional procedural recommendations are offered based on visual inspection research, although their effectiveness for hands-on bridge inspection was not explicitly evaluated in this study:

- Encourage active observation. Although visual cues are the main source of information about the condition of the structure, previous research has found that inspectors can improve their performance by touching the surface, listening for rattles or squeaks, and smelling leaks or overheating parts. Even while using only the eyes, inspectors should frequently adjust their viewing angle and distance from the inspection surface and use basic tools to increase the discriminability of flaws [21], [52]. The correlation between false calls and tape measure use observed in this study suggests that inspectors that are close enough to the inspection surface to measure the length of the crack, may also be close enough to make accurate judgments about detected indications. Additionally, these inspectors may be more willing to use their tactile sense to inform decision making. Based on the results of this study, requiring inspectors to use a tape measure may be one way to ensure that they are actively engaged with the inspection surface.
- Hold regular calibration meetings and refresher training. Calibration meetings and refresher training can help maintain a uniform understanding and application of inspection standards among the inspector population. This will help improve consistency in decision making and reduce the variability in inspection results. These trainings should be specific to the managing agency and attended by all inspectors performing inspections under their guidance. These meetings can help to address the difference in decision making observed between inspectors employed by public agencies and those employed by private consultants.

- Provide regular feedback. The usefulness of the qualitative information gathered in this study highlights the need for ongoing communication between inspectors and managers. As discussed previously, a more thorough understanding of an inspector's strategy and thought process was gathered by simply spending a day observing and talking with them. Similar interactions between inspectors and their supervisors would offer supervisors the same insights and allow them to provide feedback to improve accuracy and consistency. Although it is difficult to provide performance feedback since the true accuracy of an inspector's findings are not known, cognitive feedback has also been found to be effective in improving inspection performance [35]. Cognitive feedback may focus on an inspector's strategy, expectations, or assumptions. For instance, this type of feedback could correct an inspector's misconception about the relative cost of a missed crack and a false call. Without feedback, the same errors in judgment and decision making will be unknowingly repeated in the future.
- Rotate inspectors. Even with similar experience or training, these results show that inspectors have differing abilities to detect cracks. In fact, only one crack was detected by all 30 inspectors and only three cracks were not detected by any of the inspectors. This supports the practice of inspector rotation as a quality control measure intended to reduce the likelihood that defects are repeatedly missed [4].
- Allow adequate time to complete each inspection and encourage inspectors to use the allotted time. Adequate time should be allowed for the inspection so that inspectors do not have to rush and are able to take frequent breaks. These breaks do not need to be non-working breaks, simply alternating tasks is usually adequate to combat the vigilance decrement. The literature recommends switching tasks or taking a break every 20 to 30 minutes to maintain the necessary level of attention and focus [20].
- Establish initial and recurrent performance testing requirements. The current findings confirm previous studies which have determined that it may not be effective to rely on inspector demographics for selection or evaluation. Performance testing in a controlled environment should be utilized to confirm that an inspector can correctly apply inspection procedures in the field and achieve a satisfactory level of performance. This is discussed in greater detail in Section 3.8.4 below.

3.8.4 Performance Testing

The large variability in the inspection results and the relatively weak predictive value of many of the variables expected to correlate with performance suggest a need for performance testing for certifying bridge inspectors. Performance testing has been successfully implemented on a project-by-project basis to verify the quality and improve the consistency of nondestructive testing during steel bridge inspections [62]. In addition to the classroom training and written examination currently required under the NBIS, a practical inspection test in a controlled environment should be administered.

Effective performance testing for visual inspection should include both a quantitative and a qualitative assessment. Quantitatively, inspectors must be able to achieve a minimum passing score to be recommended for certification. Qualitatively, inspectors must demonstrate a reasonable inspection strategy and logical thought process that informs both how they inspect and where they inspect.

Currently, no standard performance measures exist for qualifying bridge inspectors for visual inspection of steel bridges. However, the data generated in this study can be used to investigate the outcomes from potential policies. To illustrate, five different hypothetical performance standards were identified, and for each performance standard, several passing criteria of varying difficulty were considered. These standards are presented below for discussion only. Additional research is necessary to establish rational performance criteria.

One option is to establish a flat detection rate for passing. For example, the inspectors would each be required to find 70% of the total cracks present on the course. This option is simple and easy to determine if an inspector meets the criterion; however, it does not consider the size or location of the cracks or the number of false calls made by the inspector. Table 3.14 summarizes the inspection performance of the passing inspectors for a range of required detection rates.

Table 3.14 Summary of results for inspector qualification based on unweighted detection rate

Required Score to Pass	Pass Rate (No. of Inspectors)	Avg. No. of Missed Cracks	Avg. No. of False Calls	Max. No. False Calls	Largest Crack Missed
50% (35 hits)	83% (25)	21	89	268	5-3/8"
60% (42 hits)	77% (23)	20	92	268	5-7/32"
70% (49 hits)	43% (13)	16	96	195	4-29/32"
80% (56 hits)	10% (3)	11	111	187	3-1/4"

A second option is to use a graduated scale for detection rates. Cracks of greater length must be detected at a higher rate than cracks of shorter length or less severity. For instance, to pass the practical exam, an inspector must find 50% of cracks shorter than 1 inch, 65% of cracks with lengths between 1 and 3 inches, and 80% of cracks greater than 3 inches long. This method stresses and presumes that inspectors should be better able to find longer cracks more easily than shorter cracks. Although crack length is considered, criticality of the crack is not directly included in the test since the location of the crack is not considered. Also, inspectors are not penalized for false calls. Table 3.15 summarizes the inspection performance of the passing inspectors for a range of required detection rates.

Table 3.15 Summary of results for inspector qualification based on weighted detection rate

Required Score to Pass (<1", 1"-3", >3")	Pass Rate (No. of Inspectors)	Avg. No. of Missed Cracks	Avg. No. of False Calls	Max. No. False Calls	Largest Crack Missed
40%, 60%, 80%	33% (10)	16	100	195	4-29/32"
50%, 65%, 80%	20% (6)	13	120	195	3-3/8"
55%, 70%, 80%	13% (4)	12	124	187	3-3/8"
65%, 75%, 80%	7% (2)	11	73	77	3-1/4"

A third option is to set the passing criteria based on crack type. Setting the criteria based on the type of crack, and how detrimental the particular crack type is for the structure, promotes a reasonable approach to bridge inspection. For example, load induced cracks grow perpendicular to the stress, and can quickly lead to member fracture. For this crack type, the detection rate for passing may be set to be 70%. For distortion induced cracks that typically grow more slowly and pose less risk, a lower detection rate, possibly 50%, may be acceptable. Table 3.16 summarizes the inspection performance of the passing inspectors for a range of required detection rates.

Table 3.16 Summary of results for inspector qualification based on crack type detection rate

Required Score to Pass (Distortion Induced, Load Induced)	Pass Rate (No. of Inspectors)	Avg. No of Missed Cracks	Avg. No. of False Calls	Max. No. False Calls	Largest Crack Missed
40%, 60%	33% (10)	17	76	187	4-29/32"
50%, 70%	13% (4)	12	92	187	3-1/2"
60%, 80%	7% (2)	11	132	187	3-1/4"

A fourth option is to evaluate inspector performance based on both detection rate and the number of false calls. Large numbers of false positives result in higher costs and can undermine the veracity of a bridge inspection program. The rating could be calculated using Equation 3.17 where n is the relative severity of a miss to a false call [63]. For example, a minimum rating of 75% could be required with a missed crack considered three times more severe than a false call. Table 3.17 summarizes the inspection performance of the passing inspectors for a range of required inspector ratings.

$$Rating = \frac{1}{2} \times \left(1 + \frac{\text{detected flaws}}{\text{total flaws}} - \frac{\text{false indications}}{n \times (\text{total indications})} \right) \quad (3.17)$$

Table 3.17 Summary of results for inspector qualification based on inspector rating

Required Rating to Pass	Pass Rate (No. of Inspectors)	Avg. No. of Missed Cracks	Avg. No. of False Calls	Max. No. False Calls	Largest Crack Missed
75% (n = 4)	60% (18)	19	81	195	4-29/32"
80% (n = 4)	23% (7)	14	76	187	3-1/2"
75% (n = 3)	40% (12)	17	70	187	4-29/32"
80% (n = 3)	7% (4)	13	53	77	3-1/2"

The fifth option is to evaluate inspector performance based on probability of detection by crack length. Under this standard, a shorter POD crack length indicates superior performance. For instance, to pass the exam, an inspector must be able to detect a 4-inch crack 90% of the time with 95% confidence. Similar to the weighted detection rate standard, this standard considers crack length, but not location. Additionally, inspectors are not penalized for false calls. Table 3.18 summarizes the inspection performance of the passing inspectors for a range of required POD crack lengths.

Table 3.18 Summary of results for inspector qualification based on probability of detection

Maximum POD Crack Length for Passing	Pass Rate (No. of Inspectors)	Avg. No of Missed Cracks	Avg. No. of False Calls	Max. No. False Calls	Largest Crack Missed
$a_{90} = 4$ in.	43% (13)	17	90	195	4-29/32"
$a_{90} = 5$ in.	63% (19)	19	91	268	5-7/32"
$a_{90/95} = 4$ in.	7% (2)	18	65	67	3-1/4"
$a_{90/95} = 5$ in.	27% (8)	16	85	195	4-29/32"

3.9 Summary

A diverse group of 30 inspectors were invited to complete a full day, hands-on inspection of 147 steel bridge specimens to establish a quantitative measure of visual inspection capability for fatigue cracks in steel bridges. The average detection rate was 65% and the average number of false calls was 90. The average 50% and 90% detection rate crack lengths were 1 inch and 5-1/2 inches, respectively. These lengths are likely larger than expected by most within the bridge inspection community, but are consistent with similar probability of detection studies in other industries [52], [64].

Inspectors employed widely varying inspection strategies; some adopted a conservative approach and categorized all indications as cracks, while others labelled only the most obvious indications as crack. The hit/call ratio varied from 13% to 75% and the number of hits was not correlated with the number of false calls. The study also considered the effect of inspector characteristics and environmental conditions on detection rate and the number of false calls. These performance measures were selected as they help describe an inspector's proficiency in both the search and decision components of visual inspection.

The variability in inspection results was substantial, and only a small amount of the variance could be described by human factors. This is not because human factors are insignificant or unimportant, but because most human factors are too complex to measure and summarize in a statistical model. Still, detection rate was found to be the more predictable performance measure, and it was most influenced by inspection time, training, temperature, and experience. Notably, the relationship between detection rate and experience was negative indicating that increased inspection experience actually reduced detection capability. Since a false call results from an error in judgment, the

variables found to influence the number of false calls were also assumed to be the variables that affect decision making. These factors include the inspector's employment sector, measuring tape use, wind spend, training, and the number of past inspections. Additionally, a slight decline in detection rate with increasing time on the inspection task was observed, but no consistent change in the number of false calls was detected.

Since the POD curves suggested that factors beyond crack length affect probability of detection, a binary logit model was developed to predict crack detection based on crack, human, and environmental characteristics. This model determined that crack detection was influenced crack length, the location or type of crack, inspector experience, inspection duration, and the elapsed time since the first inspection.

These results confirm that visual inspection of a steel bridge is an inherently difficult task. It is not a matter of "trying harder"; modifications to the current training scheme, inspection procedures, and inspection equipment are necessary to improve accuracy and consistency. Performance testing should be used to confirm the effectiveness of any changes and to ensure that all inspectors can provide the necessary inspection quality.

4. INSPECTOR TRAINING

4.1 Introduction

Based on the results from the hands-on inspections, it is clear that the current standards for visual inspection of steel bridges produce highly variable results. One possible explanation is insufficient training or lack of consistency in training. The FHWA study on the accuracy and reliability of routine bridge inspections recommended improved training on bridge behavior after noting that many inspectors could not, or did not, identify basic structural features of the bridge such as span type, support condition, or skew [5]. However, focusing on the detection of fatigue cracks in steel bridges, the currently available training courses and material seem to provide a strong background in where and how fatigue cracking may occur. All relevant training discusses the mechanics of bridge behavior, including tension and compression regions, in-plane and out-of-plane loading, and fatigue prone details. And yet, in reviewing the inspection results, it is clear that many inspectors, regardless of training and experience, may not fully understand how and where fatigue cracking occurs. A number of inspectors identified what they believed to be cracks in highly unlikely locations or orientations. For instance, many inspectors identified cracks in the plain web section of the girder specimens or parallel to the primary stress in the riveted plates. Although inspectors were regularly reminded to do their best to differentiate between fatigue cracks and other crack-like surface defects, some may have adopted an unrealistically conservative inspection strategy to limit their misses in the research environment. Still, these type of errors imply that inspectors either do not have a proper understanding of the mechanics of fatigue cracking or are unable to apply the theories taught in the classroom in the field. While the excessive numbers of false calls may be partially a construct of the inspection scenario, it does reveal a high level of uncertainty among some of the inspectors, and during actual bridge inspections, this uncertainty may produce the opposite result – fewer false calls and more missed cracks.

4.2 Observational Skills Training

To address the disconnect between training and execution, a new training course focused on the cognitive skills used during visual inspection was developed. Since there is not a practical way to apply a standardized step-by-step procedure to visual inspection, this training course objects to

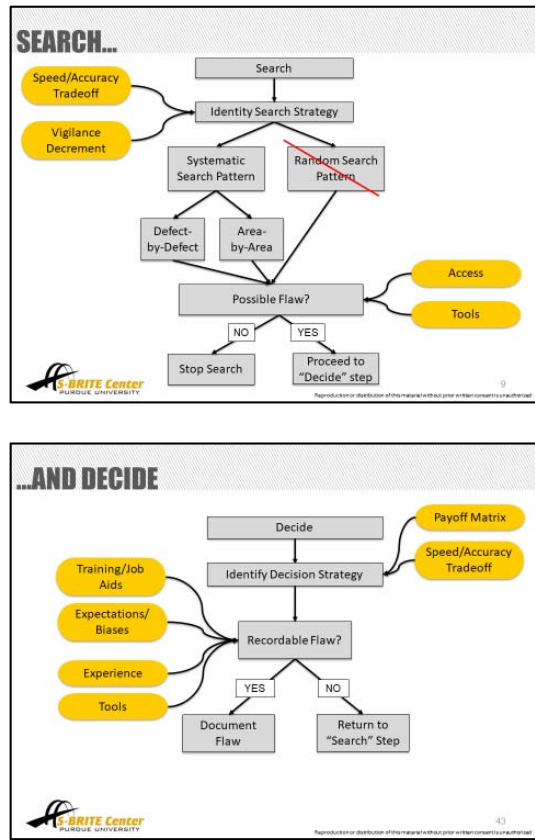
improve the application of technical knowledge in the field by teaching inspectors “how to inspect” rather than “where to inspect” or “what to inspect”. The course first divides the inspection process into four distinct subtasks (prepare, search, decide, and document) and then identifies the observation skills used in each task (perception and recognition, attention, memory, mental imaging and mental models, and judgement and decision making). These are the skills that enable an inspector to observe and interpret environmental information, compare it to previous experiences and knowledge stored in long-term memory, and use that analysis as the basis for a decision. This course covers both the theory and application of these skills and offers techniques for improvement. Additional details for the training units on each observation skill are provided below. Similar to a course developed for the International Atomic Energy Agency [65], [66], this training course is tailored to bridge inspection, although the observation skills are common to all visual inspection activities. The information is delivered through a combination of lecture, class discussion, and individual exercises. The training includes a list of good practices for steel bridge inspection based on the recommendations from visual inspection literature and the findings from the hands-on inspections.

This training is intended to be a half-day standalone course offered through the S-BRITE Center. The course material was submitted to the INDOT through SPR-3820 and is under review by the Study Advisory Committee. The projected delivery date is Fall 2019. Feedback will be solicited from course attendees and attendees will be invited to participate in the probability of detection study to evaluate the effectiveness of the training.

4.2.1 Visual Inspection Process

The training starts with an introduction to the visual inspection process including a description of the prepare, search, decide, and document subtasks. Focusing primarily on the search and decide tasks, the training discusses the findings from literature and the hands-on inspections including the utility of a systematic search and decision strategy, the speed versus accuracy trade-off, the vigilance decrement, and the influence of human factors on inspection performance. Specific details of the unit on the visual inspection process are provided in Table 4.1.

Table 4.1 Description of the visual inspection process unit

Objective	Convey how the inspection process can be broken down into a series of subtasks
Topic(s)	<ol style="list-style-type: none"> 1. Define and describe visual inspection tasks 2. Research findings and best practices 3. Introduction to cognitive skills
Exercise(s)	Computer based search and decide activities (see below)
Sample Slide(s)	 <p>The figure consists of two flowcharts illustrating the visual inspection process.</p> <p>SEARCH... Flowchart:</p> <ul style="list-style-type: none"> Starts with "Search". Leads to "Identify Search Strategy". From "Identify Search Strategy", it branches into "Systematic Search Pattern" and "Random Search Pattern". "Systematic Search Pattern" further branches into "Defect-by-Defect" and "Area-by-Area". Both "Defect-by-Defect" and "Area-by-Area" lead to "Possible Flaw?". "Possible Flaw?" leads to "NO" (Stop Search) or "YES" (Proceed to "Decide" step). Factors influencing the search process include "Speed/Accuracy Tradeoff", "Vigilance Decrement", "Access", and "Tools". <p>...AND DECIDE Flowchart:</p> <ul style="list-style-type: none"> Starts with "Decide". Leads to "Identify Decision Strategy". From "Identify Decision Strategy", it leads to "Recordable Flaw?". "Recordable Flaw?" leads to "YES" (Document Flaw) or "NO" (Return to "Search" Step). Factors influencing the decision process include "Training/Job Aids", "Expectations/Biases", "Experience", "Tools", "Payoff Matrix", and "Speed/Accuracy Tradeoff".

Two computer-based exercises were developed to allow an inspector to evaluate and improve their search and decision capabilities. These exercises were built in MathWorks MATLAB but can be delivered as standalone applications. The first exercise is solely a visual search task in which alphanumeric targets are identified among similarly shaped distractors. A similar exercise was used by Leach and Morris [31] and Koenig, et al. [67] to study cognitive factors and human performance during visual inspection. This simulation can be used to show how search performance varies with pacing (self-paced vs. machine paced), search strategy (systematic vs. random), and fault type (single fault vs. multi-fault). The second exercise is solely a decision task

Table 4.2 Description of the perception and recognition unit

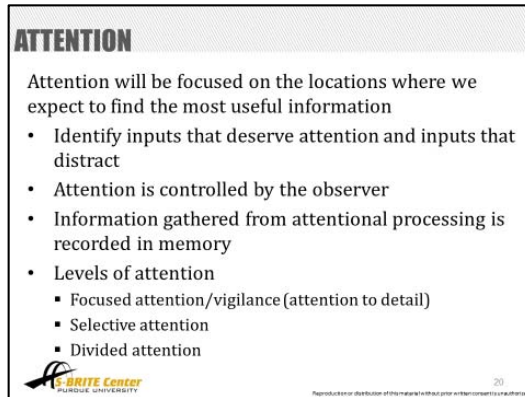
Objective	Convey how perception and recognition work and how these processes can enrich or inhibit inspection tasks
Topic(s)	<ol style="list-style-type: none"> 1. Definitions of perception and recognition 2. Rules of perception (simplicity, closure, figure-ground) 3. Illusions (<i>The dress</i>, Muller-Lyer illusion) 4. Recognition theories (template matching, prototype matching, recognition by component)
Exercise(s)	Embedded figures (identify a critical target among a distracting background)
Sample Slide(s)	

4.2.3 Attention

Attention skills enable an inspector to fully inspect all potentially important regions of the inspection surface, not just the most central or most obvious [65]. Attention skills are responsible for distinguishing between environmental inputs that deserve attention and those that do not. Sensory inputs warranting attention during a bridge inspection may be visual, auditory, or tactile. Expectations, habituation, and stress can influence attention [65]. For instance, stimuli that are

repeated frequently gradually receive less attention and stress induced by physical or mental conditions can distract or divert attention. Specific details of the unit on attention are provided in Table 4.3. This unit involves an attention to detail exercise that challenges the inspector to identify the differences between two pictures of the same structure. Similar attention to detail can be useful in detecting small changes in condition, such as a missing rivet or bolt, among repeated bridge details. A selective attention exercise uses a video developed by Simons and Chabris [69] to demonstrate how the brain selectively attends to certain stimuli while neglecting other stimuli. This highlights the importance of maintaining an open mind about the types and locations of possible defects.

Table 4.3 Description of the attention unit

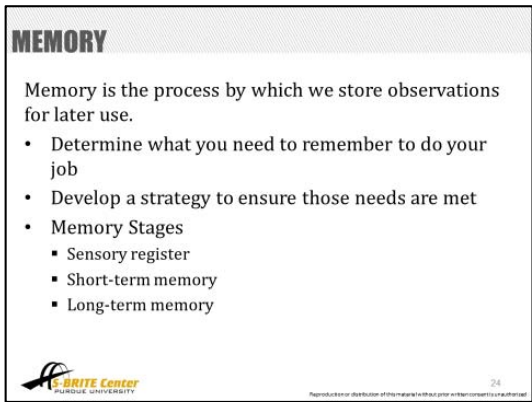
Objective	Convey how attention works and how it can be limited, especially during stressful or demanding situations
Topic(s)	<ol style="list-style-type: none"> 1. Definition of attention 2. Types of attention (focused, selective, divided)
Exercise(s)	<ol style="list-style-type: none"> 1. Attention to detail (identify the differences between two pictures) 2. The “gorilla” experiment (video demonstrating how the brain selectively attends to certain stimuli and completely misses other stimuli)
Sample Slide(s)	 <p>ATTENTION</p> <p>Attention will be focused on the locations where we expect to find the most useful information</p> <ul style="list-style-type: none"> • Identify inputs that deserve attention and inputs that distract • Attention is controlled by the observer • Information gathered from attentional processing is recorded in memory • Levels of attention <ul style="list-style-type: none"> ▪ Focused attention/vigilance (attention to detail) ▪ Selective attention ▪ Divided attention <p><small>S-BRITE Center PURDUE UNIVERSITY</small></p> <p><small>20</small></p>

4.2.4 Memory

Memory is the process by which information is stored for later use [65]. Memory skills allow an inspector to recall information about the structure, remember to do specific things during the inspection, recognize critical structural features and governing operating conditions, accurately record measurements, quantity estimates and observations, and remember to follow up on action items. Memory is also necessary for developing a knowledge base about a specific structure or

type of structure. This unit includes a memory training exercise known as dual n-back which has been found to improve short-term, or working, memory capacity [70]. This exercise challenges the inspector to retain and recall information delivered in different formats, such as auditory and visual. Similar memory skills may be necessary during a bridge inspection when an inspector is trying to remember and record both the location and length of a suspected crack. Specific details of the unit on memory are provided in Table 4.4.

Table 4.4 Description of the memory unit


Objective	Convey how memory works and its limitations; convey the importance of memory aids
Topic(s)	1. Definition of memory 2. Stages of memory (sensory register, short-term memory, long-term memory)
Exercise(s)	Dual n-back (short term memory exercise that requires the user to process both auditory and visual information and retain new information without replacing older information)
Sample Slide(s)	

4.2.5 Mental Images

Mental imaging is the perception of “remembered” information absent immediate sensory input [65]. Since mental images reflect perceptions, they will be flawed if the perceptions were flawed. Mental images can be used for remembering and organizing visual information, solving spatial problems, and interpreting observations. Mental images can be manipulated, transformed, or rotated to reveal the required information. Mental images may be useful in understanding and recalling how a bridge system responds to load. For instance, the mental images of a pin-type and roller-type bearing might help an inspector recall the type of support provided by each. Specific

details of the unit on mental imaging are provided in Table 4.5. This unit includes a mental rotation exercise challenging the inspector to mentally rotate a three-dimensional object to determine if it matches the baseline object [71]. Similar mental rotation skills allow bridge inspectors to align the drawings and details shown on construction plans with the structures they view in the field.

Table 4.5 Description of the mental images unit

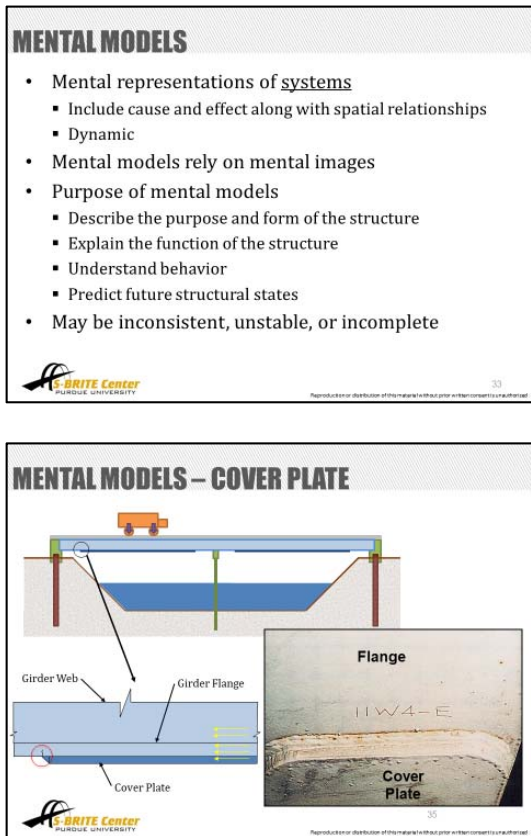
Objective	Convey what mental images are and how they can be used to organize, manipulate, and interpret spatial information during a bridge inspection
Topic(s)	1. Definition of mental images 2. Examples of mental images (pinned bearing vs. roller bearing)
Exercise(s)	Mental rotation (rotate 3D shapes until they match)
Sample Slide(s)	

4.2.6 Mental Models

Mental models are built on mental images, but include cause-and-effect relationships in addition to spatial relationships [65]. They are used to understand the purpose and form of the structure, explain its function, and predict future changes. Mental models can be updated based on observed

behavior in the field and refined with input from other inspectors and engineers. A mental model of a bridge structure would include both the structural members and their responses to applied loads. This model could be used to identify regions of the bridge that may be susceptible to cracking and which require special attention during an inspection. This unit includes an exercise challenging the inspector to use their mental models of a welded cover plate and transverse connection plate to explain why the details are susceptible to fatigue cracking, where cracks will occur, what the consequences of cracking may be, and possible repairs or retrofits. Specific details of the unit on mental models are provided in Table 4.6.

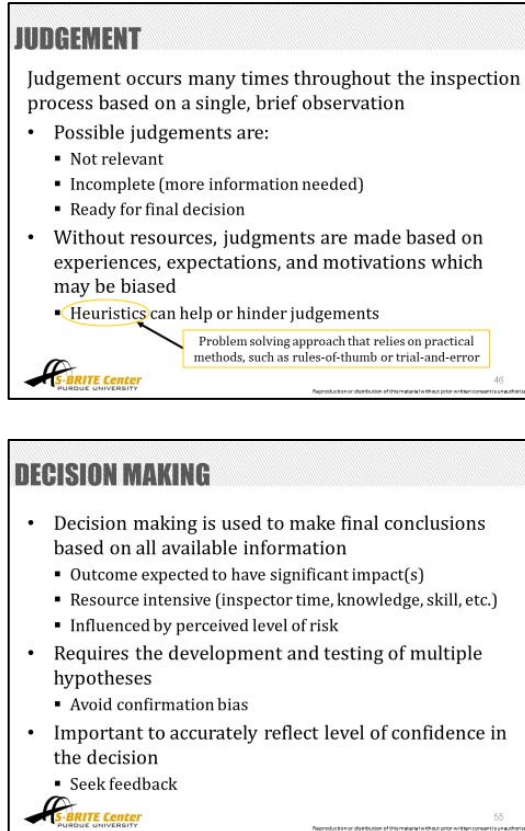
Table 4.6 Description of the mental models unit

Objective	Convey what mental models are, how they are related to mental images, and how they can be used to answer “what if” questions
Topic(s)	1. Definition of mental models 2. How to build mental models and extract information from them
Exercises	Applying mental models (connection plate detail, cover plate detail)
Sample Slides	 <p>The sample slides are as follows:</p> <p>MENTAL MODELS</p> <ul style="list-style-type: none"> • Mental representations of <u>systems</u> <ul style="list-style-type: none"> ▪ Include cause and effect along with spatial relationships ▪ Dynamic • Mental models rely on mental images • Purpose of mental models <ul style="list-style-type: none"> ▪ Describe the purpose and form of the structure ▪ Explain the function of the structure ▪ Understand behavior ▪ Predict future structural states • May be inconsistent, unstable, or incomplete <p>MENTAL MODELS – COVER PLATE</p> <p>The slide shows a cross-section of a bridge girder with a cover plate. Labels include: Girder Web, Girder Flange, and Cover Plate. A photograph of a physical cover plate is also shown, with labels for Flange and Cover Plate. The cover plate is labeled 11W4-E.</p>

4.2.7 Judgement and Decision Making

Judgement and decision making skills allow an inspector to apply logic and reasoning to observations to form conclusions about the condition of the structure [65]. In this training course, a judgement is defined as a quick evaluation made based on a single cue and a decision is defined as a final evaluation made based on a large number of cues from many different sources. The unit covers the fallibility of human judgement including mental shortcuts, or heuristics, which can lead to erroneous decisions when misapplied. For instance, the confirmation heuristic might lead an inspector that suspects a crack was caused by fatigue to take note of the cues that support this theory, such as the presence of a fatigue prone detail, while ignoring the cues that do not support this theory, such as the absence of paint or rust on the crack surface. Specific details of the judgement and decision making unit are provided in Table 4.7. This unit includes a series of word problems based on Bazerman and Moore [72] and Einhorn and Hogarth [73] intended to demonstrate how heuristics can cause errors in judgment.

Table 4.7 Description of the judgement and decision making unit

Objective	Convey how judgements and decisions are formed and how to avoid making common errors or miscalculations
Topic(s)	<ol style="list-style-type: none"> 1. Definition of judgement and decision making 2. How to identify and avoid common mistakes in judgement 3. Problem solving techniques
Exercise(s)	Judgment exercise (representativeness heuristic, availability heuristic, confirmation heuristic)
Sample Slide(s)	 <p>JUDGEMENT</p> <p>Judgement occurs many times throughout the inspection process based on a single, brief observation</p> <ul style="list-style-type: none"> • Possible judgements are: <ul style="list-style-type: none"> ▪ Not relevant ▪ Incomplete (more information needed) ▪ Ready for final decision • Without resources, judgments are made based on experiences, expectations, and motivations which may be biased • Heuristics can help or hinder judgements <p>Problem solving approach that relies on practical methods, such as rules-of-thumb or trial-and-error</p> <p>DECISION MAKING</p> <ul style="list-style-type: none"> • Decision making is used to make final conclusions based on all available information <ul style="list-style-type: none"> ▪ Outcome expected to have significant impact(s) ▪ Resource intensive (inspector time, knowledge, skill, etc.) ▪ Influenced by perceived level of risk • Requires the development and testing of multiple hypotheses <ul style="list-style-type: none"> ▪ Avoid confirmation bias • Important to accurately reflect level of confidence in the decision <ul style="list-style-type: none"> ▪ Seek feedback

4.3 Inspection Tools and Techniques Training

In speaking with inspectors during the study, it was clear that inspectors had very little exposure to actual fatigue cracks. A number of inspectors remarked that they found more cracks during this single day inspection than during their entire career, often spanning multiple decades. Consequently, a new training module was developed for inclusion in the *Inspecting Steel Bridges for Fatigue* course offered through Purdue University's S-BRITE Center and has been delivered three times (May 2017, April 2018, and October 2018). This module includes a brief introduction to the physical tools and techniques of visual inspection. It covers the steps for visual inspection,

useful inspection tools, and techniques to distinguish between paint scratches and fatigue cracks in steel bridge members. Additionally, inspectors have the opportunity to apply the training during a mock visual inspection of full size bridge specimens with actual fatigue cracks. Four sample slides from this training module are included in Figure 4.2.






<h3>STEPS FOR VISUAL INSPECTION</h3> <ol style="list-style-type: none"> 1. Visually assess the overall condition 2. Clean the surface 3. <u>Search</u> for indications of possible defects <ul style="list-style-type: none"> ▪ Use a consistent search strategy ▪ Focus on critical areas ▪ Look first with the naked eye only ▪ Then with the flashlight (ALWAYS!) ▪ Use other tools to improve visual acuity 4. <u>Decide</u> if indications are recordable defects <ul style="list-style-type: none"> ▪ Use tools to improve visual acuity 5. Record the length and location (if detected) <p> 2 <small>Reproduction or distribution of this material without prior written consent is unauthorized.</small></p>	<h3>TOOLS</h3> <ul style="list-style-type: none"> • Flashlight (and spare batteries) • Scraper, wire brush, broom for cleaning • Magnifying glass • Adjustable mirror • Measuring device • Dye penetrant kit  <p> 3 <small>Reproduction or distribution of this material without prior written consent is unauthorized.</small></p>
<h3>VISUAL INSPECTION TIPS</h3> <ul style="list-style-type: none"> • Touch the surface to feel changes in texture • Engage in active observation • View details from different angles and distances • Rotate the flashlight to vary the angle and intensity of illumination • Use compatible tools for enhanced visual acuity • Don't try to inspect while you are moving • Apply degreasing spray to see pulsing action • Don't rush; take breaks! <p> 11 <small>Reproduction or distribution of this material without prior written consent is unauthorized.</small></p>	<h3>CRACK VS. SCRATCH</h3> <p>To distinguish between a crack in the base metal and a scratch in the paint, consider:</p> <ul style="list-style-type: none"> • Location → What would cause the crack to form? • Size → Crack width will be very small (~0.006") Cracks will have a depth to them • Shape → Width/depth of scratches may change over the length Cracks will have a more well defined path • Light reflection → May see shadow or light reflection from crack • Other signs → Rust staining Changes in width/length with time <p>When in doubt, remove the paint to verify (be careful not to close the crack)</p> <p> 15 <small>Reproduction or distribution of this material without prior written consent is unauthorized.</small></p>

Figure 4.2 Select slides from the Inspection Tools and Techniques training module

4.4 Summary

The large variability in inspection performance indicates that the current training program produces inspectors with differing inspection abilities. Although the content of the existing courses appears to be adequate, it seems that some inspectors are struggling to apply the lessons taught in the classroom in the field. A new training course and a new training module were developed to address this deficiency. These trainings focus on the physical and mental factors of visual inspection and aim to teach inspectors “how to inspect” rather than “where to inspect”. The Observational Skills training course divides the visual inspection process down into four distinct subtasks and then identifies the cognitive skills used in each task. Interactive exercises are used

to demonstrate and strengthen observation skills. The Inspection Tools and Techniques training module describes the physical aspects of performing a steel bridge inspection, including effective use of inspection tools. Following the training, the inspectors have the opportunity to apply the discussed inspection techniques on full size bridge specimens with actual fatigue cracks.

5. UAS-ASSISTED VISUAL INSPECTION

5.1 Introduction

Recent advances in unmanned aircraft system (UAS) technology have made it a viable tool for state departments of transportation operating in a resource limited environment. Potential applications include infrastructure inspection, traffic surveillance, roadway assessment, accident reconstruction, and construction site monitoring [74]. For in-service bridge inspections, UAS assistance may reduce the need for extensive traffic control and specialized access equipment, thereby increasing the safety of both the inspectors and the motoring public. Additionally, the collection of high resolution video and imagery provide robust documentation of the bridge condition and allow for advanced analysis in the office. Before implementing this technology, it is critical to understand both its capabilities and limitations and identify the ideal role(s) for UAS in bridge inspection. Working with a team from Utah State University, the use of UAS for detecting fatigue cracks in steel bridge members was investigated through a series of real-time (field) and offline (desk) inspections.

The field inspections were performed by four certified bridge inspectors viewing live video from the UAS in the field and the desk inspections were performed by 19 certified bridge inspectors reviewing the recorded videos in their office. The inspection scenario and procedures followed those used during the hands-on inspections as closely as practical so that the results from the UAS inspections could be directly compared to the results from the hands-on inspections. Although many different levels of this technology exist, this project utilized an affordable, off-the-shelf UAS, such that it could be easily implemented at the state or local level.

The inspection set-up and results are discussed in detail in the following sections. The methods used to evaluate the data will be explained along with findings, recommendations, and conclusions regarding UAS-assisted inspection of steel bridge members.

5.2 Field Inspection

A small pool of bridge inspectors was invited to perform a UAS-assisted inspection of the same specimens used during the hands-on inspections discussed in Chapter 3. These inspections followed the same general process described for the hands-on inspections. Inspectors were directed to use the same level of care and attention to detail usually exhibited during a hands-on, fracture critical inspection. All of the inspectors worked with the same FAA certified pilot and the same UAS platform, the Mavic Pro from DJI. Due to the small sample size, only general conclusions can be made regarding the accuracy and consistency of UAS-assisted field inspections of steel bridge members. This program lays the groundwork for future research to develop inspection protocols and standards for UAS-assisted bridge inspections.

5.2.1 Inspection Scenario

The inspection procedures outlined in Section 3.2 for the hands-on inspections were followed as closely as practical for the UAS-assisted field inspections. The specimens were inspected in the field in real-time using a first person view monitor (tablet). The inspector was able to control the angle and exposure of the camera, while the pilot flew the UAS from specimen to specimen. A video of each specimen was recorded for use during the desk inspections. The specimens were inspected in the same predetermined order used during the hands-on inspections. A proctor from Purdue University was onsite throughout the entire inspection to ensure that the proper procedures were followed and to record inspection times, flight times, and general observations about the inspector and the inspection conditions. Figure 5.1 shows two inspectors performing the UAS-assisted field inspection.

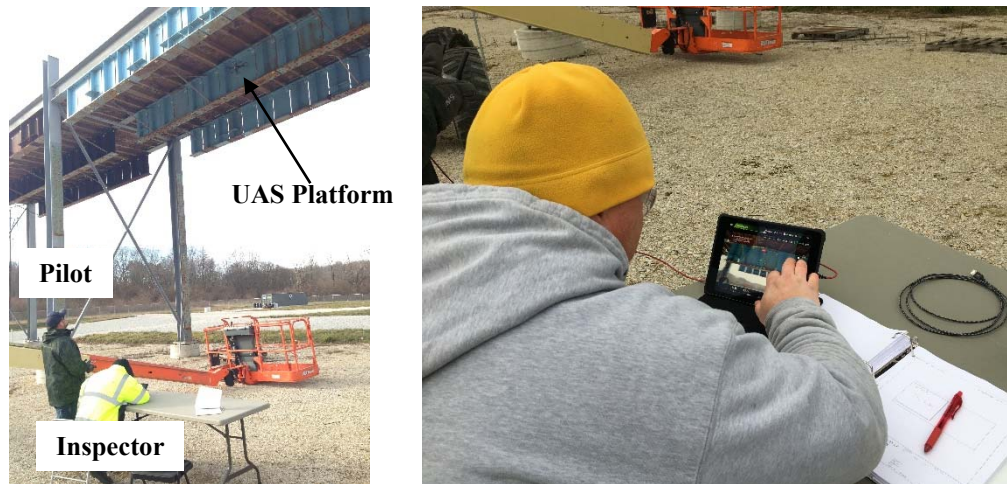


Figure 5.1 UAS-assisted field inspections of the POD specimens

On the day of the inspection, the inspector reported to the test site at approximately 8 AM. Before beginning the inspection, a brief tutorial was offered to introduce the inspectors to the mechanics of performing a bridge inspection using a UAS. This tutorial, lasting approximately 30 minutes, was performed using another bridge specimen with known fatigue cracks at the S-BRITE Center. Following the tutorial, a series of three vision tests were administered to measure normal visual acuity, near visual acuity, and contrast sensitivity. The normal visual acuity test was conducted with the Snellen eye chart, the near visual acuity test was administered with the Jaeger chart, and the contrast sensitivity test was performed with the Pelli-Robson chart. The test charts were mounted on the inspection frame as shown in Figure 5.2. The vision tests were administered per the documented procedures with the UAS hovering at the required standoff distance and the inspector verbally reciting the legible letters or words.

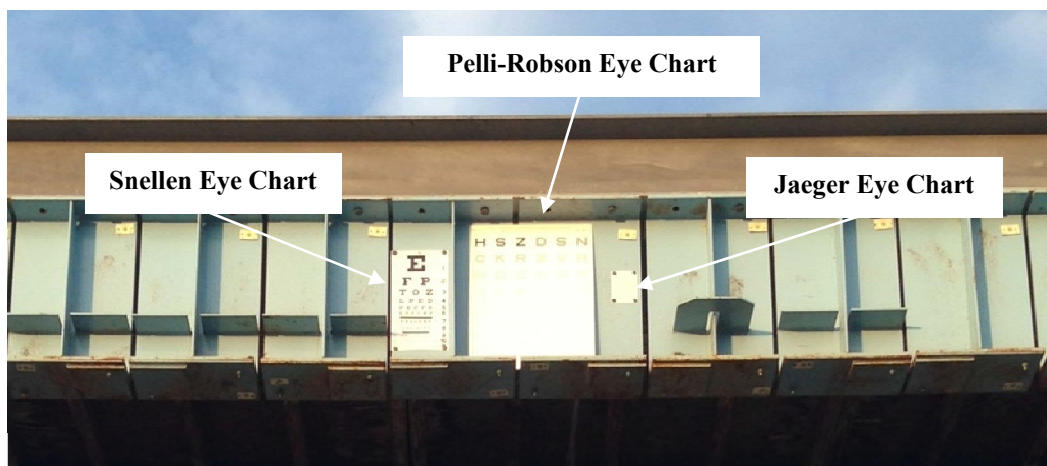


Figure 5.2 Vision test charts used during the UAS-assisted field inspections

Each inspector was given the inspection procedures, a list of assumptions about the specimens, and a blank inspection form for each specimen. Again, the inspectors were asked to record the length and location of any documented crack(s) or indicate that none were found. The inspections were self-paced, although external factors such as wind condition, battery life, and pilot fatigue necessitated frequent breaks.

Midway through the inspection day, both the inspector and the pilot were asked to complete a quantitative workload assessment administered using the NASA Task Load Index (NASA-TLX) application on a smartphone. Although originally developed for aviation tasks, the NASA-TLX has rapidly spread to a variety of industries and activities and become the standard for evaluating workload [75], [76]. The workload index is founded on six factors expected to contribute to the workload most people experience while performing most tasks. Three of the factors are related to the demands imposed on the subject (mental, physical, and temporal) and three of the factors are related to the interaction between the subject and the task (effort, frustration, and performance). The final workload index reflects the subject's perceived workload between 0 (no work) and 100 (maximum work) based on the relative contribution of each factor. While no common standard exists to categorize workload (e.g. easy, hard, *too hard*), this index provides a relative comparison across subjects performing the same task.

At the conclusion of the exercise, the inspector completed an exit survey providing information about their background and training history. The exit survey also asked for the inspector's assessment of the inspection tutorial and scenario and general impression of the quality of the UAS-assisted field inspection. The inspection documents used during the UAS-assisted field inspections are similar to those provided in Appendix A for the hands-on inspection.

The test articles for the UAS-assisted field inspections were the same as those used during the hands-on inspections; however, no inspector was able to inspect more than 85 of the 147 specimens during their UAS-assisted field inspection. All specimen types were included, although none of the overhead mounted riveted plate specimens were inspected. The UAS-assisted inspections used only the painted specimens.

Inspections were conducted outdoors on the campus of Purdue University in West Lafayette, Indiana at the S-BRITE Center between 18 December 2017 and 21 December 2017. The inspections were completed from the ground adjacent to and below the probability of detection training structure. Temperature and wind speed statistics were recorded by the KLAF weather station at the Purdue University Airport, located adjacent to the test site [56].

5.2.2 Inspector Demographics

Four inspectors (3 males and 1 female) with experience in bridge inspection participated in the UAS-assisted field inspections. Two of the inspectors worked for a state department of transportation and two worked for a private engineering and inspection firm. The inspectors worked primarily in three states, Indiana (1), Illinois (1), and Idaho (2), although the private consultants inspect bridges around the country. All of the inspectors had completed both the FHWA/NHI *Safety Inspection of In-Service Bridges* course and the *Fracture Critical Bridge Inspection* course prior to their participation in the study. The average experience of the participating inspectors was 11.1 years and the inspectors had completed an average of 15 hands-on inspections in the 12 months prior to their participation. Three of the four inspectors possessed a professional engineering license and post-secondary degree in civil engineering. Vision tests were administered to three of the four participants. None of the inspectors recorded normal (20/20) vision on the Snellen vision test or read the smallest paragraph of text on the test card for the Jaeger vision test. The average Log Contrast Sensitivity score was 1.6. Select inspector demographics are compiled in Table 5.1.

One inspector participated in both the UAS-assisted field inspection and the hands-on inspection. Due to the large number of similarly featured specimens and the passage of time, participation in the hands-on inspections was not expected to affect performance in the UAS-assisted field inspection.

Table 5.1 Inspector demographics for UAS-assisted field inspections

Inspector ID	Employer	Age (years)	Inspection Experience (years)	Professional Licensure	Number of Hands-on (Routine) Inspections	Log Contrast Sensitivity
20LC-91	Private Consultant	31	7	PE	10 (35)	1.5
19JK-92	State DOT	50	16	None	15 (340)	1.65
21UT-98	State DOT	35	5	PE	8 (25)	1.65
18YW-96	Private Consultant	47	16.5	PE	25 (800)	N/A

5.2.3 Inspection Results

The information provided on each inspection form was interpreted and categorized as outlined in Section 3.4. Since each inspector inspected a different number of specimens, detection rate was calculated based on the total number of cracks in the inspected specimens. The results from the UAS-assisted field inspections were compiled and evaluated similar to the hands-on inspection results. Inspector performance was evaluated for the entire specimen inventory and for each type of specimen (girder, welded cover plate, riveted plate) separately.

5.2.3.1 Crack Detection and False Calls

Four inspectors completed a UAS-assisted field inspection of the painted specimens. The most successful inspector achieved a detection rate of 60% (12 of 20 possible cracks) while the least successful inspector recorded a detection rate of 48% (20 of 42 possible cracks). The average detection rate was 54% and the standard deviation of detection rate was 6%. The number of false calls made during the inspections ranged from 26 to 67. The average number of false calls was 53 with a standard deviation of 19. The inspection results for each inspector are presented in Table 5.2. Note that the number of specimens inspected, and therefore the total number of possible cracks, varies among the inspectors and was lower than during the hands-on inspections.

Table 5.2. UAS-assisted field inspection results by inspector

Inspector ID	Number of Specimens Inspected (Possible Cracks)	Hits	Detection Rate	False Positives	Hit/Call Ratio
20LC-91	85 (42)	20	48%	54	27%
19JK-92	78 (38)	22	58%	26	46%
21UT-98	52 (20)	12	60%	66	15%
18YW-96	52 (20)	10	50%	67	13%

The number of false positives were compared to the number of hits for each inspector and the hit/call ratio was calculated. The highest, or best, hit/call ratio was 46% and the lowest, or worst, hit/call ratio was 13%. The average hit/call ratio was 25%, meaning that one hit was recorded for every 4 calls. As illustrated in Figure 5.3 where the inspectors are plotted in order of increasing number of hits, the number of false calls (negative y-axis) is negatively correlated with the number of hits (positive y-axis). In other words, inspectors that inspected more specimens and detected more cracks actually made fewer false calls.

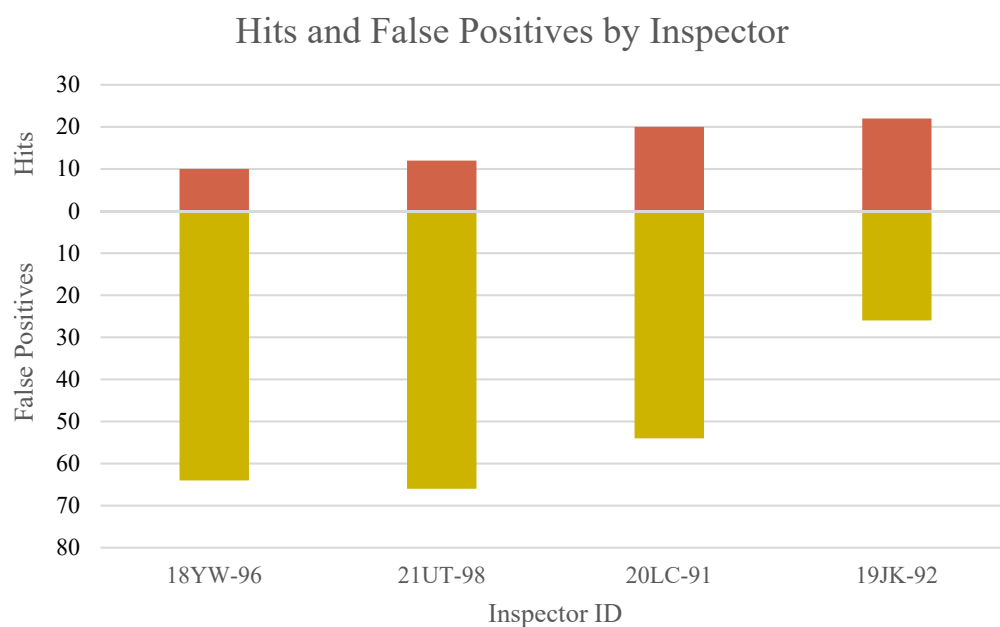


Figure 5.3 Hits and false calls by inspector during the UAS-assisted field inspections

The detection rates by crack type were also examined as shown in Figure 5.4. These inspections included between 15 and 24 out-of-plane cracks in the girder specimens, between 5 and 9 weld toe cracks in the welded cover plate specimens, and between 0 and 9 rivet hole cracks in the riveted plate specimens. Note that two inspectors, 18YW-96 and 21UT-98, did not inspect the riveted plate specimens and so the detection rate for the rivet hole cracks is labelled as “N/A”. In contrast, two inspectors, 18YW-96 and 19JK-92, inspected the welded cover plate specimens but did not detect any of the cracks and so their detection rates are labelled as 0%. The weld toe cracks had the lowest average detection rate. The camera used during these inspections was mounted beneath the body of the UAS making it difficult to inspect directly overhead. By tilting the UAS platform,

the inspectors could see the cover plate termination, but it was challenging to get a clear image of the weld.

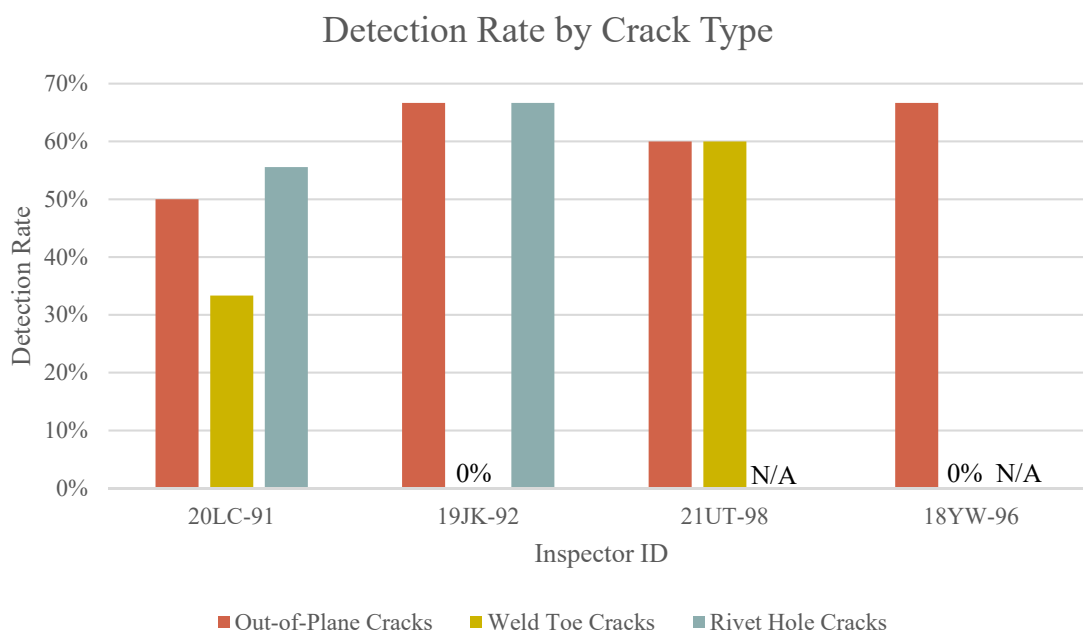


Figure 5.4 Inspector detection rates by crack type

5.2.3.2 Crack Sizing

Similar to the hands-on inspections, the inspectors participating in the UAS-assisted field inspections were asked to record both the length and location of the cracks on their inspection forms. Inspectors were provided with dimensioned drawings of the specimens to provide a reference against which to estimate crack length. Three of the four field inspectors recorded crack length estimates on their inspection forms.

Figure 5.5 shows the crack length data for the out-of-plane cracks in the girder specimens. The actual length of the crack is shown on the horizontal axis and the estimated length of the crack recorded by the inspector is shown on the vertical axis. The diagonal 1:1 reference line represents exact agreement between the actual length and the estimated length. For the majority of the cracks, the average of the estimated lengths plots below the 1:1 line indicating that the inspectors tended to underestimate crack length. This trend was also observed in the results from the hands-on

inspections. The average measurement error was -0.84 inch and the average absolute error was 1.02 inches. The average absolute error increased with crack length and the percent absolute error decreased with crack length. Table 5.3 shows the error analysis for the reported length measurement for all the cracks in the girder specimens.

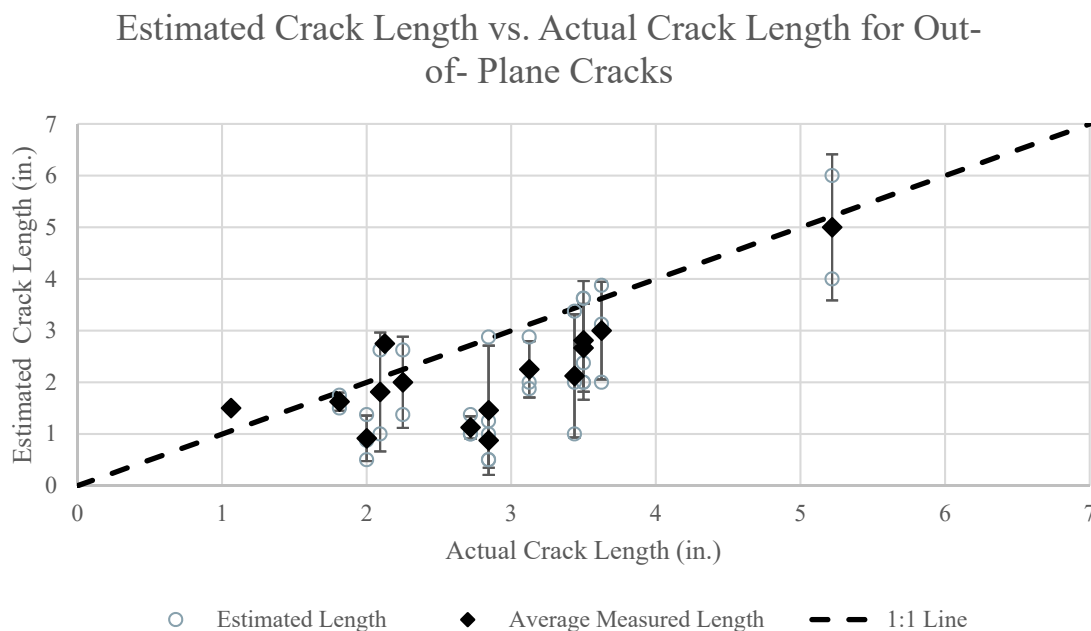


Figure 5.5 Estimated crack length versus actual crack length for out-of-plane cracks

Figure 5.6 and Figure 5.7 present the crack length data for the welded cover plate and riveted plate specimens, respectively. The crack sizing information is limited for these specimens due to the small number of inspectors and low detection rates. For most of the cracks, only one inspector provided a length estimate and they tended to overestimate the length of the weld toe cracks and underestimate the length of the rivet hole cracks. For the weld toe cracks, the average measurement error was 0.71 inch and the average absolute error was 1.29 inches. For the rivet hole cracks, the average measurement error was -0.1 inch and the average absolute error was 0.15 inch. Table 5.3 shows the error analysis for the reported length measurement for all the cracks in the welded cover plate and riveted plate specimens.

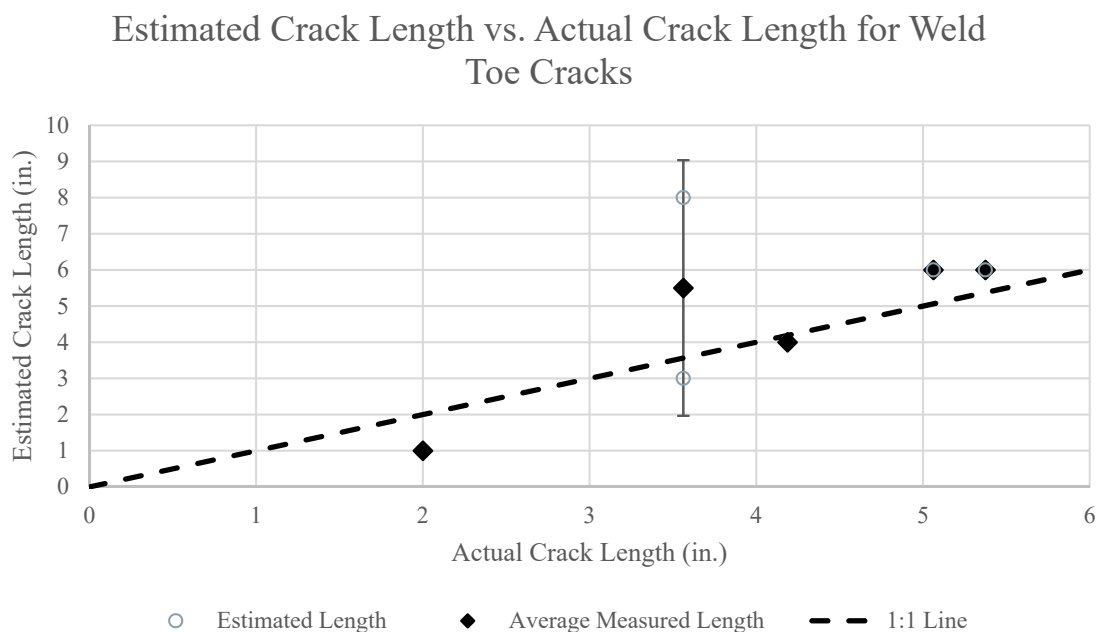


Figure 5.6 Estimated crack length versus actual crack length for weld toe cracks

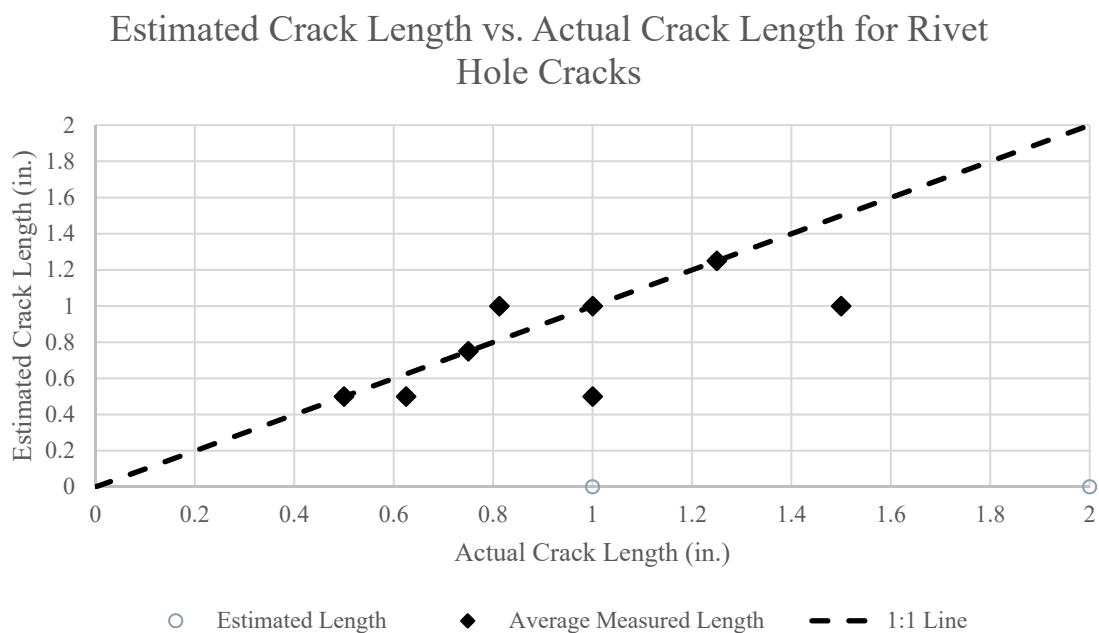


Figure 5.7 Estimated crack length versus actual crack length for rivet hole cracks

Table 5.3 Error analysis for crack length estimates from the UAS-assisted field inspections

Statistic	Out-of-Plane Cracks (24 cracks)		Weld Toe Cracks (9 cracks)		Rivet Hole Cracks (9 cracks)	
	Error (actual length – meas. length)	Absolute Error	Error (actual length – meas. length)	Absolute Error	Error (actual length – meas. length)	Absolute Error
Min./Max. (in.)	-2.44/0.78	-	-1/4.44	-	-0.5/0.19	
Average (in.)	-0.84	1.02	0.71	1.29	-0.10	0.15
St. Dev. (in.)	0.92	0.70	1.96	1.57	0.24	0.21
Average %	-28	36	14	37	-9	14

5.2.3.3 Inspection Duration

Both the field time and the inspection time were recorded during the UAS-assisted visual inspections. The field time is the total time spent in the field on the day of the inspection, excluding the time spent completing the pre-inspection tutorial, vision tests, and the post-inspection paperwork. The inspection time is the length of time the inspector was actually looking at the specimens and is approximately equal to the flight time. The difference between the field time and the inspection time represents the time spent changing batteries, waiting for flight conditions to improve, chatting, etc. and so the ratio of field time to inspection time (Time Ratio) is essentially an expression of efficiency. As this ratio increase from one, it indicates that less of the field time was spent inspecting the specimens. A summary of the field and inspection times are presented in Table 5.4. The inspection time varied between 125 and 201 minutes while field time varied from 244 minutes to 375 minutes. The ratio of field time to inspection time varied from 1.57 to 2.62. Since each inspector inspected a different number of specimens, it is also useful to consider inspection rate (inspection time divided by number of specimens inspected). Inspection rate varied from 1.71 minutes per specimen to 3.87 minutes per specimen.

Table 5.4 UAS-assisted field inspection durations by inspector

Inspector ID	Number of Specimens Inspected	Inspection Time (min.)	Field Time (min.)	Inspection Rate (min./specimen)	Time Ratio (Field Time/ Inspection Time)
20LC-91	85	145	244	1.71	1.68
19JK-92	78	199	375	2.55	1.88
21UT-98	52	201	315	3.87	1.57
18YW-96	52	125	328	2.40	2.62

5.2.3.4 NASA Task Load Index

Both the pilot and the inspector completed the NASA-TLX assessment. The workload scores and governing factors are shown in Table 5.5. The inspectors' workload scores ranged from 47 to 82 and frustration was the governing source of workload for three of the four inspectors. The pilot's workload scores varied between 62 and 72 with mental demand as the governing source of workload. These assessments agree with previous research that found mental demand and frustration to be the primary components of workload during vigilance tasks [77].

Table 5.5 NASA-TLX workload scores for the inspectors and pilot

Inspector ID	Inspector		Pilot	
	Workload Score	Governing Factor	Workload Score	Governing Factor
20LC-91	47	Mental Demand	62	Mental Demand
19JK-92	82	Frustration	68	Mental Demand
21UT-98	73	Frustration	72	Mental Demand
18YW-96	69	Frustration	66	Mental Demand

5.2.4 Crack Length Analysis

To investigate the relationship between crack detection and crack length, an abbreviated crack length analysis was performed with the results from the UAS-assisted field inspections. Due to the reduced number of cracks and inspectors, probability of detection curves were not developed. However, a general understanding of the relationship can be obtained by looking at the detection rate versus crack length bar graph.

The cracks were grouped into 1-inch length increments and the detection rate for each bin was computed. The detection rate was determined by summing the number of hits for each crack in the increment and dividing by the total number of detection attempts made. Although there is not a continuous increase in detection rate with increasing crack length, there is a clear improvement in detection for cracks over 2 inches in length as compared to cracks less than 2 inches in length. However, cracks greater than 4 inches were detected at a lower rate (50%) than cracks between 2 inches and 4 inches long (71%). Figure 5.8 shows the detection rates for each crack length increment. The number of cracks in each bin are shown on the right hand vertical axis.

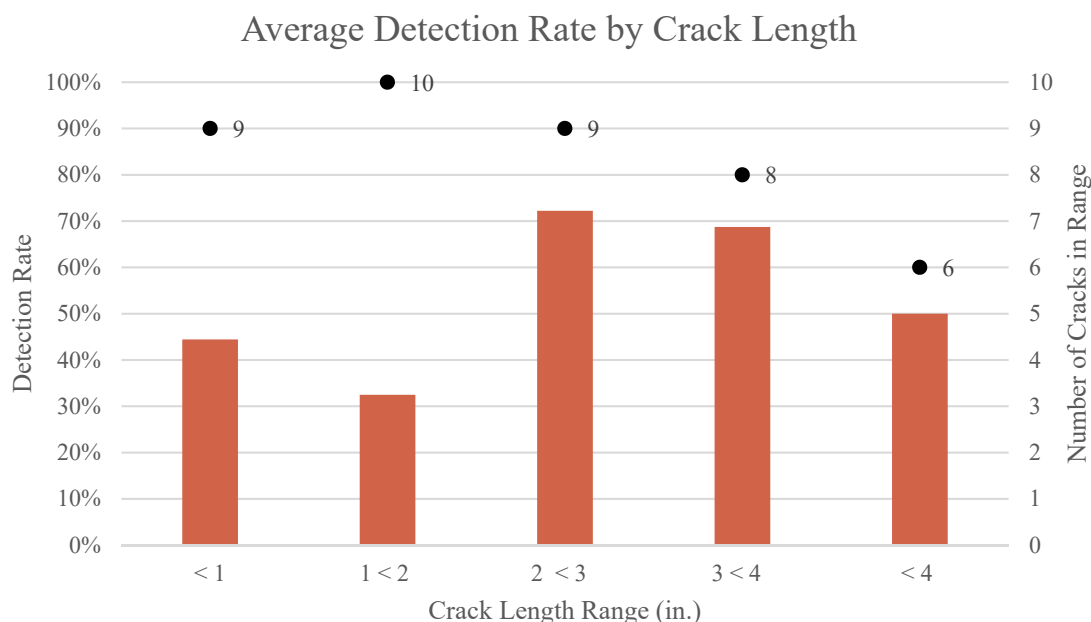


Figure 5.8 Detection rate by crack length

5.2.5 Human and Environmental Factors

Due to the limited sample size, an extensive statistical analysis of the human and environmental factors affecting inspection performance was not warranted for the UAS-assisted field inspections. Additionally, since the UAS is simply a tool used by the inspector during a visual inspection, many of the factors discussed in Section 3.7 are expected to similarly influence performance during a UAS-assisted inspection. Only the effects of wind speed were investigated since this was likely to have a more pronounced influence on performance during a UAS-assisted inspection as compared to a traditional, hands-on inspection.

While weather is an important consideration in any outdoor inspection activity, wind speed and direction has been found to be an especially critical factor in UAS-assisted structural inspections [42], [47], [78]. In particular, frequent fluctuations in wind speed and direction necessitate an increased standoff distance from the inspection surface as the likelihood of a collision increases. The average, maximum, and gust wind speeds recorded at the KLAf weather station are shown in Table 5.6. The average wind speed during the inspections ranged from 6 to 9 mph, with sustained winds up to 17 mph and gust speeds of 24 mph.

Table 5.6 Wind speeds by inspection day

Inspector ID	Average Wind Speed (mph)	Maximum Wind Speed (mph)	Gust Wind Speed (mph)
20LC-91	6	14	17
19JK-92	7	16	20
21UT-98	8	14	19
18YW-96	9	17	24

Detection rate, the number of false calls, and the ratio between field time and inspection time plotted against wind speed are shown Figure 5.9, Figure 5.10, and Figure 5.11, respectively. While no clear relationship was observed between inspection performance and wind speed, higher wind speeds were found to increase the ratio between field time and inspection time. In other words, under higher wind speeds, a reduced percentage of the field time was dedicated to inspection. This relationship was observed in the field during the inspections as increased winds, especially gusting winds, caused frequent disruptions because the pilot was unable to comfortably control the UAS within adequate proximity to the inspection specimens.

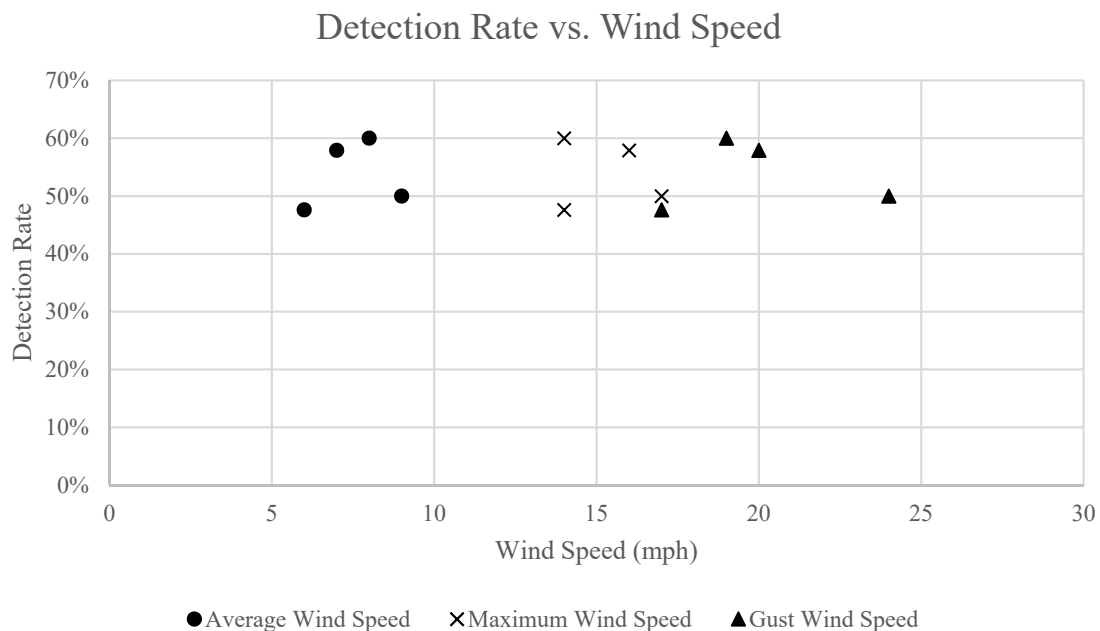


Figure 5.9 Detection rate plotted against wind speed

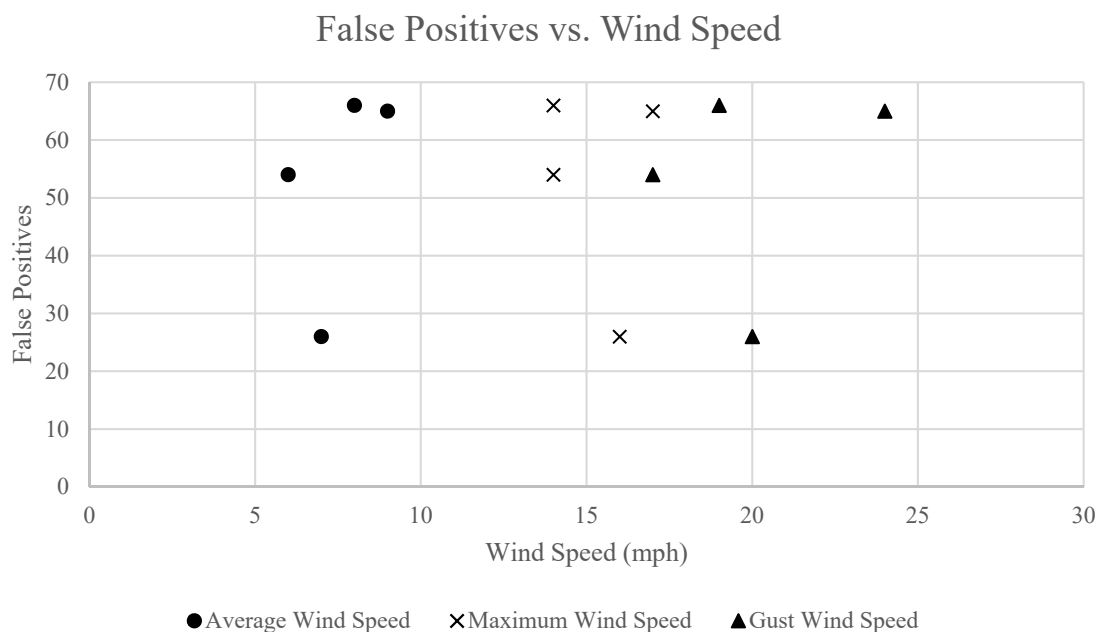


Figure 5.10 False positives plotted against wind speed

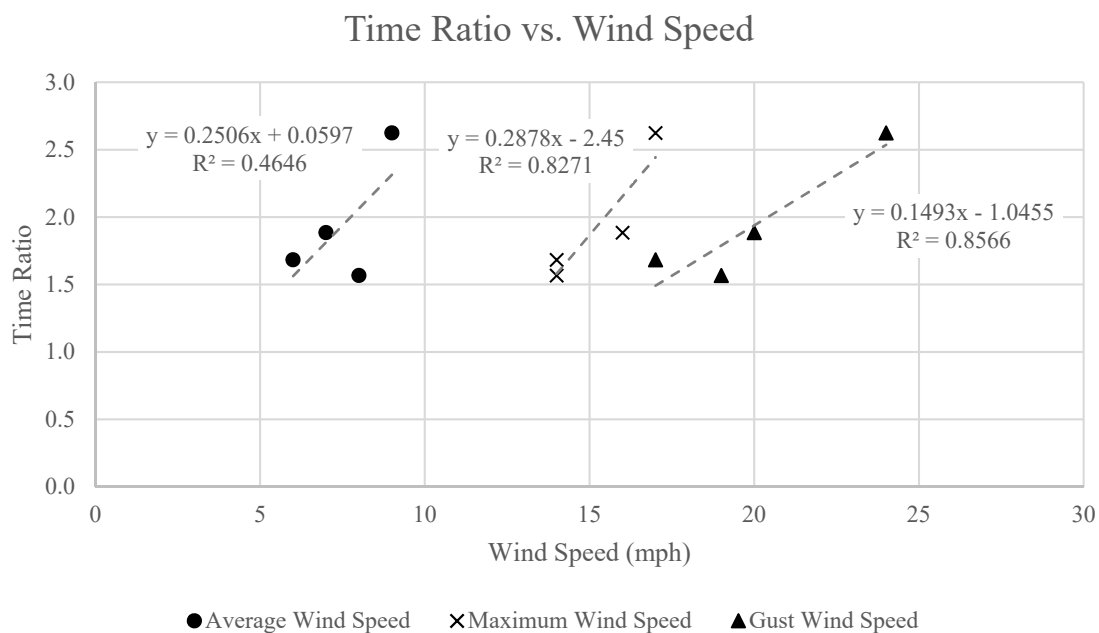


Figure 5.11 Time ratio (field time/inspection time) plotted against wind speed

5.2.6 Comparison to Hands-on Inspection

Although the sample size is limited, the results presented in Section 5.2.2 suggest a similar level of variability in performance observed during the hands-on inspections. Since these inspections were performed on consecutive days in similar weather conditions with the same pilot and UAS, inspector-to-inspector variation continues to be a significant factor in determining performance.

Table 5.7 provides a brief comparison between the results from the hands-on inspections and the UAS-assisted field inspections. This comparison is limited to the specimens that were inspected by all four inspectors during the UAS-assisted field inspections. For the girder specimens, the average performance during the UAS-assisted field inspections compared well with the hands-on inspections. The average detection rate was the same with a slight increase in the number of false calls during the UAS-assisted field inspections. Conversely, performance on the welded cover plate specimens was worse during the UAS-assisted field inspections. As mentioned previously, the location of the camera on the UAS made overhead inspection challenging. The average detection rate was lower and the number of false calls was higher during the UAS-assisted field inspections causing a reduction in the average hit/call ratio. The average inspection times were slightly longer for the UAS-assisted field inspections as compared to the hands-on inspections, although the more noticeable difference was observed in the ratio of field time to inspection time. During the hands-on inspections, 147 specimens were inspected during the full day inspection. In comparison, during the full day UAS assisted inspections, no inspector was able to inspect more than 85 specimens. In terms of crack sizing, the average error and average absolute error was greater during the UAS-assisted field inspections as compared to the hands-on inspections. Finally, separate from the numeric results, all of the inspectors participating in the UAS assisted inspections stated that they thought that “[the UAS assisted] inspection provided **worse** quality as compared to an arm’s length inspection” indicating that the lack of inspector confidence may pose a significant challenge to implementation.

Table 5.7 Average performance during UAS-assisted field and hands-on inspections

	Average from UAS-assisted field inspections - 4 inspectors (max/min in parentheses)	Average from hands-on inspections - 30 inspectors (max/min in parentheses)
Girder Specimens (36 specimens)		
Hits	10 (11/9)	10 (13/5)
False Positives	34 (50/10)	31 (94/1)
Total Cracks	15	15
Detection Rate (%)	65 (73/60)	65 (87/33)
Hit/Call Ratio (%)	27 (52/15)	33 (89/5)
Inspection time (min.)	111 (169/67)	86 (197/41)
Average Crack Sizing Error	-0.98 (0.78/-2.44)	-0.63 (7.38/-4.02)
Average Crack Sizing Absolute Error	1.1	0.93
Welded Cover Plates Specimens (16 specimens)		
Hits	1 (3/0)	3 (4/0)
False Positives	13 (19/8)	8 (20/0)
Total Cracks	5	5
Detection Rate (%)	20 (60/0)	57 (80/0)
Hit/Call Ratio (%)	7 (16/0)	40 (100/0)
Inspection time (min.)	29 (41/19)	20 (32/10)
Average Crack Sizing Error	0.67 (4.44/-1)	0.4 (8.44/-1.25)
Average Crack Sizing Absolute Error	1.55	0.72

A visual comparison of detection rate and false positives is provided in Figure 5.12 and Figure 5.13, respectively. Although the error bars are wide for the hands-on inspections, there was a statistically significant difference in detection rate between the two types of inspection. The average detection rate during the UAS-assisted field inspections was approximately one standard deviation below the average detection rate during the hands-on inspections and the two sample t -test revealed that significantly more cracks were detected during the hands-on inspections ($M = 0.699$, $SD = 0.139$) than during the UAS-assisted field inspections ($M = 0.538$, $SD = 0.048$), $t(11) = 4.62$, $p = 0.001$. In contrast, there was no statistically significant difference in the number of false calls made during the two types of inspection. Visually, the large overlap in the error bars in Figure 5.13 suggests this, and it was confirmed using the two-sample t -test.

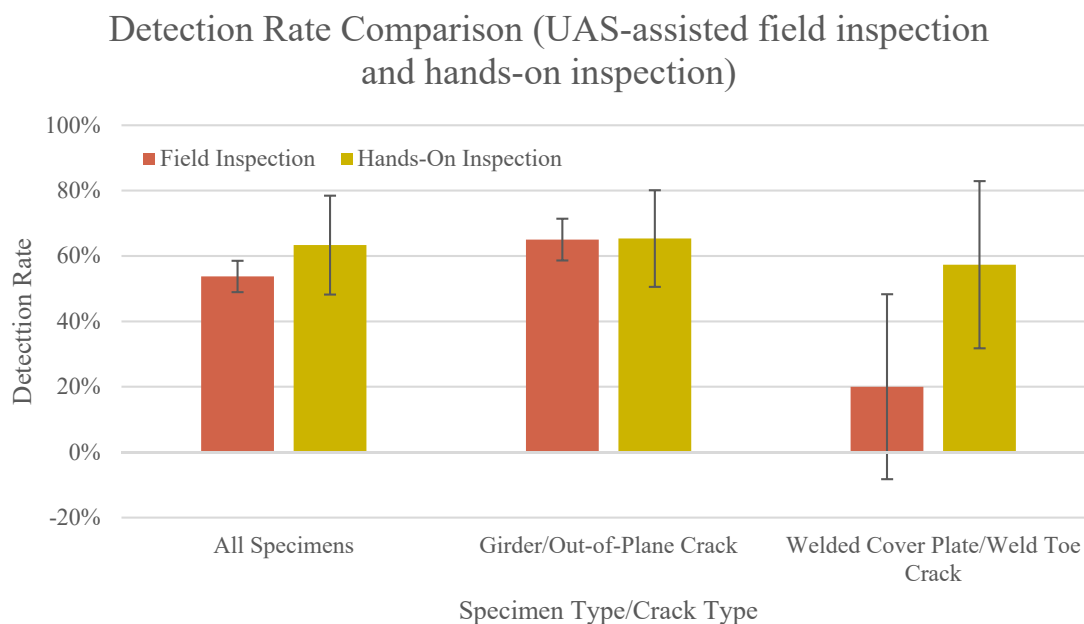


Figure 5.12 Comparison of average detection rate during the UAS-assisted field and hands-on inspections

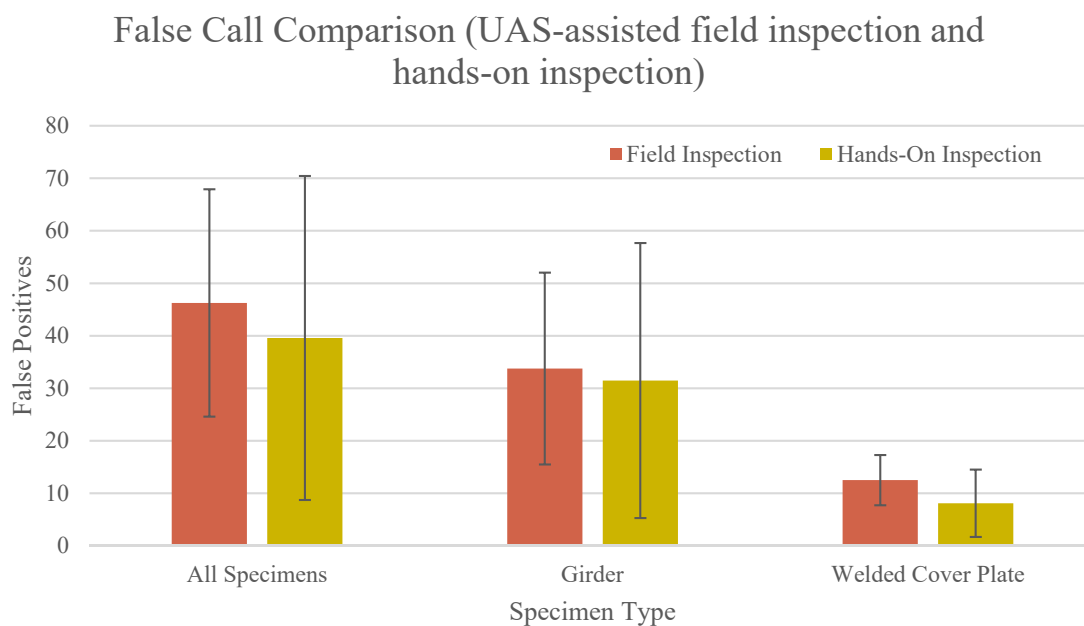


Figure 5.13 Comparison of average number of false calls made during the UAS-assisted field and hands-on inspections

As noted previously, one inspector performed both a UAS-assisted field inspection and a hands-on inspection. The results from the two inspections are presented in Table 5.8 along with the difference (hands-on inspection minus field inspection) between the two inspections. This comparison is limited to the specimens inspected during the UAS-assisted field inspection. The inspector located seven fewer cracks during the UAS-assisted inspection, but also made 18 fewer false calls. Still, the hit/call ratio during the UAS-assisted field inspection was 4% lower than during the hands-on inspection. Finally, although the inspection rate was faster during the UAS-assisted inspection, the ratio of field time to inspection time was greater. During the hands-on inspection, the inspector looked at all 147 specimens between 9:06 AM and 4:04 PM, even with a 31-minute rain delay. In contrast, during the UAS-assisted inspection, the inspector was only able to look at 52 specimens between 9:52 AM and 4:16 PM. There are a variety of explanations for the reduced inspection efficiency, including a longer lunch break and extended conversations with the research team, but the nearly constant need to change batteries in the UAS, controller, and tablet, difficulties flying in a GPS-denied environment under the timber deck, and frequent winds over 10 mph all caused significant delays that were not present during the hands-on inspection.

Table 5.8 Difference in inspection performance during UAS-assisted field and hands-on inspections for a single inspector

	Field Inspection	Hands-on Inspection	Difference
Girder Specimens (36 specimens)			
Hits	13	10	3
False Positives	67	46	21
Detection Rate (%)	87%	67%	20%
Hit/Call Ratio (%)	16%	18%	-2%
Inspection time (min.)	173	102	71
Average Crack Sizing Error	-0.60	-	-
Average Crack Sizing Absolute Error	0.71	-	-
Welded Cover Plates Specimens (16 specimens)			
Hits	4	0	4
False Positives	16	19	-3
Detection Rate (%)	80%	0%	80%
Hit/Call Ratio (%)	20%	0%	20%
Inspection time (min.)	17	23	-6
Average Crack Sizing Error	0.31	-	-
Average Crack Sizing Absolute Error	0.31	-	-

5.2.7 Discussion

Although limited in scale and scope, several of the findings from the UAS-assisted field inspections enhance the current understanding of the capabilities and limitations of this technology.

First, advocates often cite reduced inspection time as one of the expected benefits of performing civil infrastructure inspections with UAS assistance. Albeit a very small sample, this set of trials contradicts this claim. As discussed previously, none of the UAS-assisted field inspectors were able to inspect all 147 specimens during their full day inspection. In practice, a hands-on inspection would require additional time for the set-up and removal of traffic control and deployment of access equipment making the two methods more comparable in duration. Even so, it seems that one, heretofore overlooked, consideration in the planning of UAS inspections is the ratio of field time to inspection time. Although nearly continuous flight was possible due to the proximity of a reliable power source and a large store of batteries, the UAS-assisted inspection still experienced frequent pauses to change batteries, avoid swirling winds, restore GPS signal, etc. Over the four days of inspection, the maximum flight time was 18 minutes but the average flight time was about five minutes. Each interruption required the inspector and pilot to switch their focus and they were observed taking many more breaks during the UAS-assisted inspections as compared to the hands-on inspections. Some of these breaks lasted a number of minutes and extended long after the wind had settled or the battery had been changed. While frequent breaks are recommended by visual inspection research to combat the vigilance decrement, in some cases, this may reduce the time savings expected with UAS assistance.

Second, much of the current research includes only cursory mentions of the pilot's role within the inspection team. However, a successful in-depth inspection relies heavily on the pilot's comfort flying in close proximity to the structure. Therefore, it is critical that the pilot is provided with adequate tools and resources to successfully complete the flight under minimal stress. This includes training specific to bridge inspection and sufficient practice flying in confined spaces with and without GPS assistance. Additional observations from these field trials highlight the importance of understanding both the capabilities and the limitations of the pilot. Over the course of the four days, the pilot grew more comfortable and confident piloting the UAS platform in close proximity to the inspection surface. He verbally expressed this feeling and it was noticeable to the

proctor. The ratio of field time to inspection time decreased each day, from 2.62 on Day 1 to 1.57 on Day 4, in part due to the pilot's increasing ability and willingness to fly under challenging conditions. Additionally, the pilot gradually became more familiar with the inspectors' jargon and was able to direct the UAS platform to critical areas with less intervention or direction from the inspector. Even so, this pilot appreciated near constant feedback and dialogue with the inspector and expressed frustration while working with less communicative inspectors. Finally, as UAS technology evolves, longer flight times will undoubtedly become possible. However, for close range inspection missions, pilot fatigue may still limit flight times to around 15 minutes. These missions require intense concentration from the pilot and after about 15 to 20 minutes, the pilot needed a break to physically and mentally reset.

Third, no published guidance currently exists outlining the minimum system requirements for performing visual bridge inspections with the assistance of a UAS. Instead, each user is forced to develop their own set of considerations and criteria for UAS selection. These considerations may include unit size, battery life, navigation capabilities, imaging capabilities and user interface. Since none of these factors operate in isolation (e.g., a bigger and better imaging system will increase payload weight and reduce flight time), it is important to consider the trade-offs among them. The following discussion addresses the performance of the DJI Mavic Pro under each of the five considerations listed.

- (1) Unit Size: The size of the DJI Mavic Pro (14" diagonal without propellers) was satisfactory to perform the mission. The UAS was able to navigate between adjacent girder lines spaced at approximately 8 feet on center.
- (2) Battery Life: The battery life of the DJI Mavic Pro was satisfactory to perform the mission. Cited as 27 minutes [79], average battery life during the inspections was closer to 18 minutes. Additionally, six back-up batteries were on hand to ensure that a fully charged battery was always available. One note, while the battery life of the UAS was sufficient and did not cause major disruptions to the schedule, the battery life of the controllers and tablets was limited to about four hours and these devices did require charging during each full day inspection. However, battery life was only one of many factors dictating flight times during these field trials, and typically it was not the most limiting. When considering flight time, it is important to consider other factors such as wind conditions, UAS stability

and maneuverability around the bridge structure, line-of-sight from the pilot/observer to the UAS platform, and pilot stamina. The short flight times, five minutes on average, greatly reduced the efficiency of these inspections.

- (3) Navigation Capabilities: The built-in navigation system in the DJI Mavic Pro was minimally acceptable to perform the mission. Under normal conditions, the Mavic Pro relies on GPS and forward and downward facing vision systems to locate itself, hold a stable position while hovering, and avoid obstacles [79]. Without GPS, the Mavic uses a barometer to maintain vertical position and the forward and downward facing vision systems to maintain horizontal position. Both the forward and downward facing systems include two monocular sensors while the downward facing system also includes two ultrasonic sensors. For this in-depth inspection, the obstacle avoidance technology included with the forward facing sensors was deactivated to allow for a shorter standoff distance, approximately 2 to 3 feet, from the specimens. This increases the responsibility of the pilot to prevent collisions. (Note, one collision did occur during the field trials when the propeller struck an overhead bolt, but the pilot was able to safely land the UAS and no damage was incurred.) In the GPS-denied environment underneath the bridge deck, the downward facing sensors were generally adequate to maintain a stable position. At times, these sensors were unable to derive enough information from the gravel bed beneath the structure and the UAS had to be flown out into the open space to regain control.
- (4) Imaging Capabilities: The imaging capabilities of the DJI Mavic Pro were minimally acceptable to perform the mission. Since the Mavic Pro cannot capture still images and videos concurrently, only the video function was used during the field trials. The built-in camera is capable of capturing 3840 x 2160 pixels (4k) video, although this video is only streamed live at 1920 x 1080 pixels (1080p) [79]. The camera is mounted on a 3-axis gimbal beneath the body of the UAS. The location of the camera restricted the vertical field of view to 120 degrees and made overhead inspection difficult, but not impossible. Additionally, the camera could not rotate horizontally, and so the entire UAS platform had to be rotated by the pilot to adjust the horizontal field of view. Since the Mavic Pro includes only digital zoom, it was not possible to zoom in more while recording video with 4k resolution. Therefore, a short standoff distance was required to acquire detailed imagery. Finally, although the DJI Mavic Pro does not include externally mounted lights to improve

illumination, the shutter and ISO speeds can be manually controlled to improve the image quality even in poor lighting. This was generally adequate, although both the pilot and the inspector struggled in extreme light conditions (very dark or very bright).

- (5) User Interface: The DJI Mavic Pro's user interface was satisfactory to perform the mission. The pilot used the manufacturer-provided controller and the inspector viewed the live video on a 9.7-inch tablet. Both the pilot and the inspector could control the camera. The only downside was that the controller and the tablet were connected by a USB cord essentially tethering the inspector to the pilot. Additionally, the tablet could not be charged while in use since it was plugged into the controller.

5.3 Desk Inspection

The videos recorded during the four UAS-assisted field inspections were shared with 19 bridge inspectors who then performed inspections from their desks. Four inspectors reviewed the videos from 18 December, six inspectors reviewed the videos from 19 December, 4 inspectors reviewed the videos from 20 December, and five inspectors reviewed the videos from 21 December. The videos were assigned to the inspectors randomly. These inspections followed the same general procedures outlined for the hands-on and UAS-assisted visual inspections. The results from the UAS-assisted desk inspections were compared to the results from the hands-on inspections and the UAS-assisted field inspections to evaluate the relative quality and practicality of each method.

5.3.1 Inspection Scenario

The UAS-assisted desk inspections were completed at the participants' convenience between April 2018 and July 2018. The inspection procedures outlined in Section 3.2 for the hands-on inspection were followed as closely as practical for the UAS-assisted desk inspections. The inspectors were provided with the videos from one of the four field inspections, written inspection procedures, blank inspection forms, blank vision test worksheets, a confidentiality agreement, dimensioned drawings of the specimens, and an exit survey via a shared folder. Each specimen had at least one accompanying video and some specimens had multiple associated videos. The order of the videos followed the order of the inspection forms. The specimens were inspected from the inspectors' desks using their own hardware (computer and monitor) and software (media player). The

inspectors were allowed, and encouraged, to adjust the playback and display of the video as needed to perform a thorough inspection.

A brief conference call was held with each inspector prior to the start of the review to go over the inspection procedures and inspection paperwork and answer any questions. Before reviewing the specimen videos, the inspectors were asked to complete the three vision tests using the videos of the charts recorded during the field inspections. All of the inspectors were provided with the same vision test videos. Since the test card used for the near visual acuity test is two-sided, only the side with the larger text (paragraph numbers 7 to 11) was visible in the video shared with the desk inspectors.

The inspectors were instructed to spend no more than 8 hours performing the inspection to reduce the burden on their time and ensure that a reasonable and practical level of effort was committed to the inspection. The inspectors were provided with the same blank inspection forms used during the field inspections. On each form, the inspectors were directed to record the length and location of any detected crack(s) or indicate that none were found. Additionally, the inspectors were asked to record the inspection start and end times, rounded to the nearest minute, at the top of each form. Dimension drawings of the specimens were provided for their reference in estimating crack length. Midway through the inspection, the inspectors were directed to complete the NASA Task Load Index worksheets to provide a quantitative assessment of their workload. After completing the inspection, the inspectors were asked to complete a written exit survey to gather information on their education, experience, training, etc. Additionally, the exit survey included questions about the hardware and software used during the inspection. The inspection forms, vision tests forms, signed confidentiality agreement, and exit form were returned via e-mail. All of the documents provided to the inspectors are included in Appendix B.

Each set of videos included the same 54 specimens. Since only two of the four field inspectors inspected the riveted plate specimens, the videos of the riveted plate specimens from 20 December were shared with the inspectors assigned to review the 18 December videos, and the videos from 19 December were shared with the inspectors assigned to review the 21 December videos. Similar substitutions were made in a few other instances because of corrupt or missing video files. The

desk inspection included 32 girder specimens with 15 out-of-plane cracks, 14 welded cover plate specimens with 5 weld toe cracks, and 8 riveted plate specimens with 9 rivet hole cracks.

5.3.2 Inspector Demographics

The videos and inspection information were shared with 33 inspectors; however, only 19 inspectors (all males) completed the inspection and returned the inspection paperwork. Additionally, one of these 19 inspectors did not return the exit survey and so limited information on the inspector's experience and background was available.

Nine of the inspectors worked for a state department of transportation and ten worked for a private engineering and inspection firm. The inspectors worked primarily in eight states, Utah (8), Ohio (3), Wyoming (2), Illinois (2), Indiana (1), Nebraska (1), Oregon (1), and Idaho (1), although the private consultants inspect bridges around the country. Seventeen (17) of the 19 inspectors had completed both the two-week FHWA/NHI *Safety Inspection of In-Service Bridges* course and the 3.5-day FHWA/NHI *Fracture Critical Bridge Inspection* course prior to their participation in the study. Thirteen (13) of the inspectors were licensed professional engineers and/or licensed structural engineers. The average experience of the participating inspectors was 12 years and the inspectors had completed an average of 22 hands-on inspections in the 12 months prior to their participation. The average visual acuity score on the Snellen eye test was 20/25 and the Log Contrast Sensitivity scores on the Pelli-Robson vision test ranged from 1.05 to 1.80. Select inspector demographics are compiled in Table 5.9.

Additionally, two inspectors participated in both the UAS-assisted field and desk inspections and two inspectors participated in both the hands-on inspection and the UAS-assisted desk inspection. Due to the large number of similarly featured specimens and the passage of time, participation in the hands-on or UAS-assisted field inspections was not expected to affect performance in the desk inspection.

Table 5.9 Inspector demographics for UAS-assisted desk inspections

Inspector ID	Employer	Age (years)	Inspection Experience (years)	Professional Licensure	Number of Hands-on (Routine) Inspections	Log Contrast Sensitivity
18BN-82	Private Consultant	48	18	PE	25 (1000)	1.05
18RW-82	Private Consultant	Not Given	Not Given	Not Given	Not Given	1.5
18SD-87	Private Consultant	29	5	PE	50 (100)	1.5
18WJ-84	State DOT	35	5	None	12 (132)	1.8
19AS-85	State DOT	39	14	None	8 (15)	1.65
19HG-80	State DOT	65	10	PE	5 (30)	1.65
19KU-89	State DOT	42	8	None	17 (110)	1.65
19RT-88	Private Consultant	51	12	PE	15 (300)	Not Given
19SG-83	Private Consultant	57	20	SE, PE	10 (200)	1.65
19YD-89	Private Consultant	56	34	PE	57 (49)	1.35
20OD-83	Private Consultant	46	20	PE	80 (60)	1.65
20RJ-83	Private Consultant	44	21	SE, PE	0 (2)	1.8
20QH-83	State DOT	43	1	None	6 (130)	1.65
20YF-86	Private Consultant	30	8	SE, PE	20 (35)	1.65
21DF-85	State DOT	39	8	PE	25 (50)	1.35
21JG-81	State DOT	52	10.67	None	22 (500)	1.65
21EH-80	State DOT	65	7	SE	10 (180)	1.5
21BN-85	State DOT	36	6	PE	8 (25)	1.65
21ET-84	Private Consultant	29	4	PE	9 (200)	1.35

5.3.3 Inspection Results

The information provided by the inspectors on each inspection form was interpreted and categorized as outlined in Section 3.4. Inspector evaluation considered 29 visually detectable cracks with lengths between 1/2 and 5-7/32 inches in 54 painted specimens. The results from the UAS-assisted field inspections were compiled and evaluated similar to the hands-on inspection results. Inspector performance was evaluated for the entire specimen inventory and for each type of specimen (girder, welded cover plate, riveted plate) separately.

5.3.3.1 Crack Detection and False Calls

Nineteen (19) inspectors completed a UAS-assisted deck inspection of the painted specimens. The most successful inspector detected 23 of the 29 possible cracks (79%). The least successful inspector detected 6 of the 29 possible cracks (21%). On average, the inspectors detected 17 cracks

(57%), meaning 12 cracks were not located. The standard deviation of detection rate was 16% indicating that the majority of the inspectors detected between 12 and 21 of the cracks. The number of false calls made during the inspections ranged from 9 to 214. The average number of false calls was 70 with a standard deviation of 57. A summary of the results is provided in Table 5.10 and the individual result for each inspector is shown in Table 5.11.

Table 5.10 Summary of results from UAS-assisted desk inspections

Specimen Type/Crack Type	Detection Rate			Number of False Calls		
	Mean (%)	Standard Deviation (%)	Minimum/Maximum (%)	Mean	Standard Deviation	Minimum/Maximum
All Specimens	57	16	21/79	70	57	9/214
Girder/Out-of-Plane	61	14	33/80	52	39	8/154
Welded Cover Plate/Weld Toe	24	26	0/80	14	13	0/43
Riveted Plate/Rivet Hole	70	24	11/100	4	7	0/26

Table 5.11 UAS-assisted desk inspection results by inspector

Inspector ID	Hits	Possible Cracks	Detection Rate	False Calls	Hit/Call Ratio
18BN-82	13	29	45%	55	19%
18RW-82	14	29	48%	79	15%
18SD-87	6	29	21%	9	40%
18WJ-84	18	29	62%	92	16%
19AS-85	20	29	69%	30	40%
19HG-80	18	29	62%	87	17%
19KU-89	21	29	72%	145	13%
19RT-88	9	29	31%	11	45%
19SG-83	15	29	52%	65	19%
19YD-89	19	29	66%	35	35%
20OD-83	17	29	59%	39	30%
20RJ-83	23	29	79%	51	31%
20QH-83	14	29	48%	42	25%
20YF-86	22	29	76%	194	10%
21DF-85	15	29	52%	83	15%
21JG-81	13	29	45%	39	25%
21EH-80	20	29	69%	214	9%
21BN-85	16	29	55%	21	43%
21ET-84	23	29	79%	42	35%

The number of false positives were compared to the number of hits for each inspection and the hit/call ratio was calculated. The highest, or best, hit/call ratio was 45% and the lowest, or worst, hit/call ratio was 9%. The average hit/call ratio was 25%, meaning that one hit was recorded for every 4 calls. As illustrated in Figure 5.14 where the inspectors are plotted in order of increasing number of hits, the number of false calls (negative y-axis) is positively correlated with the number of hits (positive y-axis). In other words, inspectors that detected more cracks also made false calls.

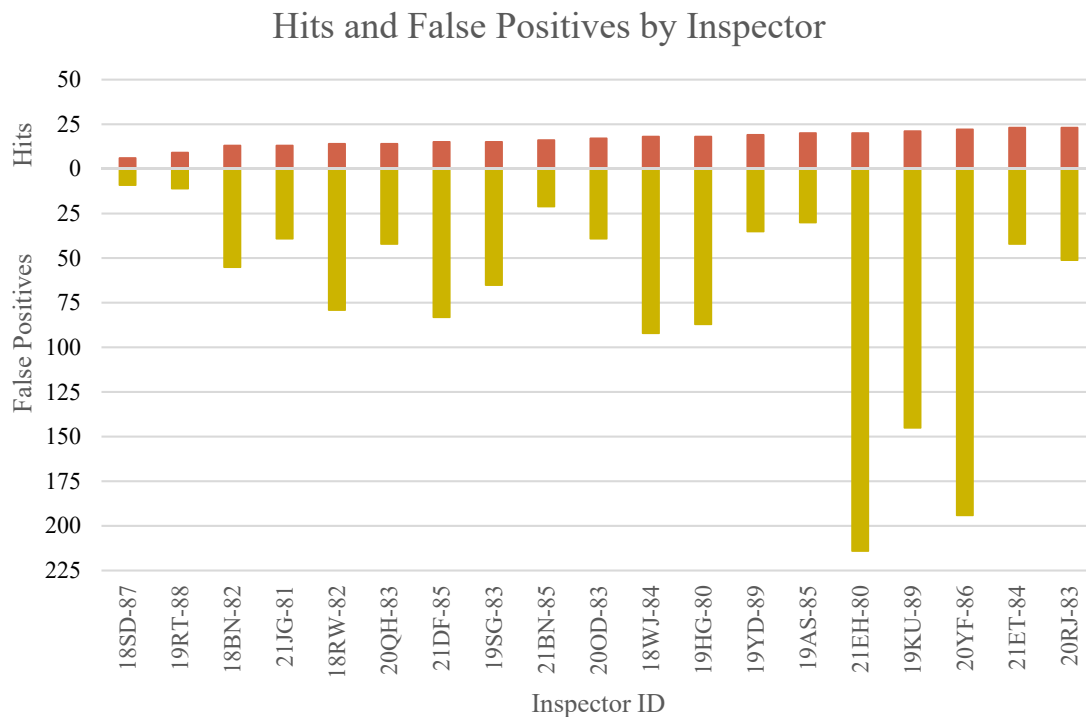


Figure 5.14 Hits and false calls by inspector during the UAS-assisted desk inspections

The detection rates and number of false calls by specimen and crack type were also examined. The course included 15 out-of-plane cracks in the girder specimens, 5 weld toe cracks in the welded cover plate specimens, and 9 rivet hole cracks in the riveted plate specimens. The detection rates by crack type are shown for each inspector in Figure 5.15. As shown in Table 5.10, the weld toe cracks had the lowest average detection rate, 24%, while the average detection rates on the out-of-plane and rivet hole cracks were 61% and 70% , respectively. As discussed for the field inspections, the mounting location of the camera made it difficult to get a clear image of the weld because the plates were positioned overhead. Thirteen (13) of the 19 inspectors recorded their

highest detection rate on the rivet hole cracks and four inspectors recorded their highest detection rate on the out-of-plane cracks. One inspector recorded the same maximum detection rate on the rivet hole cracks and the out-of-plane cracks and one inspector recorded the same maximum detection rate on the out-of-plane cracks and the weld toe cracks. Seventeen (17) of the 19 inspectors recorded their lowest detection rate on the weld toe cracks.

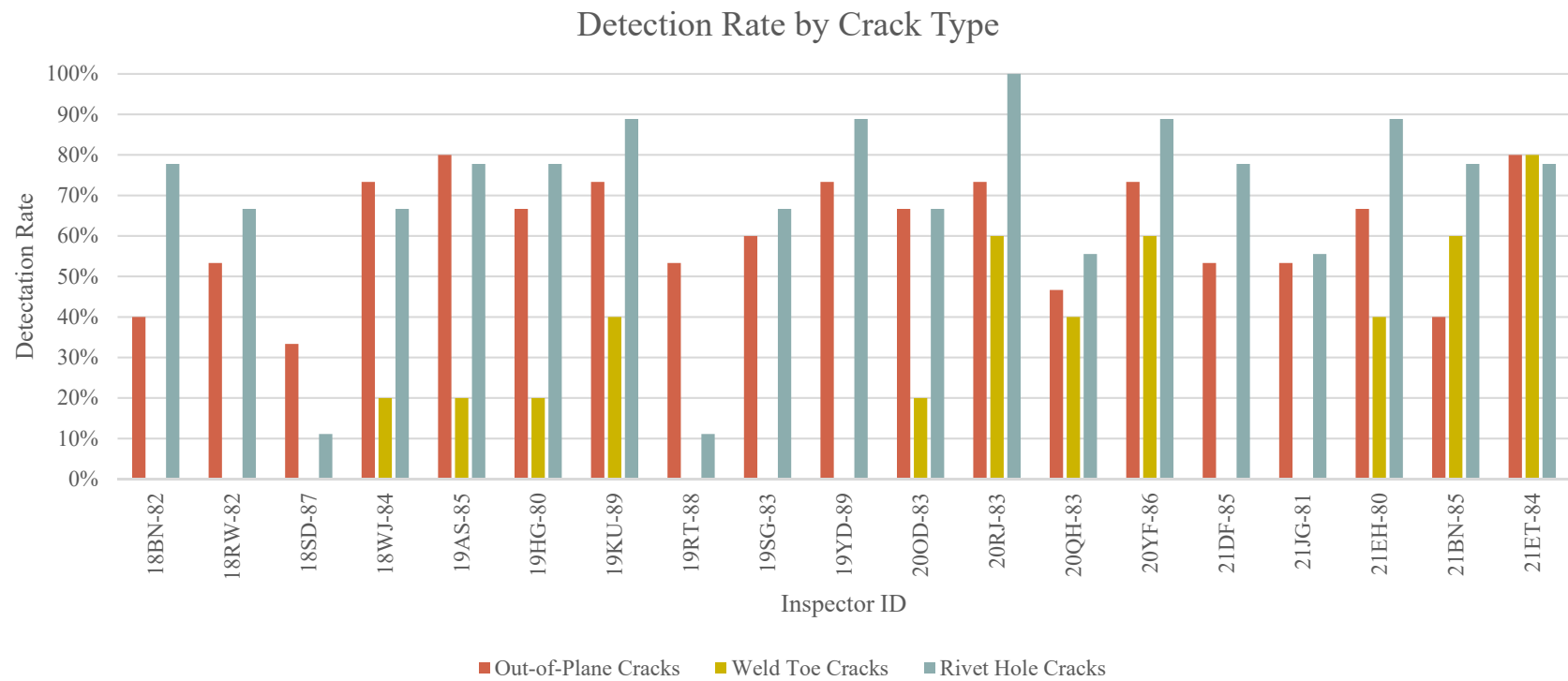


Figure 5.15 Inspector detection rates by crack type

Since four different sets of videos were used, inspection performance by video set was investigated. A summary of performance by video date is shown in Table 5.12. The highest average detection rate and number of false calls was recorded among the inspectors that reviewed the videos from 20 December while the lowest average detection rate and number of false calls was recorded among the inspectors that reviewed the videos from 18 December. Performance on any single day was not significantly different from performance on the other three days.

Table 5.12 Average UAS-assisted desk inspection performance by video set

Video Date (no. of inspectors that reviewed the videos)	Detection Rate			Number of False Calls		
	Mean (%)	Standard Deviation (%)	Minimum/Maximum (%)	Mean	Standard Deviation	Minimum/Maximum
18 December (4 inspectors)	44	17	21/62	59	37	9/92
19 December (6 inspectors)	59	15	31/72	62	49	11/145
20 December (4 inspectors)	66	15	48/79	82	75	39/194
21 December (5 inspectors)	60	14	45/79	80	78	21/214

5.3.3.2 Crack Sizing

Similar to the hands-on inspections, the inspectors participating in the UAS-assisted inspections were asked to record both the length and location of the cracks on their inspection forms. Inspectors were provided with dimensioned drawings of the specimens to provide a reference against which to estimate crack length. Eleven (11) of the 19 field inspectors recorded crack length estimates on their inspection forms. Two of the inspectors that did not estimate crack length explained a physical scale mounted on the inspection surface was needed to accurately estimate size, while the remaining six inspectors did not provide an explanation.

Figure 5.16 shows the crack length data for the out-of-plane cracks in the girder specimens. The actual length of the crack is shown on the horizontal axis and the estimated length of the crack recorded by the inspector is shown on the vertical axis. The diagonal 1:1 reference line represents exact agreement between the actual length and the estimate length. For the majority of the cracks, the average of the estimated lengths plots below the 1:1 line indicating that the inspectors tended to underestimate crack length. This trend was also observed in the results from the hands-on and

UAS-assisted desk inspections. The average measurement error was -1.12 inches and the average absolute error was 1.33 inches. The average absolute error increased with crack length and the percent absolute error remained constant with crack length. Table 5.13 shows the error analysis for the reported length measurement for all the cracks in the girder specimens.

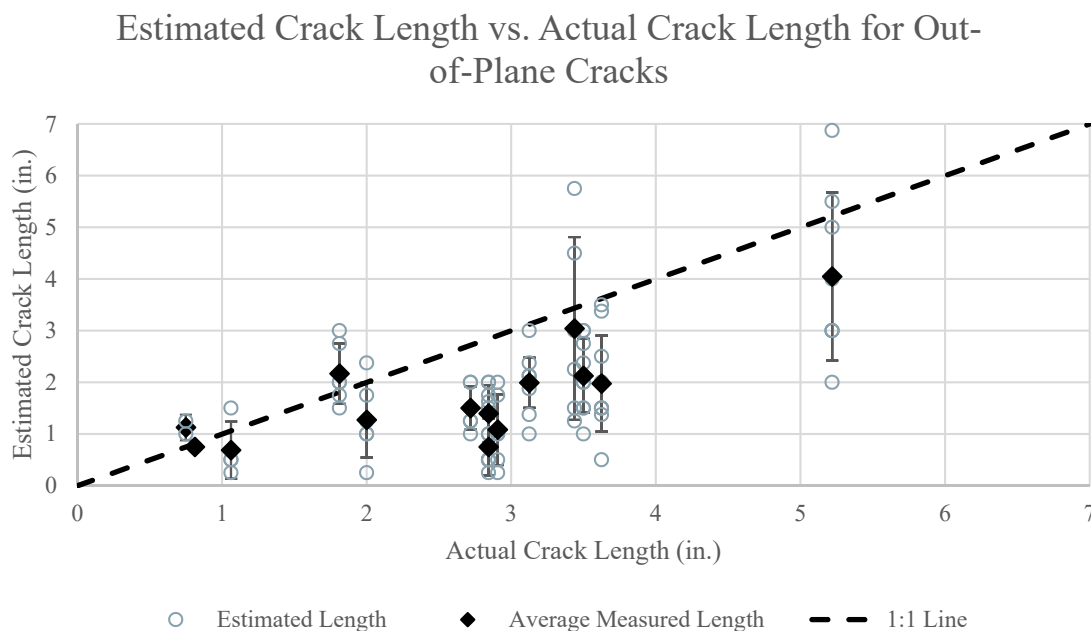


Figure 5.16 Estimated crack length versus actual crack length for out-of-plane cracks

Figure 5.17 presents the crack length data for the welded cover plate specimens. Crack sizing information is limited for these cracks due to the limited number of cracks and low detection rates. The average estimated crack length was larger than the actual crack length for two of the cracks and smaller than the actual crack length for two of the cracks. The average measurement error was -0.36 inch and the average absolute error was 1.28 inches. Table 5.13 shows the error analysis for the reported length measurement for all the cracks in the welded cover plate specimens.

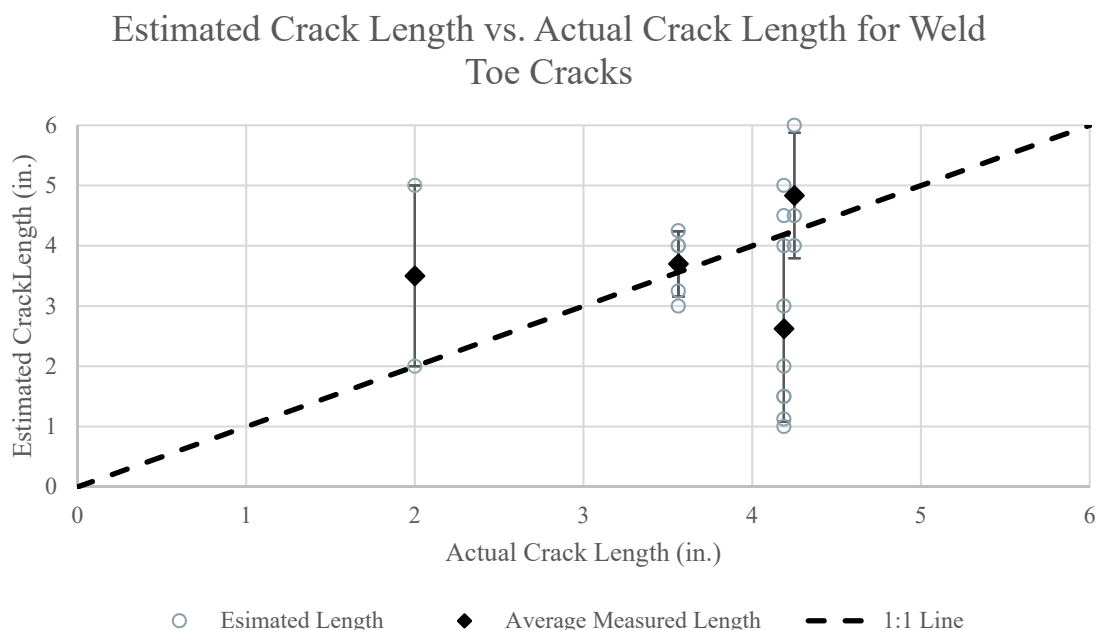


Figure 5.17 Estimated crack length versus actual crack length for weld toe cracks

Figure 5.18 presents the crack length data for the rivet holes cracks in the riveted plate specimens. Similar to the girder specimens and in contrast to the crack sizing results from the hands-on inspections, the average of the estimated lengths of the rivet hole cracks is generally below the 1:1 line indicating that the inspectors had a tendency to underestimate crack length. The average measurement error was -0.04 inch and the average absolute error was 0.31 inch. The average absolute error increased with crack length and the percent absolute error decreased with crack length. Table 5.13 shows the error analysis for the reported length measurement for all the cracks in the riveted plate specimens.

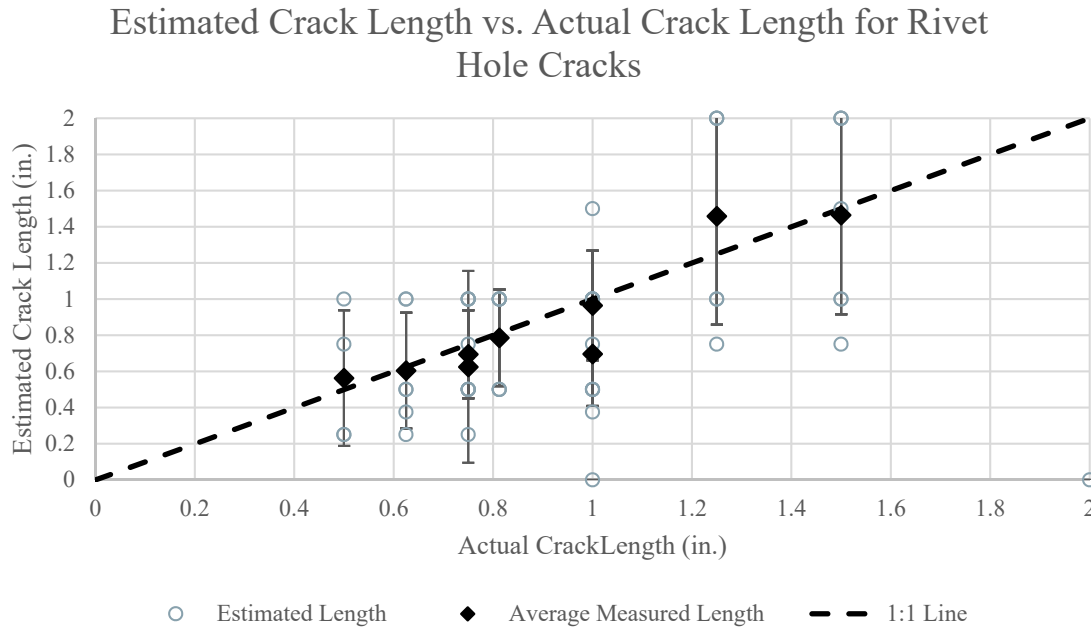


Figure 5.18 Estimated crack length versus actual crack length for rivet hole cracks

Table 5.13 Error analysis for crack length estimates from UAS-assisted desk inspections

Statistic	Out-of-Plane Cracks (15 cracks)		Weld Toe Cracks (5 cracks)		Rivet Hole Cracks (9 cracks)	
	Error (actual length – meas. length)	Absolute Error	Error (actual length – meas. length)	Absolute Error	Error (actual length – meas. length)	Absolute Error
Min./Max. (in.)	-3.22/2.31	-	-3.19/3	-	-0.75/0.75	
Average (in.)	-1.12	1.33	-0.36	1.28	-0.04	0.31
St. Dev. (in.)	1.06	0.78	1.69	1.13	0.38	0.21
Average %	-35	45	-3	37	-4	35

5.3.3.3 Inspection Duration

Both the inspection duration and the length of the inspection videos were recorded during these inspections. The average time to complete the inspection was 245 minutes and the standard deviation was 88 minutes. The fastest inspector completed the inspection in just over two hours (127 minutes) and the longest inspection lasted just over seven hours (427 minutes). Since inspection time was least partly related to the length of the provided inspection videos, a ratio between inspection duration and video length was also calculated for each inspector. As the ratio

increases from one, it indicates that the that the inspector was pausing, slowing down, or re-watching the videos more frequently. This ratio varied between 1.28 and 5.32, with an average of 2.46. Considering each set of inspection videos separately, the average time ratio increased as the length of the inspection videos decreased indicating that longer inspection videos required less additional time to complete the desk inspection. Therefore, this ratio does not provide an exact comparison between inspectors that reviewed different videos; but can be used to compare inspectors that reviewed the same set of videos. A summary of the field and inspection times are presented in Table 5.14.

Table 5.14 UAS-assisted desk inspection durations by inspector

Inspector ID	Inspection Duration (minutes)	Video Length (minutes)	Inspection Duration/Video Length
18BN-82	218	76	2.86
18RW-82	140		1.84
18SD-87	127		1.67
18WJ-84	405		5.32
19AS-85	427	103	4.14
19HG-80	249		2.41
19KU-89	347		3.36
19RT-88	155		1.50
19SG-83	249		2.41
19YD-89	179	70	1.73
20OD-83	131		1.87
20RJ-83	170		2.43
20QH-83	312		4.46
20YF-86	258	182	3.69
21DF-85	261		1.44
21JG-81	241		1.33
21EH-80	284		1.56
21BN-85	263		1.45
21ET-84	233		1.28

5.3.3.4 NASA Task Load Index

After completing their review of the girder specimens, the inspectors were directed to complete the NASA-TLX worksheets included with their inspection forms. Eighteen (18) of the 19 inspectors returned their completed worksheets. The workload scores and governing factors are shown in Table 5.15. The workload scores ranged from 27 to 82 with an average of 61. Similar to

the workload evaluations from the field inspectors, the governing source of workload was determined to be mental demand or frustration for 17 of the 18 desk inspectors. While the relationship between performance and workload was investigated in this research, it is important to remember that workload is not just a measure of effort, it represents the cost (e.g., fatigue, stress, boredom, injury, etc.) incurred on the subject to meet the mission requirements. Therefore, when selecting between two inspectors with equal performance, the inspector with the lower workload score may be better suited for the task. For example, inspectors 19AS-85 and 18SD-87 both achieved a hit/call ratio of 40%. However, their workload ratings were 64 and 42, respectively. This means that even though the inspectors achieved the same level of performance, inspector 19AS-85 did so at a higher personal cost.

Table 5.15 NASA-TLX workload scores for the UAS-assisted desk inspectors

Inspector ID	Workload Score	Governing Factor	Inspector ID	Workload Score	Governing Factor
18BN-82	72	Frustration	19YD-89	47	Mental Demand
18RW-82	82	Frustration	20OD-83	58	Mental Demand
18SD-87	42	Mental Demand	20RJ-83	59	Frustration
18WJ-84	58	Mental Demand	20QH-83	67	Mental Demand
19AS-85	64	Effort	21DF-85	54	Mental Demand
19HG-80	65	Mental Demand	21JG-81	66	Mental Demand
19KU-89	63	Mental Demand	21EH-80	27	Frustration
19RT-88	69	Mental Demand	21BN-85	78	Frustration
19SG-83	68	Mental Demand	21ET-84	64	Mental Demand

5.3.4 Crack Length Analysis

To investigate the relationship between detection and crack length, an abbreviated crack length analysis was performed on the results from the UAS-assisted desk inspections. Due to the reduced number of cracks and inspectors, probability of detection curves were not developed. However, a general understanding of the relationship between crack length and detection can be obtained by looking at the detection rate versus crack length bar graph.

Each crack had the potential to be detected 19 times, once by each participant. Two cracks were detected by all 19 inspectors, while one crack was not detected by any participants. The two cracks that were detected by all the inspectors were over 3 inches in length and located at the bottom of a

transverse stiffener. The undetected crack was 3-1/4 inches long and located along the weld toe of a tapered cover plate specimen. This crack was also not detected during the hands-on inspections. The smallest crack, measuring 1/2 inch, was detected by 7 of the 19 inspectors. This crack was located at a rivet hole. Figure 5.19 shows each crack, sorted by length, and the number of times it was correctly identified by an inspector. Visually, it appears that there was little correlation between crack size and number of detections.

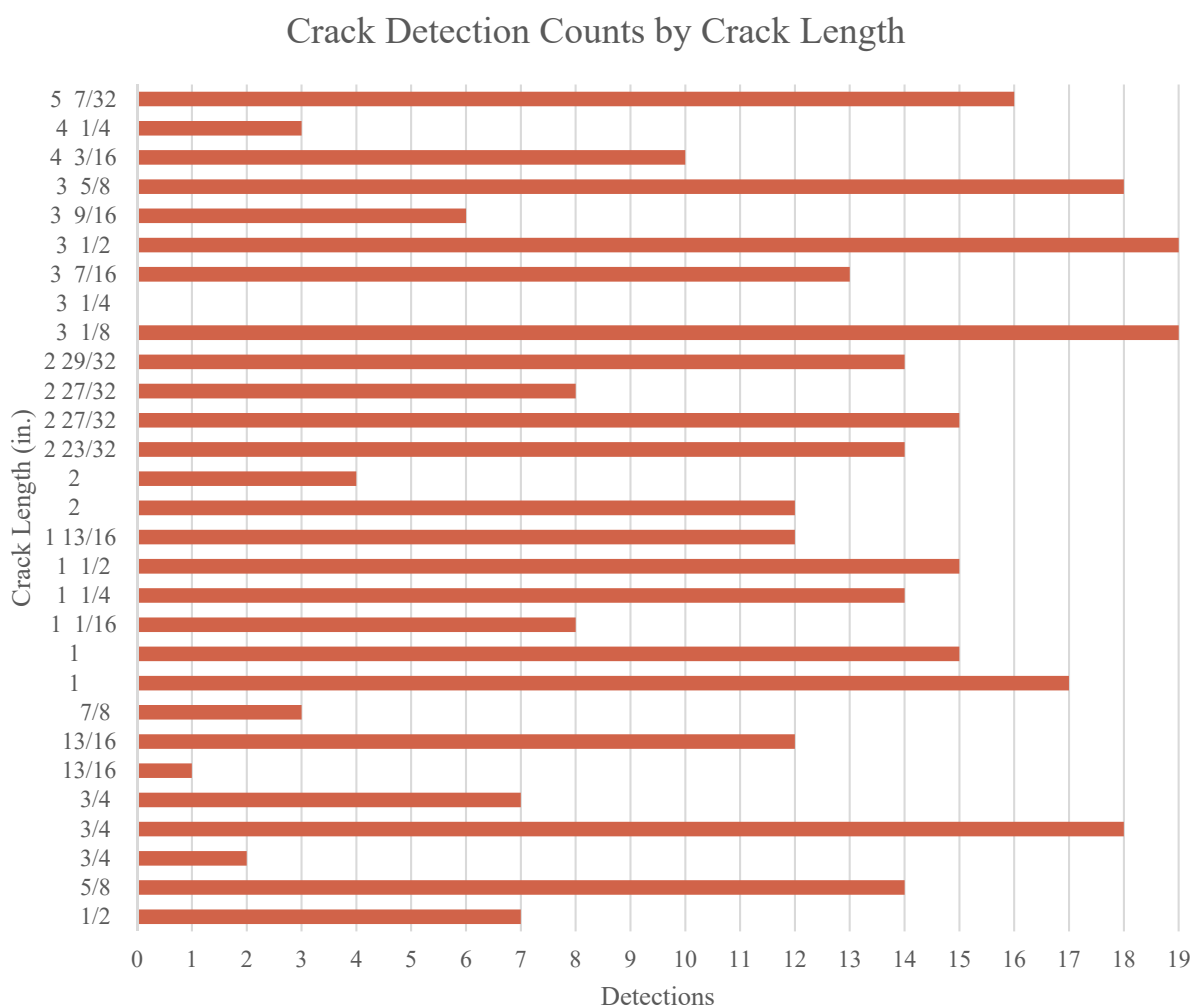


Figure 5.19 Number of crack detections by crack

The cracks were grouped into 1-inch length increments and the detection rate for each bin was computed. The detection rate was determined by summing the number of hits for each crack in the increment and dividing by the total number of detection attempts made. Figure 5.20 shows the

detection rates for each length increment. The bar graph does not show a clear relationship between crack detection and crack length. For instance, cracks between 1 inch and 2 inches in length were detected more frequently, 71%, than cracks longer than 4 inches in length, 51%. The shortest crack length range, less than 1 inch, had the lowest detection rate, 42%. The number of cracks in each size range are shown on the right hand vertical axis.

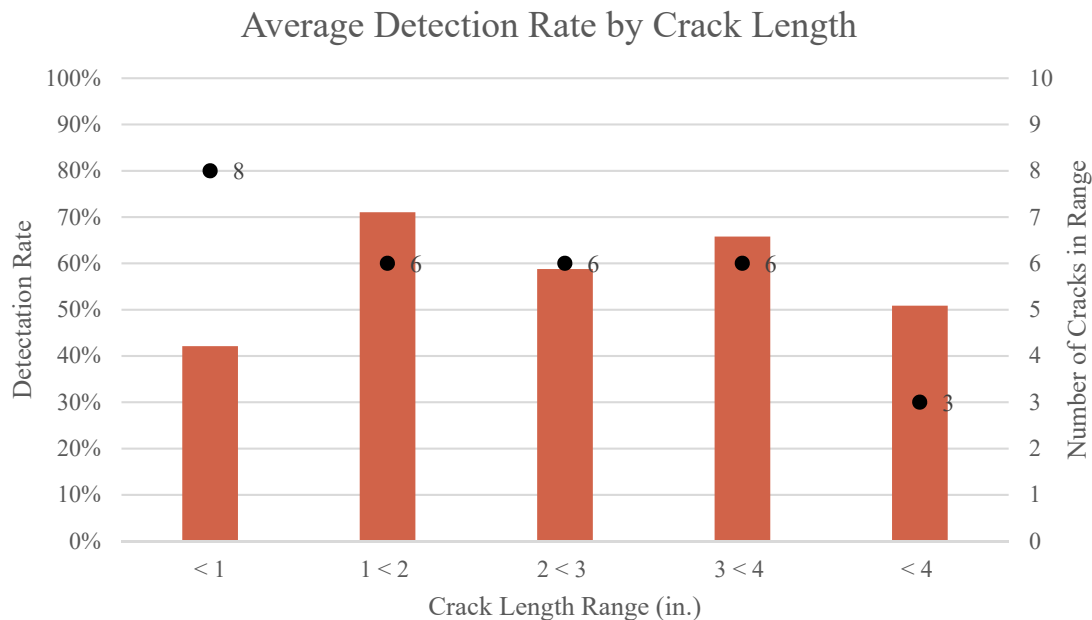


Figure 5.20 Detection rate by crack length

5.3.5 Human and Environmental Factors

As was done with the results from the hands-on inspections, a series of complementary statistical models was applied to the inspection results from the UAS-assisted desk inspections in order to identify key factors that influenced inspection performance. Environmental conditions, specimen characteristics, and inspector attributes were considered in the statistical models. Performance was evaluated based on detection rate and the number of false calls. Descriptive statistics for the independent and dependent variables are shown in Table 5.16.

Table 5.16 Descriptive statistics for dependent and independent variables

Variable Description	Min./Max. Values	Mean of Observations	Standard Deviation of Observations
Performance Measures (dependent variables)			
Detection rate (all specimens)	0.21/0.79	0.57	0.16
False calls (all specimens)	9/214	70	57
Specimen Characteristics (independent variables)			
Crack length	0.5/5.22	2.17	1.33
Out-of-plane crack (1 if the crack is an out-of-plane crack, 0 otherwise)	-	0.517	-
Weld toe crack (1 if crack is a weld toe crack, 0 otherwise)	-	0.172	-
Rivet hole crack (1 if crack is at a rivet hole, 0 otherwise)	-	0.311	-
Environmental Conditions (independent variables)			
VLC media player (1 if the inspector used VLC media player to review the videos, 0 otherwise)	-	0.778	-
Monitor size (diagonal measured in inches)	12/60	22.9	12.9
Monitor resolution (thousands of pixels)	1049/2304	1629	456
December 18 (1 if the inspector reviewed the inspection videos from 18 December, 0 otherwise)	-	0.211	-
December 19 (1 if the inspector reviewed the inspection videos from 19 December, 0 otherwise)	-	0.316	-
December 20 (1 if the inspector reviewed the inspection videos from 20 December, 0 otherwise)	-	0.211	-
December 21 (1 if the inspector reviewed the inspection videos from 21 December, 0 otherwise)	-	0.263	-
Inspector Attributes (independent variables)			
Inspection experience (yrs.)	1/34	11.8	8.15
Age	29/65	44.8	11.3
Inspection duration (min.)	127/427	245	85.9
Normal visual acuity (Snellen eye exam)	3/11	6.63	2.29
Near visual acuity (Jaeger eye exam)	7/11	8.21	1.51
Log Contrast Sensitivity (Pelli-Robson eye exam)	1.05/1.8	1.56	0.19
No. of routine inspections performed in the last 12 months (total)	2/1000	173	240
No. of routine inspections performed in last 12 months (steel)	0/250	55.4	72.1
No. of hands-on inspections performed in last 12 months	0/80	21.1	20.9
No. of training courses attended (out of 8 listed on exit survey)	2/6	3.78	1.22
NASA-TLX workload score	27/82	61.2	13
Professional licensure (1 if licensed PE or SE, 0 otherwise)	-	0.684	-
Employer (1 if employed by a private consultant, 0 otherwise)	-	0.526	-
Previous experience with UAS-assisted bridge inspection (1 if the inspector had previous experience, 0 otherwise)	-	0.333	-
Introduction to Element Level Bridge Inspection (1 if inspector had taken the course, 0 otherwise)	-	0.667	-

Initially, a simple univariate analysis was used to identify statistically significant correlations between inspection performance and the independent variables and the independent two-samples *t*-test was used to determine which independent variables were significant in discriminating between higher and lower performing inspectors. Then, a binary logit model was used to determine which factors, beyond crack length, affected the likelihood of detection for an individual crack. Due to the small sample size, the multivariate regression models were not developed to predict detection rate or false calls. Unless otherwise noted, statistical significance was determined using a two-tailed test with a 95% confidence level and goodness of fit was evaluated using the adjusted R-squared (R^2) or rho-squared (ρ^2) statistic.

5.3.5.1 Univariate Regression and Two Samples *t*-Test

The univariate linear regression analysis and independent two-samples *t*-test were used to investigate the influence of the individual and environmental factors on the performance measures. The form and function of these models are discussed in Section 3.7.1. Table 5.17 shows the *p*-values for all the combinations of performance measures and independent variables. Note that none of the *p*-values are less than 5% which was the threshold used to indicate significance in this study. Since not all of the inspectors provided the requested information on the exit survey, the number of observations used for each analysis is shown in the rightmost column.

Table 5.17 *p*-values for the univariate linear regression analyses

Independent Variables	Performance Measures		Number of Observations
	Detection Rate	False Calls	
Inspection Duration	0.089	0.183	19
Age	0.913	0.500	18
Inspection Experience	0.606	0.364	18
No. of Training Courses	0.771	0.212	18
No. of Routine Inspections (Steel)	0.131 (0.081)	0.666 (0.559)	18 (15)
No. of Hands-On Inspections	0.263	0.376	18
Normal Visual Acuity	0.280	0.447	19
Near Visual Acuity	0.364	0.738	19
Log Contrast Sensitivity	0.322	0.722	18
NASA-TLX Workload Score	0.620	0.065	18
Monitor Size	0.543	0.526	11
Monitor Resolution	0.541	0.637	14

As shown in Figure 5.21, detection rate increased with increasing false positives. In other words, inspectors that made more calls also found more cracks. This trend was not observed during the hands-on inspections.

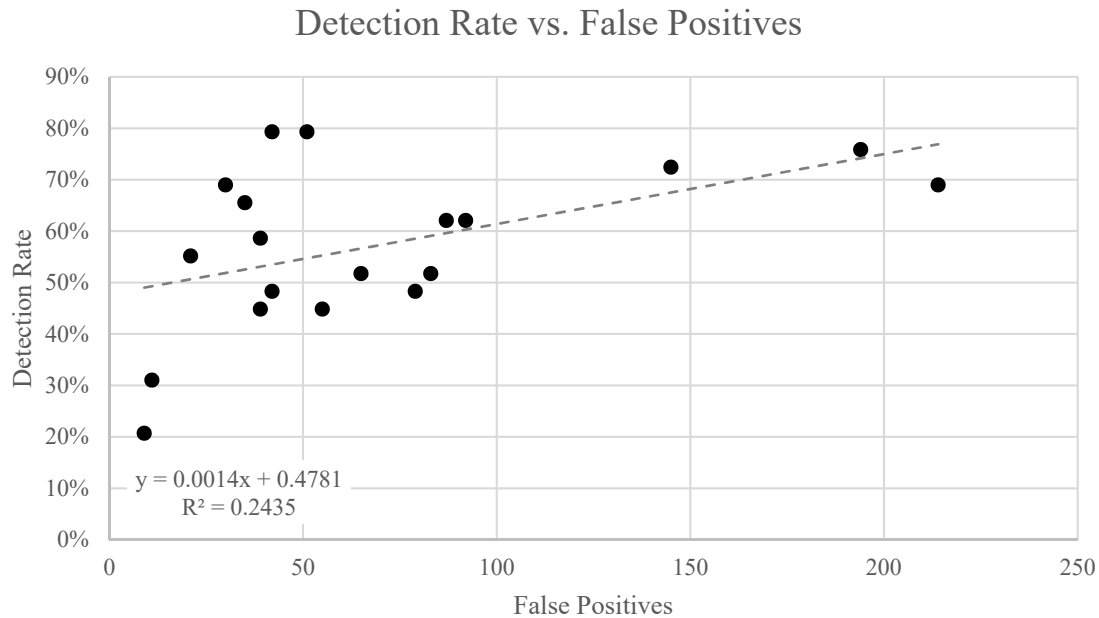


Figure 5.21 Detection rate plotted against the number of false calls

As shown in Figure 5.22, detection rate decreased as the number of routine steel bridge inspections performed in the previous 12 months increased. The *t*-test revealed that inspectors that had performed less than 25 routine inspections of steel bridges in the previous 12 months ($M = 0.690$, $SD = 0.090$) detected significantly more cracks than inspectors who had performed at least 25 routine inspections of steel bridges in the previous 12 months ($M = 0.548$, $SD = 0.148$), $t(12) = 2.31$, $p = 0.040$. A similar trend was observed between detection rate and the number of hands-on inspections performed in the previous 12 months, although this relationship was not statistically significant. The *t*-test also indicated that inspectors that spent less than 50% of the average work week performing inspections ($M = 0.659$, $SD = 0.121$) detected significantly more cracks than inspectors that spent at least 50% of the average work week performing inspections ($M = 0.514$, $SD = 0.165$), $t(15) = 2.16$, $p = 0.047$. The FHWA found a similar relationship between detection of weld crack indications and the number of annual bridge inspections and supposed that inspectors that inspect more bridges per years may be less likely to perform a thorough inspection or less familiar with in-depth inspection procedures [5].

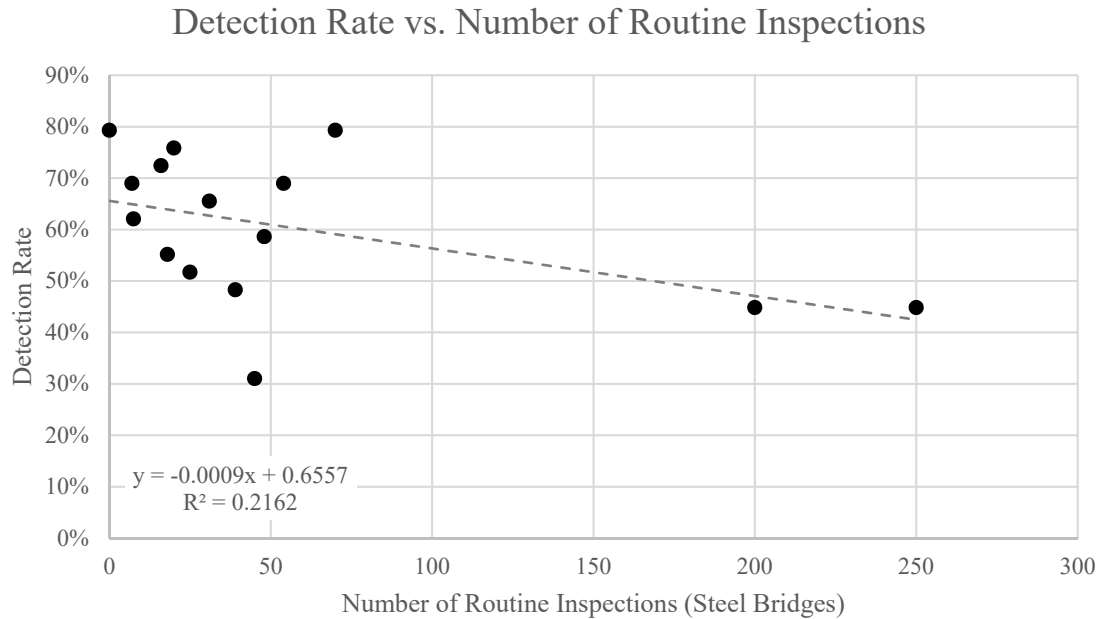


Figure 5.22 Detection rate plotted against the number of routine steel bridge inspections performed in the previous 12 months

Finally, the two sample t -test indicated that previous experience with UAS-assisted bridge inspections, NASA-TLX, and normal visual acuity could be considered significant in discriminating between inspectors that found more or less cracks. The six inspectors with previous experience with UAS-assisted bridge inspections ($M = 0.695$, $SD = 0.080$) detected significantly more cracks than those that had no previous experience with UAS-assisted bridge inspections ($M = 0.520$, $SD = 0.161$), $t(15) = 3.08$, $p = 0.008$. Inspectors with a NASA-TLX workload score below 65 ($M = 0.627$, $SD = 0.162$) detected significantly more cracks than inspectors with a NASA-TLX workload score above 65 ($M = 0.463$, $SD = 0.077$), $t(15) = 2.88$, $p = 0.011$. Inspectors that were able to read the row of letters equivalent to 20/20 vision in the Snellen vision test video ($M = 0.667$, $SD = 0.086$) detected significantly more cracks than inspectors that were not able to read this line of letters in the Snellen vision test video ($M = 0.531$, $SD = 0.167$), $t(16) = 2.34$, $p = 0.033$.

Similar to the results from the hands-on inspections, no single independent variable showed a statistically significant correlation with the number of false positives. The t -test indicated that the only variables that could be considered significant in discriminating between inspectors that made more and less false calls were inspection duration and completion of the *Introduction to Element*

Level Bridge Inspection training course. Inspectors who spent more than 4 hours completing the inspection ($M = 92$, $SD = 65.7$) made significantly more false calls than inspectors who spent less than 4 hours ($M = 40.1$, $SD = 23$), $t(13) = 2.42$, $p = 0.031$. Although this trend was not observed in the hand-on inspections, it has been documented in previous visual inspection studies [21], [26]. Inspectors who had completed the *Introduction to Element Level Bridge Inspection* training course ($M = 88.1$, $SD = 64.5$) made significantly more false calls as compared to the six inspectors that had not completed this training course ($M = 32.8$, $SD = 15.3$), $t(13) = 2.81$, $p = 0.015$. A similar trend was observed in the hands-on inspections as discussed in Section 3.7.2.

5.3.5.2 Binary Logit Model

A binary logit model was used to investigate which factors, beyond crack length, influenced probability of detection during the UAS-assisted desk inspections. Instead of predicting the performance of an inspector, this analysis estimates the likelihood that an individual crack will be detected. Since one inspector did not provide adequate background information, this model is based on 522 observations (18 inspectors x 29 cracks). Because two discrete outcomes, hit or miss, are possible for each observation, a binary logit model is appropriate. The form and function of this model is discussed in Section 3.7.3. The standard binary logit model was modified to include an individual specific disturbance term to account for the unobserved effects caused by repeated observations from the same observer. Random parameters were also considered in this model, but were found to be insignificant based on the standard deviations of the estimated parameters.

Many combinations of factors were analyzed to determine which variables significantly influenced the likelihood of detection. For this population of inspectors, the probability of detection for each crack is best described by a function considering crack length, crack type, the number of false calls made by the inspector, and whether or not the inspector had previous UAS-assisted bridge inspection experience. All variables were considered fixed across the population, while the constant term was allowed to vary to account for random effects. The results from this model are shown in Table 5.18.

Table 5.18 Results from the binary logit model estimating probability of detection

Variable	Estimated Parameter (St. Dev.)	Standard Error Estimate (St. Dev.)	t-statistic (St. Dev.)	P(> t) (St. Dev.)
Constant	-0.570 (0.539)	0.191 (0.115)	-2.98 (4.67)	0.003 (0.000)
Crack Length (in.)	0.828	0.116	7.11	0.000
Out-of-Plane Crack (1 if out-of-plane crack, 0 otherwise)	-1.50	0.274	-5.49	0.000
Weld Toe Crack (1 if weld toe crack, 0 otherwise)	-3.78	0.377	-10.02	0.000
Number of False Calls	0.004	0.002	2.48	0.013
Previous UAS-assisted Bridge Inspection Experience (1 if the inspectors had previous experience, 0 otherwise)	0.587	0.198	2.97	0.003
Number of Observations	522			
Log-likelihood of constant	-369			
Log-likelihood at convergence	-269			
Adjusted ρ^2	0.256			

As was done in Section 3.7.3, the marginal effects, indicating how each parameter affects the probability of detection, were derived. A larger marginal effect indicates a greater influence on the likelihood of detecting the crack while a smaller marginal effect indicates a lesser influence. The marginal effects for the model parameters are shown in Table 5.19.

Table 5.19 Marginal effects for the parameters in the binary logit model

Variable Description	Avg. Marginal Effect (Std. Dev.)
Crack Length (inches)	0.163 (0.044)
Out-of-Plane Crack (1 if out-of-plane crack, 0 otherwise)	-0.262 (0.095)
Weld Toe Crack (1 if weld toe crack, 0 otherwise)	-0.548 (0.166)
Number of False Calls	8.84E-4 (2.37E-4)
Previous UAS-assisted Bridge Inspection Experience (1 if the inspector had previous experience, 0 otherwise)	0.116 (0.029)

Again, the binary logit model gives insight into how crack characteristics, human factors, and environmental conditions interact to affect the likelihood that a specific crack will be detected. In contrast to the binary logit model from the hands-on inspections, the influence of crack length was fixed across the inspector population in this the model. On average, a one-inch increase in crack length resulted in a 16% increase in probability of detection. Notably, this is approximately the same as the crack length marginal effect calculated from the hands-on inspections in Section 3.7.3.

Supporting the previous discussion about the challenges of inspecting overhead with the UAS platform used in this study, the binary logit model suggests that the weld toe cracks in the welded cover plates were less likely to be detected as compared to the out-of-plane cracks in the girder specimens or the rivet hole cracks in the riveted plate specimens. Compared to the rivet hole cracks, this model estimates that out-of-plane and weld toe cracks were 26% and 55% less likely to be detected, respectively.

As the univariate linear regression analysis showed, inspectors that made more false calls also detected more cracks during the UAS-assisted desk inspections. The binary logit model supports this and predicts that an additional false call increases the likelihood of detecting a crack by 0.08%, regardless of crack length.

As was found in the two samples *t*-test, the binary logit model revealed that individual cracks were more likely to be found by inspectors with previous experience with UAS-assisted bridge inspection than by inspectors without previous experience. On average, a crack was 12% more likely to be detected by an inspector with previous UAS-assisted bridge inspection experience.

For this model, the restricted log-likelihood of constants only ($LL(\beta_{RC})$) is -368, the unrestricted log-likelihood of the proposed model ($LL(\beta_U)$) is -269, the number of degrees of freedom (v) is 5, and the X^2 statistic is 199. This exceeds the critical value of 25.7 and a confidence level of over 99.99% is achieved. In other words, the probability that the unrestricted model provides a superior fit to the restricted model by chance alone is very small (less than 0.01%).

5.3.5.3 Computer System

Due to the nature of the desk inspections, it is important to investigate the effect that the computer system (hardware and software) may have on the results. This is roughly equivalent to considering the influence of visual acuity or a magnifying glass during a hands-on inspection. The inspectors were permitted to use any computer system available to them to perform the desk inspection. Inspectors were made aware of VLC media player, which is freely available for download and allows the user to adjust the picture (zoom, contrast, brightness, etc.) while viewing the video [80], however they were not required to use this software. The exit survey included a series of questions

to collect information about the selected system. No significant correlation was found between inspection performance and any single system parameter, however this is likely due to the large variety in systems and system settings and the small sample size. Additionally, many of the inspectors did not respond to all the questions on the exit survey, so only partial system information was available.

The exit survey included a number of questions regarding the inspector's computer hardware, including screen and video card make and model, monitor size, and display settings (resolution, color depth, brightness, scaling, etc.). Twelve (12) of the 19 inspectors provided the size of their primary and/or secondary display. The display diagonals ranged from a 12-inch laptop screen to a 60-inch conference room projection screen. Although there was no clear relationship between display size and inspection performance, the inspector with the smallest reported screen size made the fewest number of calls (most missed cracks and least number of false calls). Fourteen (14) of the 19 inspectors provided the resolution of their primary and/or secondary display. The specified display resolutions varied from 1366 x 768 pixels (1049k) to 1920 x 1200 pixels (2304k). Since none of the inspectors used a monitor with 4k resolution, the videos were down sampled to match the resolution of the display in all cases. Although not statistically significant, inspectors using a display with a resolution greater than 1440k detected more cracks and made fewer false calls than inspectors using a display with a resolution of 1440k or less.

Fourteen (14) of the 19 inspectors used VLC media player, two of the inspectors used Windows Media Player, and three inspectors did not specify which media player they used to review the videos. On the exit survey, a number of common playback features were listed and inspectors were asked to indicate which of these features they used during their review of the videos and how useful these features were from 1 (not very useful) to 5 (very useful). Table 5.20 shows the number of inspectors that used each playback feature and its average usefulness rating for the VLC media player and other media player. All of the inspectors that used the VLC media player used the pause, rewind, zoom, and brightness adjustment features. These features all had average usefulness ratings above 4. These features could be used to improve the visibility of cracks as shown in Figure 5.23. Among the inspectors that used a media player other than VLC, the ability to zoom, rewind, and capture still images were rated as the most useful features.

Table 5.20 Usefulness ratings for video playback features

Feature	VLC media player		Other media player	
	Usefulness Rating	Frequency of Use (out of 14)	Usefulness Rating	Frequency of Use (out of 4)
Pause	4.57	14	5	3
Rewind	4.5	14	5	2
Fast-forward	3.92	12	3.33	3
Decrease playback speed	3.56	9	1	1
Increase playback speed	2.5	8	3	2
Zoom	4.29	14	3	1
Brightness Adjustment	4.07	14	-	0
Contrast Adjustment	4	13	-	0
Color Adjustment	3.77	13	-	0
Saturation Adjustment	3.5	12	-	0
Still Image	3.43	7	5	1

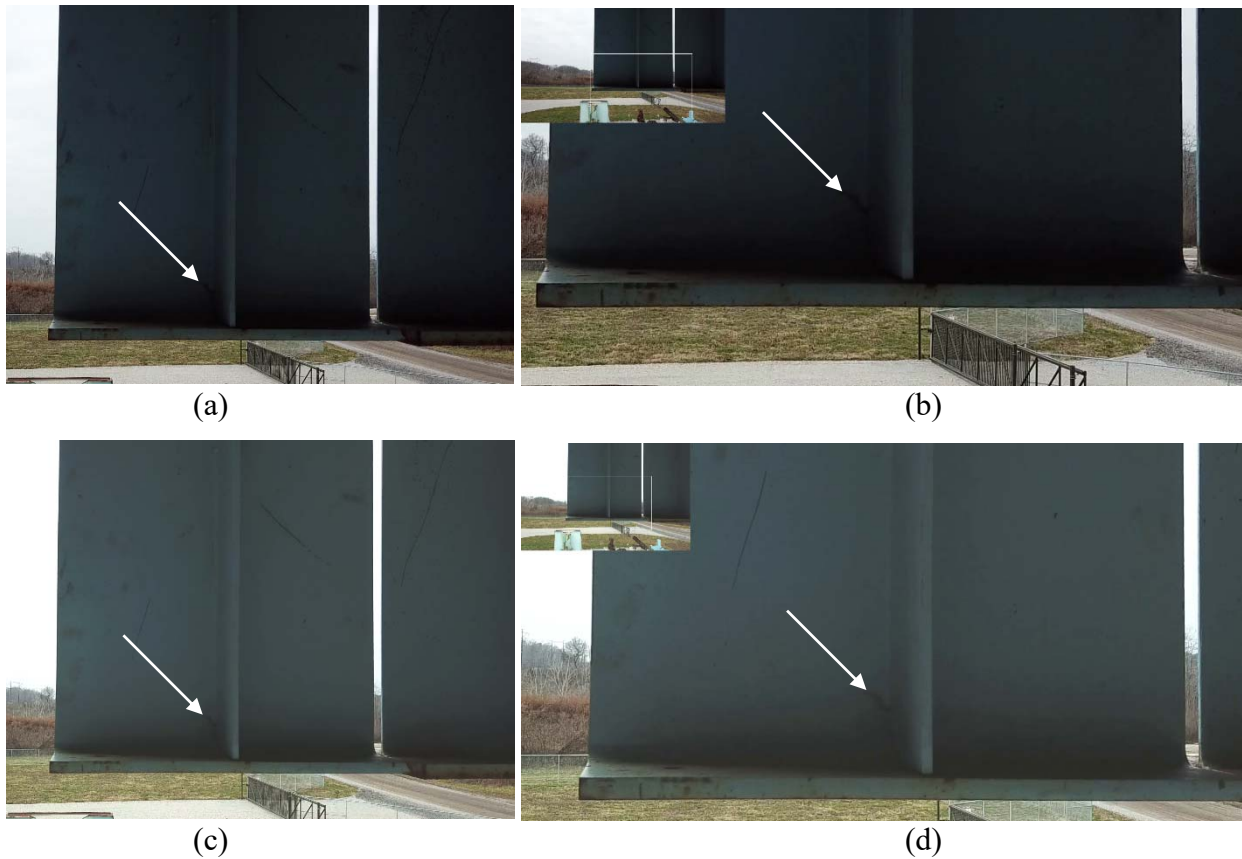


Figure 5.23 Still image from a UAS video (a) uncorrected, (b) zoomed in on the crack-prone region at the bottom of the transverse stiffener, (c) with brightness adjusted, and (d) zoomed in on the crack-prone region with brightness adjusted

On the exit survey, all of the inspectors indicated that they were either satisfied or very satisfied with the quality of the playback, although a few inspectors specified additional technical capabilities that they wished they would have had available during their desk inspection. Of the inspectors that used VLC media player, two inspectors stated that they would have liked easier zoom controls, one inspector mentioned smoother fast forward and rewind functions, and one suggested the ability to bookmark times and record comments within the video. Among the inspectors that did not use VLC media player, one inspector stated that they would have liked to have the ability to zoom and adjust the brightness of the video while it was playing. Not specific to the media player, one inspector commented that touchscreen capability may have made the process easier and another mentioned that it would have been helpful to see the flight information (elevation, heading, speed, etc.) from the UAS.

5.3.5.4 Inspector Assessment

For the majority of the participants in this study, this was their first experience using UAS as a tool during visual inspection. And since these inspectors would be directly affected by the integration of UAS into traditional inspection practices, it is critical to gauge their interest in and support of the new technology.

The exit survey included five questions to elicit feedback on the inspection method and scenario and gather recommendations for improving the quality of UAS-assisted desk inspections. Results from the exit survey are presented below in a question-by-question format. The questions and answer choices are repeated exactly as they appeared on the exit survey. A brief discussion of the results follows each question.

Q1. How did your effort level during this task compare to your effort level during a typical bridge inspection? (*circle one*)

More effort

Similar effort

Less effort

Q2. How did your focus level during this task compare to your focus level during a typical bridge inspection? (*circle one*)

More focused

Similar focus

Less focused

Question 1 and Question 2 were included to obtain a sense of the relative effort and focus levels required for a UAS-assisted bridge inspection as compared to a typical bridge inspection. Note that the questions were not specific about whether a typical bridge inspection referred to a routine inspection or a hand-on inspection, so there may have been some differences in interpretation among the respondents. Eighteen (18) of the 19 inspectors provided a response to these two questions, although one inspector circled both “Similar effort” and “Less effort” for Question 1. The responses to both questions are summarized in Figure 5.24. The responses for effort level were nearly uniform across the levels with six inspectors indicating that the desk inspection required less effort than a typical bridge inspection, five inspectors responding that it required similar effort, and seven inspectors stating that it required more effort. In contrast, the majority of the inspectors stated that they had the same focus level during the desk inspection as during a typical bridge inspection. Two inspectors responded that they were less focused during the desk inspection and three inspectors indicated that they were more focused. No correlation was found between inspection performance and the responses to these two survey questions.

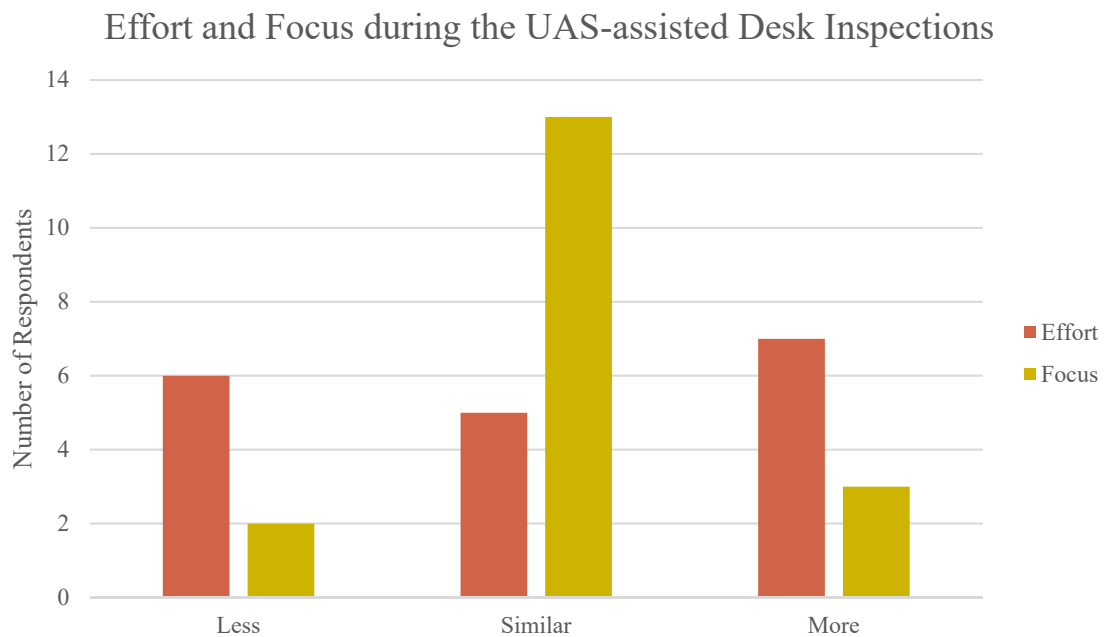


Figure 5.24 Effort and focus level during UAS-assisted desk inspection compared to traditional bridge inspection

Q3. Do you thinking this inspection provided _____ quality as compared to a UAS inspection performed live in the field? (*circle one to fill in the blank*)

Worse Better Similar

Please briefly explain the reason(s) for this selection.

Question 3 was intended to gather information on the inspectors' assessment of the relative quality of the UAS-assisted desk inspection as compared to a UAS-assisted field inspection. Since many of the inspectors had not previously participated in a UAS-assisted field inspection, these responses were based primarily on speculation, not first-hand experience. Sixteen (16) of the 19 inspectors responded to this question, with one inspector specifically abstaining because he had not witnessed a UAS-assisted field inspection. The responses to this question are summarized in Figure 5.25. The responses were relatively evenly divided among the three choices with five inspectors each indicating that they thought the desk inspection provided worse or similar quality to a field inspection and six inspectors stating that they thought the desk inspection provided better quality. One of the two inspectors that participated in both the UAS-assisted desk and field inspections responded that the desk inspection provided better quality than the field inspection while the other inspector felt that the desk inspection provided worse quality. The inspectors that believed the desk inspection provided better quality than a field inspection cited the following advantages:

- Ability to rewind, fast forward, adjust the display, etc.
- Less time pressure
- Comfortable inspection environment (not subject to glare, wind, extreme temperature, etc.)
- Access to larger screens to display the videos

The inspectors that believed the desk inspection provided worse quality than a field inspection cited the following drawbacks:

- No control over what imagery is captured or how it is captured
- Lack of perspective, difficult to orient the details and defects
- Tedious to review all the videos, hard to stay focused and engaged

Q4. Do you thinking this inspection provided _____ quality as compared to an arm's length inspection? (circle one to fill in the blank)

Worse Better Similar

Please briefly explain the reason(s) for this selection.

Similar to Question 3, Question 4 was intended to gather information on the inspectors' assessment of the relative quality of the UAS-assisted desk inspection as compared to a hands-on inspection. One note, under current federal law, a UAS-assisted inspection cannot legally replace a hands-on inspection, but the inspectors' impressions still provide insight into the relative strengths and weaknesses of the two inspection methods. The responses to this question are summarized in Figure 5.25. Eighteen (18) of the 19 inspectors responded to this survey question with the majority indicating that the quality of the UAS-assisted desk inspection was worse than a hands-on inspection. Two inspectors indicated that the quality of the UAS-assisted desk inspection was similar to the quality of a hands-on inspection and one inspector responded that the quality of the desk inspection was better than a hands-on inspection. One the two inspectors that participated in both the UAS-assisted desk and hands-on inspections responded that the desk inspection provided worse quality while the other inspector felt that the desk inspection provided similar quality. The inspectors that believed the desk inspection provided similar or better quality than a hands-on inspection cited the following advantages:

- Able to pay closer attention to detail in the office
- Does not require traffic control for specialized access equipment
- Provides permanent record of the inspection
- Improved access to certain details

The inspectors that believed the desk inspection provided worse quality than a hands-on inspection cited the following drawbacks:

- Cannot interact with the inspection surface
- Poor video quality
- Difficult to make accurate measurements
- Less control over what areas are inspected

Even the inspectors that responded that the UAS-assisted desk inspection provided similar or better quality as a hands-on inspection acknowledged that in certain situations a hands-on inspection would always be necessary.

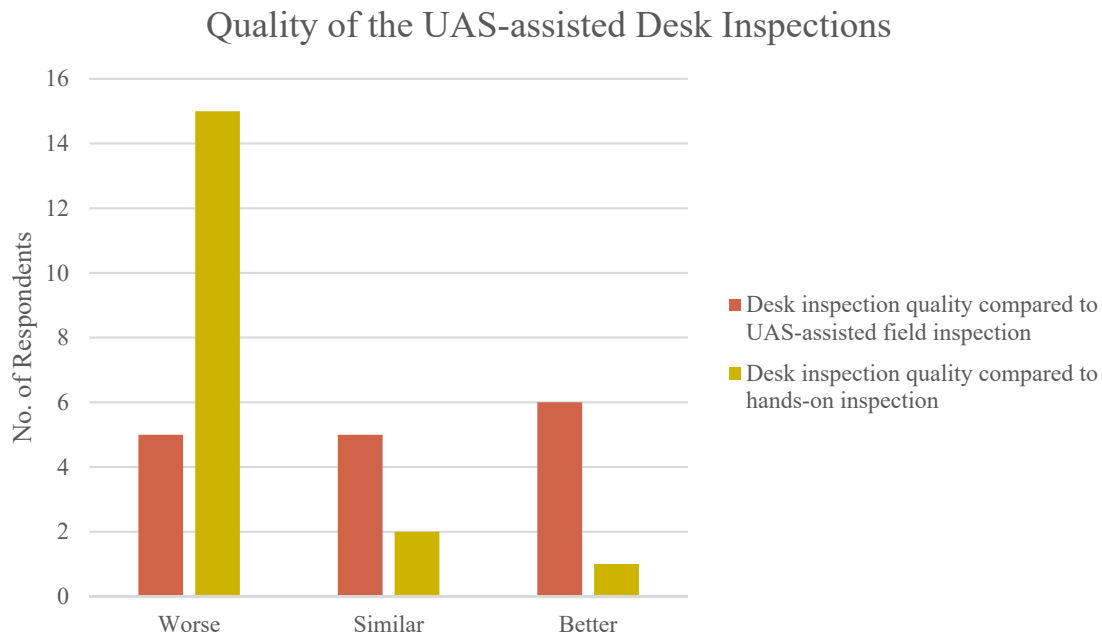


Figure 5.25 Quality of a UAS-assisted desk inspection compared to a UAS-assisted field inspection and a hands-on inspection

Q5. Do you have any suggestions or recommendations to improve performance during a UAS inspection?

Question 5 was an open ended question soliciting recommendations for improving the quality of a UAS-assisted desk inspection. Thirteen (13) of the 19 inspectors provided a response to this question. Some of the inspectors simply included their impression of the inspection, while others provided detailed recommendations for improving the process. The majority of the suggestions fell into three broad categories: UAS technology, video context, and inspection mission. The comments are summarized in Table 5.21. For inspection technology, the inspectors recommended using an inspection-specific UAS with better stability and a higher quality camera. A few inspectors noted that providing more context and references points in the videos would eliminate confusion in the office and reduce errors during the inspection. Finally, the general consensus

among this group of inspectors was that the UAS used in this study would be most effective during a routine inspection, instead of an in-depth inspection.

Table 5.21 Inspector recommendations for improving UAS-assisted desk inspections

UAS Technology	<ul style="list-style-type: none"> • Use UAS platform with better stability • Use camera with higher resolution and improved zoom capability • Collect still images instead of video • Use UAS platform that can fly closer to inspection surface • Use UAS platform with headlamp and/or flash to illuminate inspection surface • Use camera mounted on a pole to photograph areas that hard to access with the UAS
Video Content/Context	<ul style="list-style-type: none"> • Collect imagery at a distance and close up to provide perspective • Label key locations on the bridge (i.e. “FB3, Low Sta.”) • Place or project scale onto the element being inspected for reference
Inspection Mission	<ul style="list-style-type: none"> • Use UAS for routine inspections or as “first pass” before a hands-on inspection • Use UAS for linear or planar elements (long riveted chords, bridge cables, gusset plates, etc.) • Use UAS to create 3D models of specific details • Use UAS to create inspection record

5.3.6 Comparison to UAS-assisted Field Inspection

Since the same specimens were used for the desk and field inspections, a direct comparison can be made between the two methods. While the proctor observations and exit survey responses captured some of the perceived strengths and weaknesses of each approach, the numeric results can be used to quantitatively evaluate the relative accuracy of the inspection strategies.

Table 5.22 provides a brief comparison between the results from UAS-assisted desk and field inspections. The comparison was made over the 54 specimens common between the two inspections. It is important to note that this is not a comparison between a specific field inspector and a group of desk inspectors because the provided video files were not all directly linked to a single field inspector. Instead, this is a comparison between the desk inspectors that reviewed the video and the field inspector that recorded the video. For instance, since the field inspector on 18 December did not inspect the riveted plate specimens, the desk inspectors assigned to the 18 December video set were also given video files from 20 December. Therefore, the comparison

was made between the 18 December desk inspectors and the 20 December field inspector for the riveted plate specimens only. This means that the results presented for the field inspections in this section will not exactly match the results presented in Section 5.2.2.

Over all the specimens, the average performance during the UAS-assisted desk inspections compared well with the UAS-assisted field inspections. The average detection rate for the out-of-plane cracks was slightly higher during the field inspections while the average detection rate for the weld toe and rivet hole cracks was slightly higher during the desk inspections. For all the specimens, the average number of false calls was greater during the desk inspections as compared to the field inspections. Combined, these differences resulted in a lower hit/call ratio for the girder specimens and a higher hit/call ratio for the welded cover plate and riveted plate specimens during the desk inspections. In terms of crack sizing, no overarching trends were observed as the average absolute crack sizing error. The error was larger for the out-of-plane and rivet hole cracks and smaller for the weld toe cracks during the UAS-assisted desk inspections as compared to the field inspection.

Table 5.22 Average performance during UAS-assisted desk and field inspections

	Average from UAS-assisted desk inspections - 19 inspectors (max/min in parentheses)	Average from UAS-assisted field inspections - 4 inspectors (max/min in parentheses)
Girder Specimens (32 specimens)		
Hits	9 (12/5)	10 (11/9)
False Positives	52 (154/8)	33 (51/9)
Total Cracks	15	15
Detection Rate (%)	61 (80/33)	65 (73/60)
Hit/Call Ratio (%)	20 (44/6)	28 (55/15)
Average Crack Sizing Error	-1.12 (2.31/-3.22)	-0.84 (0.78/-2.44)
Average Crack Sizing Absolute Error	1.33	1.02
Welded Cover Plates Specimens (14 specimens)		
Hits	1 (4/0)	1 (3/0)
False Positives	14 (43/0)	11 (14/8)
Total Cracks	5	5
Detection Rate (%)	24 (80/0)	20 (60/0)
Hit/Call Ratio (%)	12 (57/0)	8 (20/0)
Average Crack Sizing Error	-0.34 (3/-3.19)	0.67 (4.44/-1)
Average Crack Sizing Absolute Error	1.21	1.55
Riveted Plate Specimens (8 specimens)		
Hits	6 (9/1)	6 (6/5)
False Positives	4 (26/0)	4 (6/2)
Total Cracks	9	9
Detection Rate (%)	70 (100/11)	61 (67/56)
Hit/Call Ratio (%)	74 (100/24)	60 (75/45)
Average Crack Sizing Error	-0.04 (0.75/-0.75)	-0.1 (0.19/-0.5)
Average Crack Sizing Absolute Error	0.31	0.15

A visual comparison of detection rate and false calls is provided in Figure 5.26 and Figure 5.27, respectively. The large overlap between the standard deviation error bars for the two inspection methods suggests that there was no significant difference in performance, and this was confirmed using the two sample *t*-test.

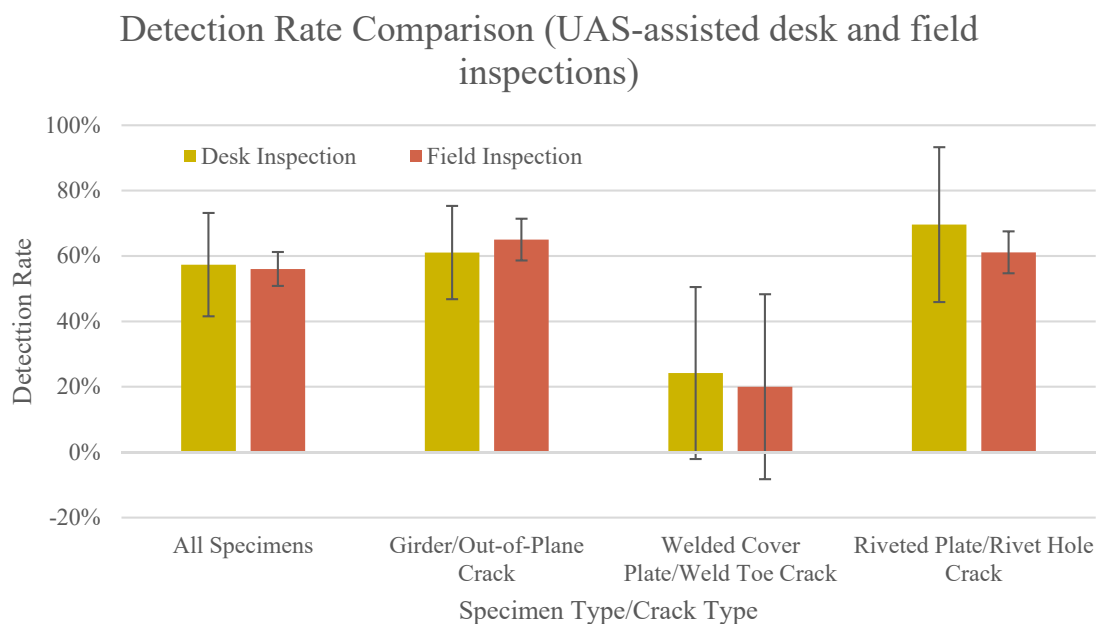


Figure 5.26 Comparison of average detection rate during the UAS-assisted desk and field inspections

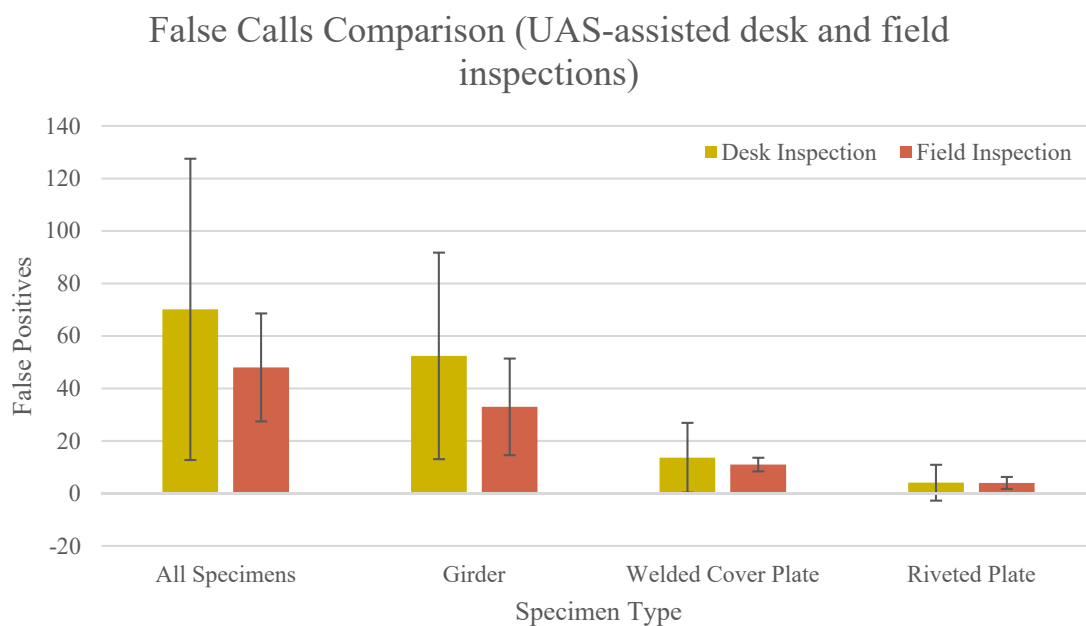


Figure 5.27 Comparison of average number of false calls made during the UAS-assisted desk and field inspections

Since two of the field inspectors also performed a desk inspection review of the same videos, a direct comparison of their performance was made. The results from the two inspections are presented in Table 5.23 along with the difference (field inspection minus desk inspection) between the two inspections. This comparison is limited to the specimens that were inspected during the UAS-assisted field inspection.

Both inspectors made a smaller number of total calls during the desk inspection as compared to their field inspection. Inspector 1 detected four fewer cracks and made 11 fewer false calls for an overall reduction in the hit/call ratio during the desk inspection. Inspector 2 detected three fewer cracks and made 43 fewer cracks during the desk inspection. This resulted in an overall increase in the hit/call ratio for Inspector 2. It is notable that both inspectors made fewer false calls during the desk inspection, while the average number of false calls was higher during the desk inspections. Inspector 1 did not estimate crack length during either the desk or the field inspections so no comparison in accuracy could be made. Similar to the trend observed between the average absolute errors for all the inspectors, Inspector 2 was less accurate in estimating the length of the out-of-plane cracks and more accurate in estimating the length of the weld toe cracks during the desk inspections. Both inspectors experienced a higher level of workload during the desk inspection as compared to the field inspection.

Table 5.23 Difference in inspection performance during UAS-assisted desk and field inspections for two inspectors

	Inspector 1			Inspector 2		
	Field Inspection	Desk Inspection	Difference	Field Inspection	Desk Inspection	Difference
NASA TLX	66	72	-6	72	78	-6
Girder Specimens (32 specimens)						
Hits	10	6	4	9	6	3
False Positives	46	40	6	50	17	33
Detection Rate (%)	67	40	27	60	40	20
Hit/Call Ratio (%)	18	13	5	15	26	-11
Average Crack Sizing Error	-	-	-	-1.2	-2.02	0.81
Average Crack Sizing Absolute Error	-	-	-	1.43	2.02	-0.59
Welded Cover Plates Specimens (14 specimens)						
Hits	0	0	0	3	3	0
False Positives	19	14	5	13	3	10
Detection Rate (%)	0	0	0	60	60	0
Hit/Call Ratio (%)	0	0	0	19	50	-31
Average Crack Sizing Error	-	-	-	-0.58	0.08	-0.67
Average Crack Sizing Absolute Error	-	-	-	0.58	0.21	0.38

5.3.7 Comparison to Hands-on Inspection

Similar to the comparison made between the UAS-assisted field and field inspections in Section 5.2.6, a comparison can be made between the 19 UAS-assisted desk inspections and the 30 hands-on inspections discussed in Chapter 3. Additionally, the results from the two inspectors that performed both the UAS-assisted desk inspection and the hands-on inspection can be compared to provide a direct evaluation of the relative accuracy of the inspection strategies.

Table 5.24 provides a brief comparison between the results from UAS-assisted desk inspections and the hands-on inspections. The comparison was made over the 54 specimens common between the two inspections. Since this comparison only includes about one third of the specimens included in the hands-on inspections, the results presented in Table 5.24 do not exactly match the results presented in Section 3.5.

For all specimens, the average performance during the hands-on inspections was superior to the average performance during the UAS-assisted desk inspections. The average detection rate for all three cracks types was higher during the hands-on inspections as compared to the desk inspections. The greatest difference was observed for the weld toe cracks as their overhead location made detection with the UAS platform used during this study challenging. Similarly, the average number of false positives was higher during the desk inspections as compared to the hands-on inspections for all three specimen types. Combined, these changes resulted in a higher average hit/call ratio for the hands-on inspections. In terms of crack sizing, the average absolute crack sizing error was larger in the desk inspections than the hands-on inspections for all three crack types.

Table 5.24 Average performance during UAS-assisted desk and hands-on inspections

	Average from UAS-assisted desk inspections - 19 inspectors (max/min in parentheses)	Average from hands-on inspections - 30 inspectors (max/min in parentheses)
Girder Specimens (32 specimens)		
Hits	9 (12/5)	10 (13/5)
False Positives	52 (154/8)	30 (85/1)
Total Cracks	15	15
Detection Rate (%)	61 (80/33)	65 (87/33)
Hit/Call Ratio (%)	20 (44/6)	34 (89/6)
Average Crack Sizing Error	-1.12 (2.31/-3.22)	-0.63 (7.38/-4.02)
Average Crack Sizing Absolute Error	1.33	0.93
Welded Cover Plates Specimens (14 specimens)		
Hits	1 (4/0)	3 (4/0)
False Positives	14 (43/0)	8 (18/0)
Total Cracks	5	5
Detection Rate (%)	24 (80/0)	57 (80/0)
Hit/Call Ratio (%)	12 (57/0)	41 (100/0)
Average Crack Sizing Error	-0.36 (3/-3.19)	0.4 (8.44/-1.25)
Average Crack Sizing Absolute Error	1.28	0.72
Riveted Plate Specimens (8 specimens)		
Hits	6 (9/1)	8 (9/3)
False Positives	4 (26/0)	2 (11/0)
Total Cracks	9	9
Detection Rate (%)	70 (100/11)	84 (100/33)
Hit/Call Ratio (%)	74 (100/24)	86 (100/30)
Average Crack Sizing Error	-0.04 (0.75/-0.75)	0.11 (1.19/-0.75)
Average Crack Sizing Absolute Error	0.31	0.21

A visual comparison of detection rate and false positives is provided in Figure 5.28 and Figure 5.29, respectively. Although the error bars are wide for both inspection methods, there was a statistically significant difference in inspector performance during the two types of inspection. The average detection rate during the UAS-assisted desk inspections was approximately one standard deviation below the average detection rate during the hands-on inspections, and the average number of false calls during the UAS-assisted desk inspections was approximately one standard deviation greater than the average number of false calls made during the hands-on

inspections. The t -test revealed that significantly more cracks were detected during the hands-on inspections ($M = 0.699$, $SD = 0.139$) than during the UAS-assisted desk inspections ($M = 0.574$, $SD = 0.158$), $t(34) = 2.83$, $p = 0.008$. Similarly, the t -test showed that significantly less false calls were made during the hands-on inspections ($M = 39$, $SD = 29.9$) than during the UAS-assisted desk inspections ($M = 70.2$, $SD = 57.4$), $t(24) = 2.19$, $p = 0.039$.

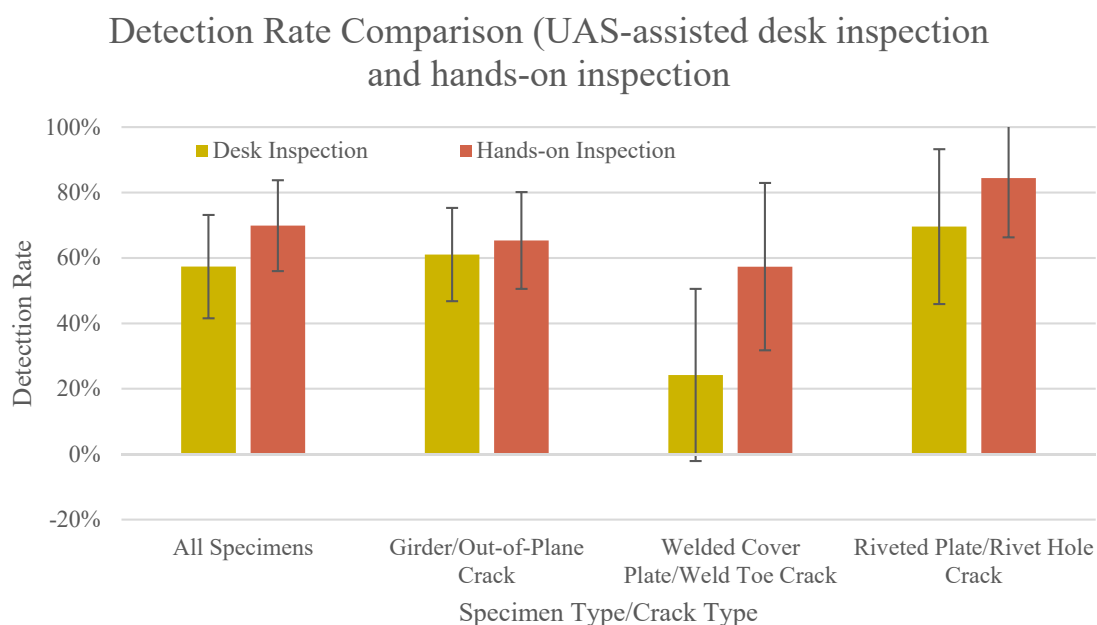


Figure 5.28 Comparison of average detection rate during the UAS-assisted desk and hands-on inspections

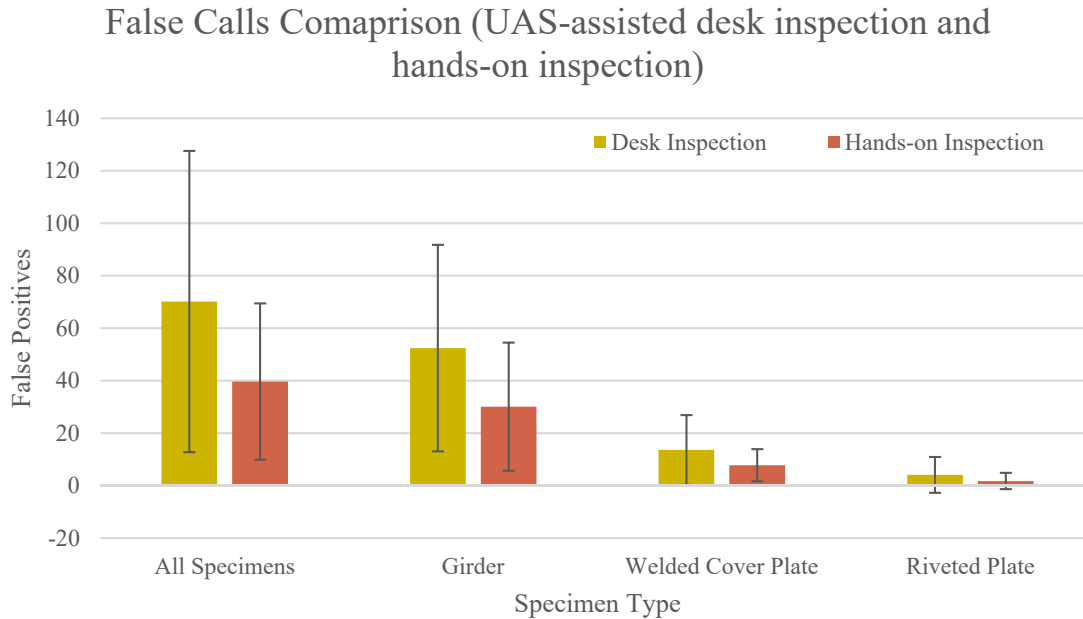


Figure 5.29 Comparison of average number of false calls made during the UAS-assisted desk and hands-on inspections

Since two of the desk inspectors also performed a hands-on inspection, a direct comparison of their performance on the same specimens was made. The results from the two inspections are presented in Table 5.25 along with the difference (hands-on inspection minus desk inspection) between the two inspections. This comparison is limited to the specimens that were inspected during the UAS-assisted desk inspection.

For both inspectors, the hit/call ratio achieved during the hands-on inspection was higher than the hit/call ratio achieved during the desk inspections. Despite this similarity, the inspectors utilized seemingly opposite inspection strategies to transition from the traditional hands-on inspection to the less familiar desk inspection. Inspector 1 made fewer overall calls during the desk inspection as compared to the hands-on inspection, suggesting a more lenient inspection strategy or change in bias. As a result, Inspector 1 detected 13 fewer cracks but also made 25 fewer false calls during the desk inspection. In contrast, Inspector 2 made more overall calls during the desk inspection as compared to the hands-on inspection, suggesting a more conservative inspection strategy or change in sensitivity. As a result, Inspector 2 detected only 1 less crack but recorded 148 more false calls during the desk inspection. Inspector 1 did not estimate crack length during the desk inspections

so no comparison in accuracy could be made. Similar to the trend observed between the average absolute errors for all the inspectors, Inspector 2 was less accurate in estimating crack length during the desk inspection.

Table 5.25 Difference in inspection performance during UAS-assisted desk and hands-on inspections for two inspectors

	Inspector 1			Inspector 2		
	Hands-On Inspection	Desk Inspection	Difference	Hands-On Inspection	Desk Inspection	Difference
Girder Specimens (32 specimens)						
Hits	13	6	7	11	11	0
False Positives	65	40	25	22	130	-108
Detection Rate (%)	87	40	47	73	73	0
Hit/Call Ratio (%)	17	13	4	33	8	26
Average Crack Sizing Error	-0.60	-	-	-0.64	-0.67	0.03
Average Crack Sizing Absolute Error	0.71	-	-	0.69	0.96	-0.27
Welded Cover Plates Specimens (14 specimens)						
Hits	4	0	4	4	3	1
False Positives	15	14	1	13	38	-25
Detection Rate (%)	80	0	80	80	60	20
Hit/Call Ratio (%)	21	0	21	24	7	16
Average Crack Sizing Error	0.31	-	-	0.50	-0.88	1.38
Average Crack Sizing Absolute Error	0.31	-	-	0.75	1.31	-0.56
Riveted Plate Specimens (8 specimens)						
Hits	9	7	2	8	8	0
False Positives	0	1	-1	11	26	-15
Detection Rate (%)	100	78	22	89	89	0
Hit/Call Ratio (%)	100	88	13	42	24	19
Average Crack Sizing Error	-	-	-	0.16	0.35	-0.19
Average Crack Sizing Absolute Error	-	-	-	0.23	0.35	-0.12

5.3.8 Discussion

Although limited in scope, a couple findings from the UAS-assisted desk inspections warrant further discussion as they could influence UAS implementation within the bridge inspection industry.

First, one unexplored topic in visual inspection research is how the shift from field time to office time due to UAS assistance may affect current and future bridge inspectors. A study by the Oregon Department of Transportation and Oregon State University found that the use of UAS technology during bridge inspections may reduce field time by 20% and increase office time by 30% [47]. While UAS assistance is not expected or intended to replace bridge inspectors, this technology could cause a fundamental shift in an inspector's work day as more office time and less field time is required. For many bridge inspectors, the physical nature of the job may be part of its appeal, and so the transition to a more office-based job may be displeasing. Additionally, the results from the desk inspection showed that the inspectors currently performing the most routine and hands-on inspections did not detect the most cracks during the desk inspections. While it is possible that these inspectors were more complacent and less thorough because they perform so many inspections, this trend was not observed in the results from the hands-on inspections. Alternatively, this trend could imply that the current bridge inspector population is not well suited for the office-based review. Future research should consider how the proposed shift in inspection technique could affect the bridge inspector population and possible implications for the agencies overseeing these inspections.

Second, the results from the desk inspections suggest that inspectors had more difficulty distinguishing between true cracks and cracklike surface defects during these inspections as compared to the hands-on inspections. The average hit/call ratio and the hit/call ratios from the two inspectors that performed both the hands-on and the desk inspections were lower during the desk inspections. In practice, this means that in order to achieve the same detection rate during a UAS-assisted desk inspection as during a hands-on inspection, more false calls will be made. This is noteworthy since it is likely that any UAS-assisted inspection protocols will include the requirement to verify the findings with hands-on inspection, at least initially. Therefore, the high number of false calls may reduce the benefits of UAS assistance since substantial hands-on

inspection will still be required to separate the true positives from the false positives. This study also found that the number of false positives increased with inspection duration. Since more indications are likely to be found during a longer inspection, it follows that more false calls may also be made. Additionally, inspectors have been found to relax their response criterion as time on the task increases [26]. Therefore, a time limit could be recommended or imposed to reduce the number of false positives and simulate the temporal pressure experienced during field inspections. However, this may reduce the number of detections along with the number of false calls.

5.4 Recommendations

The results from the UAS-assisted bridge inspections show a similar level of variability in performance as seen in the results from the traditional hands-on inspections. This suggests that human factors continue to have a significant influence on inspection performance, regardless of inspection method. Since many of the recommendations included in Section 3.8 aim to improve visual inspection consistency, these would apply to UAS-assisted bridge inspections as well as hands-on inspections. The recommendations in this section focus on the unique aspects of UAS-assisted inspection. Again, the recommendations will be divided into three categories: inspection equipment and environment, inspector training, and inspection procedures.

5.4.1 Equipment

Inspection equipment specific to a UAS-assisted inspection includes the UAS platform and the computer system used to review the videos. Since this study included only a single UAS platform, it is not possible to identify minimum system requirements or the features which did and did not influence performance. The performance of the DJI Mavic Pro was qualitatively in Section 5.2.7. Although the platform was found to be generally satisfactory, the average performance during the UAS-assisted inspections was worse than during the hands-on inspections, indicating that an improvement in technology may be necessary. Based on the quantitative inspection results, the qualitative evaluation of the DJI Mavic Pro, and inspector feedback, it is recommended that a front-mounted camera equipped with optical zoom be used for in-depth bridge inspections. This will provide a full field of view (straight upward to straight downward) and allow for the collection of detailed, high resolution imagery. It is recommended that the camera system should be able to

capture high resolution still images and record video simultaneously, although this capability is not widely available at this time. Additionally, it is recommended that the UAS platform should be able to hover in the necessary position with minimal intervention from the pilot. This was not possible during the field trials, and the instability of the UAS platform likely affected both the quality and the efficiency of the inspections. Since stability depends on a variety of factors, such as platform weight and power, the onboard navigation system, and environmental conditions, it is recommended that the flight crew establish the minimum acceptable operating conditions for the specific UAS platform and inspection mission during test flights prior to performing the actual inspection.

5.4.2 Training

Training specific to the mechanics of performing a UAS-assisted bridge inspection should be provided to both the inspector and the pilot. For pilots, the training should include basic bridge terminology, common bridge defects, and critical regions of the bridge. For inspectors, the training should include the basics of UAS technology, the capabilities and limitations of the specific platform, and the camera controls (if applicable). These trainings should include a short classroom lecture, flight demonstration, and hands-on practice. The pilots should be given ample time to practice flying close to and beneath a bridge structure. Limited research has been performed specific to training UAS pilots for bridge inspections, although a group of researchers at the Florida Institute of Technology determined that between 1.75 and 2.75 hours was required to train a new pilot to operate a UAS for a high mast light tower inspection [42].

For desk inspections, additional training should be provided to ensure that inspectors know how to properly adjust the display settings (resolution, brightness, etc.) on their screen or monitor and operate the features of the media player. This training should also include both a demonstration and a hands-on exercise. The results from the vision tests suggest that the inspectors possessed varying abilities to manipulate the videos to improve the clarity of the image and the correlation between normal visual acuity and detection rate implies that the inspectors who were more successful in improving the quality of the videos during the vision tests also found more cracks during the inspection.

5.4.3 Inspection Procedures

The following procedural recommendations are based on the quantitative results from the UAS-assisted desk and field inspections, the inspectors' assessments of the inspection scenario, proctor observations, and the literature review. Although the effectiveness of these recommendations was not explicitly evaluated in this study, they represent the best practices available at this time. Future research should continue to evaluate and adjust these recommendations.

- Collect videos and still images during the inspection. Both still images and videos should be collected during the inspection. The videos should include both distance and close-range recordings to provide a frame of reference and context. At a minimum, still images should be captured of suspected defects since these higher resolution images will retain their clarity as the image is enlarged. Alternately, the camera could be programmed to automatically capture still images at a predetermined time interval. This may reduce the responsibility of the pilot and/or inspector, but will increase the amount of data collected during the inspection. The UAS platform should allow for both still images and videos to be captured concurrently without intervention from the operator.
- Perform inspection in the field and review results in the office. The findings from this study suggest that for an in-depth inspection, the inspection should be performed in the field during the flight(s). A desk inspection may be performed after the field inspection to reduce the number of calls, although the number of true positives is likely to decline along with the number of false positives. Alternative approaches, such as having a dual certified pilot/inspector collect the information in the field and review it in the office may be effective, however further research is needed to validate this approach.
- Coordinate inspection equipment with inspection mission and inspection conditions. During the inspection planning phase, the inspection mission, equipment, and conditions should be identified and coordinated. The objectives of the inspection will establish the necessary UAS capabilities and operating conditions. The inspection should not be performed if the actual environmental conditions differ from the expected conditions identified during planning. Test flights before the inspection may be necessary to establish a range of allowable operating conditions for the proposed mission and equipment.
- Confirm findings from UAS-assisted inspections with hands-on inspection. A hands-on inspection should be performed to verify the findings from the UAS-assisted inspection

before planning maintenance or repair activities. If the number of false positives in practice is found to be less than the number of false positives in this study, this requirement could be gradually relaxed.

- Establish initial and recurrent performance testing requirements. Similar to the recommendation made for hands-on inspections, performance testing in a controlled environment should be utilized to confirm that, together, the pilot, inspector, and UAS platform can achieve a satisfactory level of performance. Additional research under Transportation Pooled Fund (TPF) Study No. 5(387) aims to develop standards, protocols, and testing requirements for UAS in a variety of civil engineering applications, including bridge inspection.

5.5 Summary

Unmanned aircraft systems (UAS) have the potential to dramatically change how bridge inspections are performed. Proponents cite the reduced need for traffic control and access equipment as the primary advantages; however, without quantitative data to validate the quality of these inspections, decision makers are left with little beyond vendor literature on which to base deployment decisions. In this project, the use of UAS for detecting fatigue cracks in steel bridge members was investigated through a series of real-time (field) and offline (desk) inspections. The inspections were performed using the same specimens and inspection procedures as the hands-on inspections to allow for a direct comparison of performance.

Four inspectors were invited to perform a UAS-assisted field inspection of the POD specimens at the S-BRITE Center. All inspections were completed with the same pilot operating the same UAS platform. During these inspections, the defection rates ranged from 48% to 60% and the number of false calls varied between 26 and 67. The average detection rate during the UAS-assisted field inspections was significantly lower than the average detection rate during the hands-on inspections, while there was no significant difference between the number of false calls made during the two types of inspection. The greatest decrease in performance was observed in the detection of weld toe cracks in the welded cover plates mounted overhead. Notably, the UAS-assisted inspections were much less efficient as none of the field inspectors inspected all 147 specimens included in the hands-on inspections. Overall, the UAS platform used for these inspections, the DJI Mavic

Pro, was found to be marginally acceptable for detecting fatigue cracks in steel bridge members. Judged on five considerations, the unit size, battery life, and user interface were found to be acceptable, while the navigation and imaging systems and need improvement.

Nineteen (19) inspectors were invited to perform a UAS-assisted desk inspection of the videos captured during the four field inspections. The participants were randomly assigned to four groups and provided with a link to download the appropriate set of videos. During these inspections, the average detection rate was 57% and the average number of false calls was 70. A similar variability was seen in the results from these 19 inspectors as was seen in the results from the 30 hands-on inspectors. The hit/call ratio varied from 9% to 45% and the number of hits was correlated with the number of false calls. Again, only a small amount of the variance would be explained by the human and environmental factors. A linear regression analysis and two-sample *t*-test were used to determine that detection rate was influenced by the number of bridge inspections performed in the previous 12 months, the percentage of time spent performing bridge inspections, previous experience with a UAS-assisted bridge inspection, NASA task load index, and normal visual acuity. The number of false calls was found to be correlated with inspection duration and completion of the *Introduction to Element Level Bridge Inspection* training course. A binary logit model predicted that the probability of detection for a specific crack was influenced by crack length, the location or type of crack, the number of false positives made by the inspector, and previous experience with UAS-assisted bridge inspection. The average performance during the UAS-assisted desk inspections was not statistically different from the average performance during the field inspections, but it was significantly worse than the average performance during the hands-on inspections.

6. TRUSS CHORD INSPECTION AND LOAD RATING

6.1 Introduction

In order to gain a better understanding of the effects of corrosion and section loss on the structural capacity of built-up steel tension members, two segments of the lower truss chord from a deck truss approach span were removed from the Winona Bridge in Winona, Minnesota and transported to Purdue University for testing. To complement the findings from the fatigue crack detection inspections, these bridge specimens were used to investigate variability in the inspection and evaluation of severely corroded steel tension members. A small round robin inspection and load rating study was performed with certified bridge inspectors and practicing load rating engineers. This process evaluated two separate, but related, sources of variability within the inspection and load rating process. The variability in each task was controlled such that variability in the load ratings was not compounded by variability in the inspection findings.

The inspection and load rating set-up and results are discussed in detail in the following sections. The methods used to evaluate the data will be explained along with findings, recommendations, and conclusions regarding the visual inspection and load rating of corroded steel bridge members.

6.2 Inspection

A small pool of bridge inspectors was invited to perform a detailed inspection of one of the truss chords between, but not including, the two gusset plates, as shown in Figure 6.1. Inspectors were directed to use typical inspection procedures to evaluate the condition of the chord and provide the necessary thickness measurements to support a load rating calculation. The inspection was divided into three separate tasks to gain insight into the individual inspection strategies employed by the inspectors, but also to provide consistent data that could be directly compared to evaluate variability. The variability was evaluated by comparing the inspector measurements to each other and to the reference measurements determined by the author.

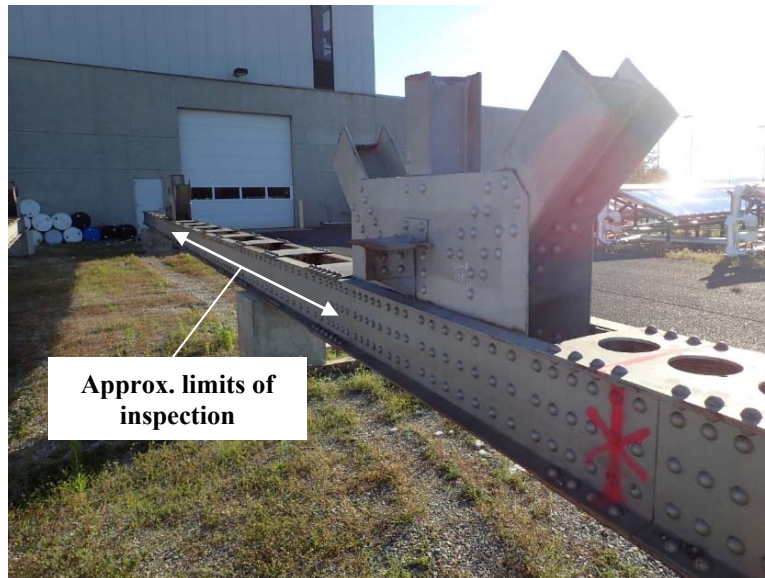


Figure 6.1 Truss chord elevation view (looking from joint L4 to L2)

6.2.1 Description of Specimen

The specimen used for this inspection was the bottom chord of a deck truss from a bridge constructed in 1941 in Winona, Minnesota. The chord is built up from 2-C15x40 channels with 3/8-inch cover, or web, plates. All connections are riveted and the channel sections are connected with intermittent 3/8-inch batten plates. Significant pack rust has formed between the web plates and the channels causing severe distortion of the web plates in some locations. The elevation view and cross sections of the specimen are provided in Figure 6.2 and photos of typical corrosion, section loss, and pack rust along the chord are shown in Figure 6.3.

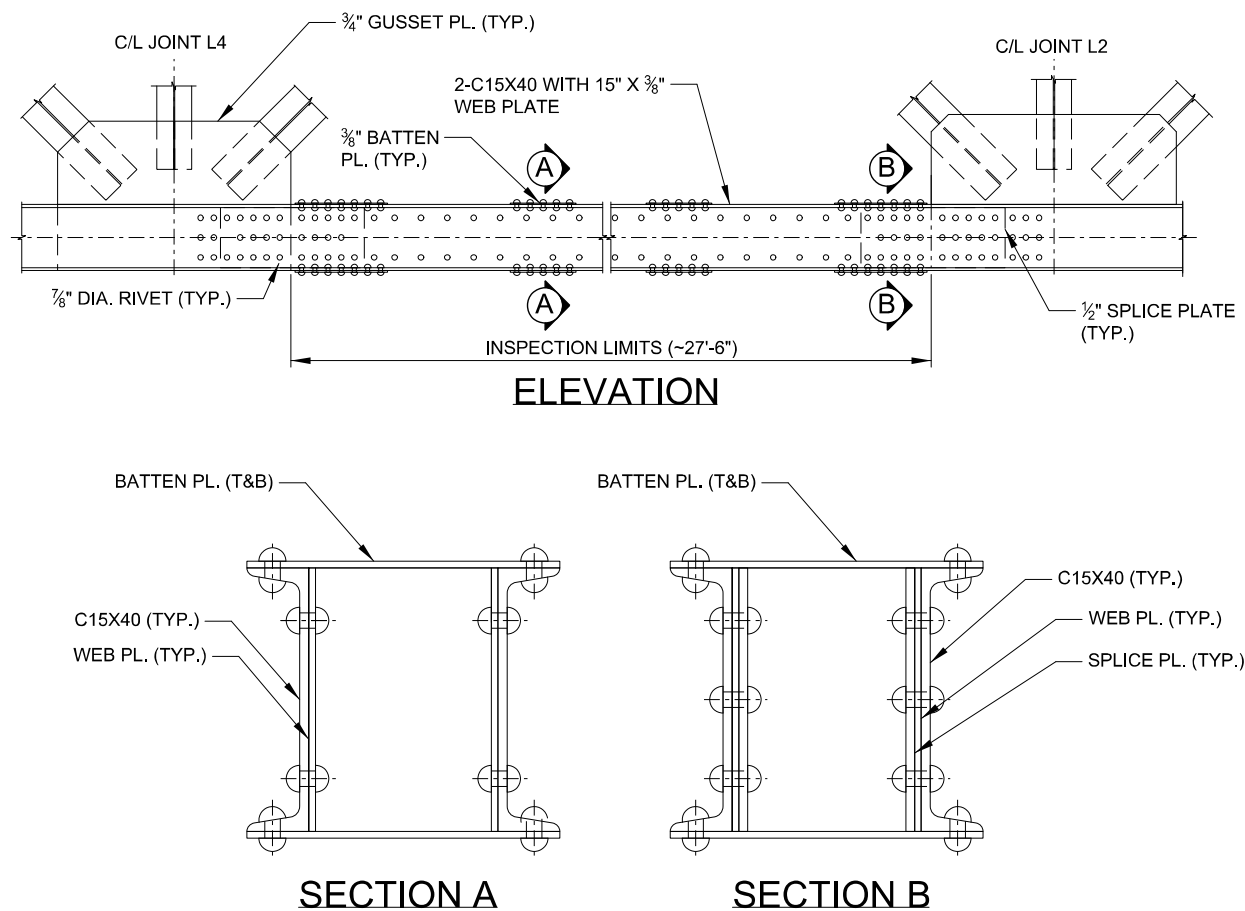


Figure 6.2 Elevation and section views of the truss chord



Figure 6.3 View of typical corrosion damage resulting in section loss and pack rust

6.2.2 Inspection Scenario

The inspection was intended to represent a typical hands-on inspection of a fracture critical bridge member. Prior to the inspection, all of the inspectors were read the same set of instructions and provided with a set of blank inspections forms. During each inspection, the proctor recorded the weather condition, the start and end times for each task, the tools used by the inspector, and general firsthand observations of the inspector's activities. After the inspection, the inspectors were asked to complete a written exit survey to gather information on their education, experience, and training. The inspection procedures, inspection forms and exit survey are provided in Appendix C.

The inspection was divided into three separate tasks. Before beginning each task, the proctor read the same set of instructions to the inspector and answered any questions. Task 1 was the most unstructured task. In this task, the inspectors were asked to identify the critical section for measuring section loss within the inspection limits. The inspectors were allowed to use any tools that they brought with them to complete this task and a time limit of 30 minutes was imposed. This time limit was intended to encourage the inspectors to approach the task with a practical and realistic strategy. In Task 2, the inspectors were asked to estimate the remaining thickness of the truss chord at two specific locations. The locations were identified on the inspection forms and marked on the chord. The inspectors were allowed to use any tools they brought with them to complete the task and no time limit was imposed. In Task 3, the inspectors were asked to identify the critical section for measuring section loss within a 28-inch segment of the chord spanning between two adjacent batten plates. The inspectors did not need to record the remaining thickness at this location, but could use any tools they brought with them to complete the task. No time limit was imposed. The three tasks were designed to yield inspection results that could be compared across the inspectors, while also providing a realistic representation of the variability in inspection strategies.

The inspections were completed outside at Bowen Laboratory on the campus of Purdue University in West Lafayette, Indiana between April 2018 and July 2018. The inspections were completed from the ground adjacent to the truss chord and no specialized access equipment was required. The average temperatures on the days of the inspections ranged from 47°F to 80°F. Figure 6.4 shows two of the inspectors taking thickness measurements during their hands-on inspections.



Figure 6.4 Inspectors taking thickness measurements of the truss chord

6.2.3 Inspector Demographics

Five inspectors (4 males and 1 female) participated in the inspection round robin. Two of the inspectors were from federal agencies, one was from a state department of transportation, and two worked for a private engineering and inspection firm. Participating inspectors were based in three states, Louisiana (1), Illinois (2) and Minnesota (2), although the private consultants perform bridge inspections around the country. The inspectors ranged in age from 25 to 59 and their experience varied from 1 year to 18 years. The average experience of the group was 7.2 years. Four of the five inspectors possessed a post-secondary degree in engineering and were either an engineer-in-training or a professional engineer. Four of the five inspectors had completed the FHWA/NHI 2-week course *Safety Inspection of In-service Bridges*. Only one inspector had received additional training specific to estimating section loss in steel members. The inspectors had performed between 0 and 60 hands-on inspections during the 12 months prior to their participation in the study and the percentage of time spent performing inspections ranged from less than 1% to 95%. Select inspector demographics and inspection conditions are summarized in Table 6.1.

Table 6.1 Inspector demographics and inspection conditions

	Inspector 1	Inspector 2	Inspector 3	Inspector 4	Inspector 5
Temperature (°F)	55	55/47	75	80	80
Weather	Overcast, light rain	Clear	Overcast, humid	Hot, humid	Hot, humid
Age	59	29	45	27	25
Gender	Male	Female	Male	Male	Male
Employer	Federal Agency	Federal Agency	State DOT	Private Consultant	Private Consultant
Years of Inspection Experience	10	1	18	5	2
No. of Routine Inspections in Previous 12 Months	4	1	80	20	30
No. of Hands On Inspections in Previous 12 Months	4	0	60	4	1
Professional Licensure	PE	FE	None	FE	FE
Education	Master's Degree	Bachelor's Degree	Bachelor's Degree	Bachelor's Degree	Bachelor's Degree

6.2.4 Inspector Response Evaluation

Determining the remaining member area based on the inspectors' recorded thickness measurements was not a straightforward task. In many cases, assumptions about the intentions of the inspectors had to be made since the inspectors often just noted measurements on the sketches without providing locations or dimensions. Two of the inspectors explained that common practice is to take the measurements in the field and then calculate section loss in their office, so they may not be accustomed to making their field notes intelligible to others. For consistency in this study, a numerical approximation similar to the middle Riemann sum method was used to calculate the remaining area from the thickness measurements. Each recorded thickness measurement was assumed to be the measurement at the center of a rectangular shaped increment as shown in Figure 6.5 and the total remaining area of the member was calculated by summing the areas of the increments. As the number of increments increases, the length of the increments gets shorter and the sum of the areas eventually approaches the true area of the member. The number and length of the increments was determined based on the number of thickness measurements recorded by the inspector; more thickness measurements resulted in shorter increments. When the exact location of the measurements was not provided, they were assumed to be evenly spaced, and when

no thickness measurements were provided, the design thickness of the component was used to calculate the remaining area. The rivet holes were ignored during all area calculations.

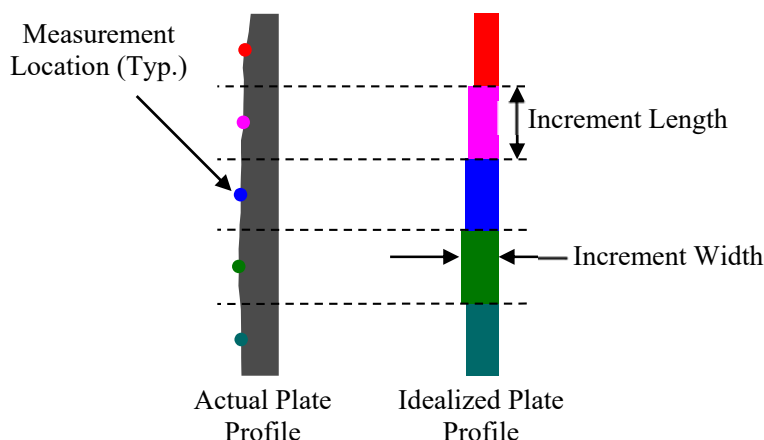


Figure 6.5 Illustration of the method used to calculate remaining member area from field measurements

6.2.5 Inspection Results

6.2.5.1 Task 1

In Task 1, the inspectors were asked to identify the most critical section for estimating the remaining capacity of the tension chord within the inspection limits. Four of the five inspectors successfully completed Task 1. The remaining inspector provided thickness measurements along the length of the chord, but did not provide enough detail to determine the remaining section at any single location. A summary of the results for this task is shown in Table 6.2. Based on the measurements provided, the remaining area estimates ranged from 32.4 in² to 35 in². Assuming an original gross cross section of 34.7 in² based on the construction plans, the percent “loss” estimates range from -6.4% to +0.9%.

The quickest inspector completed Task 1 in 17 minutes and the slowest inspector required 47 minutes. Although a time limit of 30 minutes had been set for this task, two inspectors were allowed to exceed this to provide useable data. All of the inspectors used an ultrasonic thickness gauge to complete this task. Four of the five inspectors used a measuring device (folding ruler or tape measure) and a camera during this task. One inspector used digital calipers and another inspector used a hammer.

Table 6.2. Summary of Task 1 results

	Inspector 1	Inspector 2	Inspector 3	Inspector 4	Inspector 5
Remaining Member Area (in ²)	32.4	INCOMPLETE	34.6	33.1	35
Duration (min.)	47	30	24	39	17
Location of Critical Section (in inches measured from Joint L2)	129	INCOMPLETE	17.25	130	120

Figure 6.6 shows the locations along the chord that the inspectors identified as the most critical section for determining remaining capacity. Three of the four inspectors that successfully completed Task 1 identified a critical section within a 10-inch region of the chord. This region of the chord had experienced significant section loss and distortion due to pack rust near the top edge of the cover plates as shown in Figure 6.7. The fourth inspector determined that the critical section was beneath the batten plate located closest to joint L2.

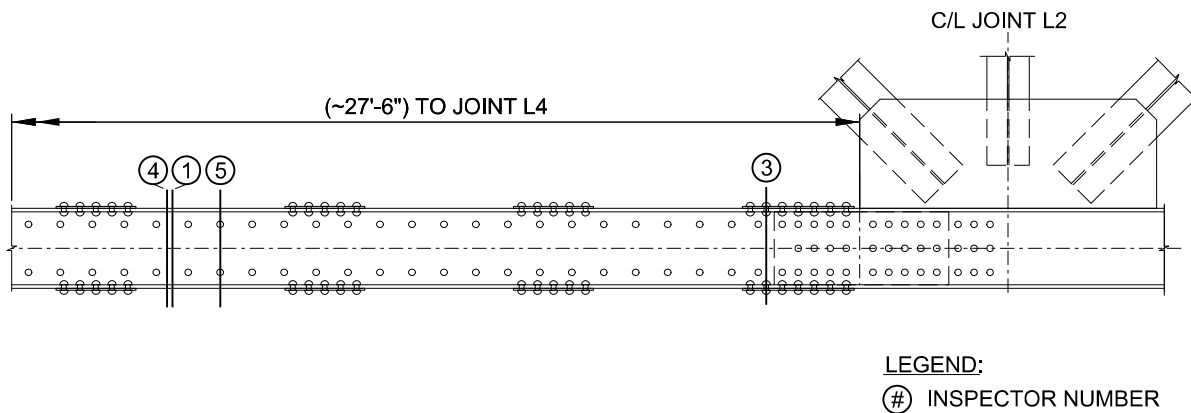


Figure 6.6 Critical section locations identified by the inspectors during Task 1



Figure 6.7 Truss chord near the critical section identified by three inspectors

One of the largest sources of variability within this task was the location and quantity of thickness measurements recorded by the inspectors. Only two of the five inspectors recorded thickness measurements on all four components (two channels and two cover plates) of the built-up chord. Inspector 1 provided measurements at approximately 1-inch intervals along the depth of each component and Inspector 3 provided measurements at approximately 2-1/2-inch intervals along the components. In contrast, Inspector 4 provided seven thickness measurements along just one of the channels and one of the cover plates and Inspector 5 provided three thickness measurements along one of the channels and no thickness measurements along either of cover plate. Since the amount of section loss in the chord is relatively small, especially in the channel components, the lack of measurements had a limited influence on the remaining area calculations. However, in members with more severe section loss, the difference in measurement frequency may cause significant variability in area estimates.

Although not necessary for this task, only two of the five inspectors provided notes or measurements related to the distortion in the cover plate. In the vicinity of the critical sections selected by Inspectors 1, 4, and 5, the thickness of the pack rust is more than 1 inch at the top and bottom of the member. For tension members, distortion due to pack rust is typically ignored during engineering evaluations since design standards allow for areas of localized yielding and tension

members are self-stabilizing. Therefore, the lack of measurements or notes would likely not affect the load rating [53]. However, for compression members, the distortion could pose more significant issues and may need to be accounted for by the load rating engineer. Distortion in the cover plates may reduce their effectiveness and lower the overall capacity of the member [53]. In this case, the inspector should include more detailed notes, photos, and measurements. For this inspection, the inspectors were not explicitly told that this was tension member, but were told it was a fracture critical lower truss chord which implies that it is a tension member. It is possible they understood that the distortion measurements were not needed for the load rating of a tension member or it is possible that they were not aware of the importance for any loading type. In general, inspectors should be instructed to err on the side of providing too much information and allow the rating engineers to determine what condition information should be included in their analysis.

6.2.5.2 Task 2

In Task 2, the inspectors were instructed to provide thickness measurements for the truss chord at the two locations shown in Figure 6.8.

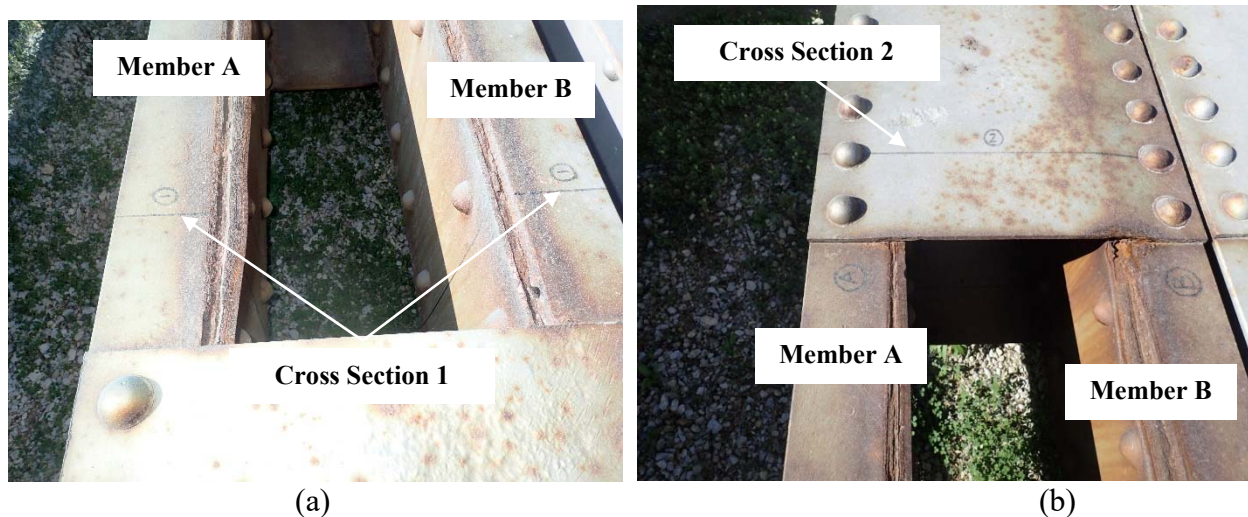


Figure 6.8 Truss chord at (a) Cross Section 1 and (b) Cross Section 2

Following the inspections, reference values were developed by the author after disassembling the chord. Measurements were taken with an ultrasonic thickness gauge at 1/2-inch increments along each cover plate and channel and the remaining area was calculated using the middle Riemann sum approach discussed in Section 6.2.4. In total, the reference measurements were taken at a

minimum of 6 locations along each flange, 28 locations along the channel web, and 31 locations along the cover plate. A summary of member areas estimated from the reference measurements is provided in Table 6.3. The remaining area determined from the thickness measurements was 33.27 in² at Cross Section 1 and 33.63 in² at Cross Section 2. The disassembled pieces of the truss are shown in Figure 6.9.

Table 6.3 Summary of reference areas

Member Description	Cross Section 1		Cross Section 2	
	Member A Areas (in ²)	Member B Areas (in ²)	Member A Areas (in ²)	Member B Areas (in ²)
Web	7.70	8.01	8.09	8.05
Top Flange	1.83	1.83	1.85	1.83
Bottom Flange	1.81	1.83	1.84	1.83
Channel (Web + Flanges)	11.33	11.68	11.78	11.71
Cover Plate	4.99	5.27	5.24	4.90
Channel + Cover Plate	16.32	16.94	17.02	16.61
Total Cross Section	33.27		33.63	



Figure 6.9 Disassembled pieces from Cross Sections 1 and 2

All five inspectors successfully completed Task 2. A summary of the results for this task is shown in Table 6.4. At Cross Section 1, the remaining area estimates ranged from 33.42 in² to 35.66 in². At Cross Section 2, the remaining area estimates ranged from 34.16 in² to 35.70 in². Compared to

the reference values, all inspectors overestimated the remaining area at both locations as shown in Figure 6.10. The reference area calculated based on the reference measurements is shown on the x-axis and the estimated area calculated based on the inspector's measurements is shown on the y-axis. The diagonal 1:1 reference line represents exact agreement between the reference area and the estimated area. All of the estimated areas plot above the 1:1 line indicating that the estimated areas exceed the reference area. The percent error ranged from 0.5% to 7.2% at Cross Section 1 and from 1.0% to 6.2% at Cross Section 2.

The duration of this task ranged from 5 minutes to 31 minutes. The average time to complete this task was 19 minutes. The quickest inspector measured the remaining thickness at Cross Section 2 as part of Task 1, and so they did not repeat that as part of Task 2. All of the inspectors used an ultrasonic thickness gauge to complete this task. Four of the five inspectors used a measuring device (tape measure or folding ruler) during this task and one inspector used a flashlight to during this task.

Table 6.4 Summary of Task 2 results

	Inspector 1	Inspector 2	Inspector 3	Inspector 4	Inspector 5	Reference
Cross Section 1 (in ²)	33.4	34.3	34.6	35.7	35.1	33.3
Cross Section 2 (in ²)	34.3	34.9	34.6	35.7	34.0	33.6
Duration (min.)	29	31	5	19	13	N/A

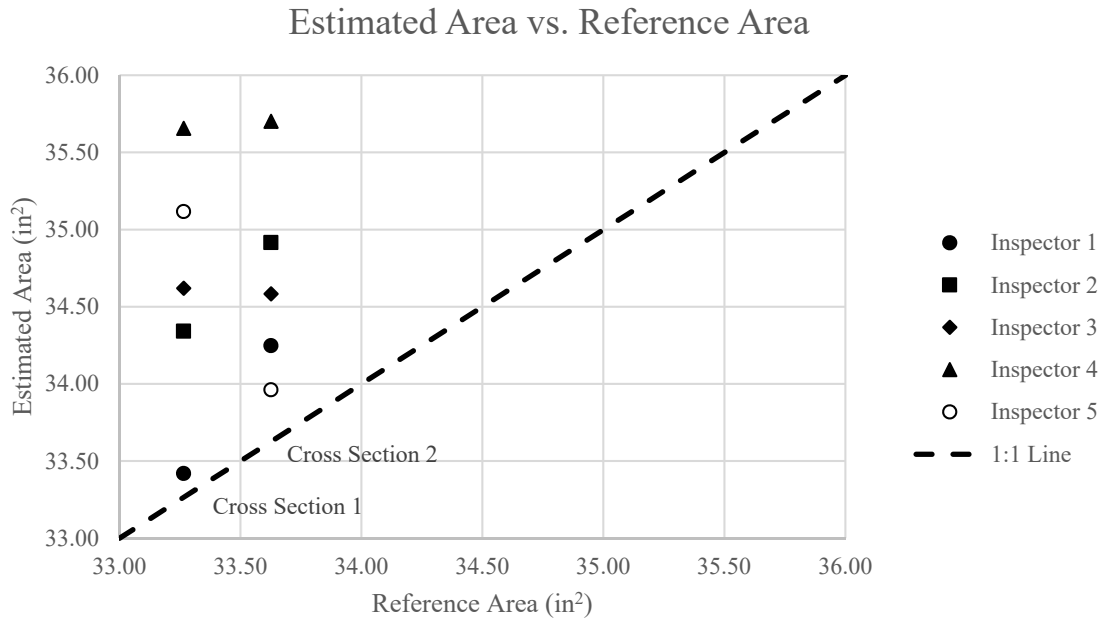


Figure 6.10 Estimated area plotted against reference area for Cross Section 1 and 2

Similar to Task 1, the location and number of thickness measurements varied significantly among the inspectors. Again, Inspector 1 provided measurements at approximately 1-inch intervals along the depth of each component and Inspector 3 provided measurements at approximately 2-1/2-inch intervals. Inspector 2 took measurements at approximately 5-inch intervals along the channels and 2-1/2-inch intervals along the cover plates. Inspector 4 recorded the thickness of the channels at 1-1/2-intervals but took only one measurement along the cover plates. Inspector 5 recorded the thickness of the channels and cover plates at a single location.

Considering the chord components separately provides a clearer look at the variability in the measurements. Measurement statistics for the cover plates are provided for Cross Sections 1 and 2 in Table 6.5 and Table 6.6, respectively. For each inspector, these tables show the cover plate area calculated from the inspector's measurements, the minimum recorded thickness measurement, and the number of recorded thickness measurements. The variability in remaining area estimates, as indicated by the standard deviation, was largest at Cover Plate B in Cross Section 2 where the smallest estimate of the remaining area of this member was 4.9 in² and the largest estimate was 6.27 in². The average of these measurements was 5.4 in² and the standard deviation was 0.61 in². This cross section was difficult to inspect because it was beneath the batten plate near Joint L2 and

this plate had experienced complete section loss near the top edge as shown in Figure 6.11. Conversely, the variability in remaining area was smallest at Cover Plate A in Cross Section 2. The remaining area estimates ranged from 5.3 in² to 5.65 in² with a mean of 0.5 in² and a standard deviation of 0.15 in². Although this plate was also obscured by the batten plate, it was in relatively good condition and inspectors were able to provide a reasonable estimate of the remaining thickness without performing a careful visual inspection.

Table 6.5 Measurement statistics for the cover plates at Cross Section 1

	Inspector 1	Inspector 2	Inspector 3	Inspector 4	Inspector 5	Reference
Calculated Area, in ² (Plate A/Plate B)	5.11/4.92	5.1/5.29	5.19/5.36	5.63/5.63	5.58/5.81	4.99/5.27
Min. Thickness Measurement, in. (Plate A/Plate B)	0.04/0.13	0.160/0.265	0.207/0.329	0.375/0.375	0.378/0.394	0.12/0.26
Number of Measurements (Plate A/Plate B)	16/16	6/6	6/6	1/1	1/1	31/32

Table 6.6 Measurements statistics for the cover plates at Cross Section 2

	Inspector 1	Inspector 2	Inspector 3	Inspector 4	Inspector 5	Reference
Calculated Area, in ² (Plate A/Plate B)	5.36/4.9	5.3/5.4	5.52/4.73	5.63/5.63	5.65/6.27	5.24/4.9
Min. Thickness Measurement, in. (Plate A/Plate B)	0.31/0.16	0.305/0.275	0.329/0	0.375/0.375	0.383/0.425	0.30/0
Number of Measurements (Plate A/Plate B)	16/16	6/5	6/6	1/1	1/1	31/31



Figure 6.11 Cover plates at Cross Section 2

Table 6.7 shows the percent error for each cover plate area estimate. Errors exceeding 5% are shown in red. As expected, the area calculated from the inspector's measurements approaches the reference area as the number of measurements increase. In general, Inspectors 1, 2, and 3 provided sufficient thickness measurements such that the estimated remaining area was within 5% of the reference area. Of note, only one of the five inspectors recorded the depth of the cover plates, 14-3/4 inches, on the inspection forms. Although the construction plans call for 15-inch by 3/8-inch web plates, the measured depth of these plates in their current condition is between 14-3/4 and 14-7/8 inches. If no depth dimension was recorded, the plan dimension was used to calculate the remaining area.

Table 6.7 Percent error for cover plate area estimates

	Inspector 1	Inspector 2	Inspector 3	Inspector 4	Inspector 5
Cover Plate A, Cross Section 1	2.4%	2.3%	4.0%	12.8%	11.8%
Cover Plate B, Cross Section 1	-6.6%	0.3%	1.7%	6.8%	10.3%
Cover Plate A, Cross Section 2	2.3%	1.1%	5.2%	7.3%	7.8%
Cover Plate B, Cross Section 2	0.1%	10.2%	-3.5%	14.8%	28.0%
Average Absolute Error	2.8%	3.5%	3.6%	10.4%	14.5%

Table 6.8 through Table 6.12 present the descriptive statistics for the areas estimated from the inspectors' measurements. These tables include the reference area, along with the average estimated area, the standard deviation and coefficient of variation (COV) of the estimates, the minimum and maximum estimated areas, and the number of inspectors that provided thickness measurements (n). In all but one instance, the average estimated area exceeded the reference area.

Table 6.8 Descriptive statistics for Member A at Cross Section 1

	Cover Plate	Channel Web	Channel Top Flange	Channel Bottom Flange	Total Channel	Total Member
Reference Area (in ²)	4.99	7.70	1.83	1.81	11.3	16.3
Average Area (in ²)	5.32	8.07	1.91	1.94	11.9	17.3
Standard Deviation of Area (in ²)	0.258	0.225	0.125	0.069	0.158	0.379
COV	0.049	0.028	0.065	0.035	0.013	0.022
Minimum Area (in ²)	5.10	7.86	1.70	1.83	11.8	16.9
Maximum Area (in ²)	5.63	8.33	2.03	2.03	12.2	17.8
n	5	5	5	5	5	5

Table 6.9 Descriptive statistics for Member B at Cross Section 1

	Cover Plate	Channel Web	Channel Top Flange	Channel Bottom Flange	Total Channel	Total Member
Reference Area (in ²)	5.27	8.01	1.83	1.83	11.7	16.9
Average Area (in ²)	5.40	8.10	1.95	1.93	12.0	17.4
Standard Deviation of Area (in ²)	0.341	0.238	0.056	0.087	0.252	0.516
COV	0.063	0.029	0.029	0.045	0.021	0.030
Minimum Area (in ²)	4.92	7.69	1.87	1.79	11.6	16.5
Maximum Area (in ²)	5.81	8.30	2.03	2.03	12.2	17.8
n	5	5	5	5	5	5

Table 6.10 Descriptive statistics for Member A at Cross Section 2

	Cover Plate	Channel Web	Channel Top Flange	Channel Bottom Flange	Total Channel	Total Member
Reference Area (in ²)	5.24	8.09	1.85	1.84	11.8	17.0
Average Area (in ²)	5.49	8.20	1.85	1.85	11.9	17.4
Standard Deviation of Area (in ²)	0.155	0.109	0.264	0.268	0.476	0.436
COV	0.028	0.013	0.143	0.145	0.040	0.025
Minimum Area (in ²)	5.30	8.07	1.38	1.37	11.1	16.7
Maximum Area (in ²)	5.65	8.31	2.03	2.03	12.2	17.8
n	5	5	5	5	5	5

Table 6.11 Descriptive statistics for Member A at Cross Section 2

	Cover Plate	Channel Web	Channel Top Flange	Channel Bottom Flange	Total Channel	Total Member
Reference Area (in ²)	4.90	8.05	1.83	1.83	11.7	16.6
Average Area (in ²)	5.38	8.25	1.81	1.84	11.9	17.3
Standard Deviation of Area (in ²)	0.613	0.119	0.352	0.251	0.526	0.422
COV	0.114	0.014	0.194	0.136	0.044	0.024
Minimum Area (in ²)	4.73	8.12	1.19	1.39	11.0	16.9
Maximum Area (in ²)	6.27	8.40	2.03	1.96	12.2	17.9
n	5	5	5	5	5	5

Table 6.12 Descriptive statistics for the truss chord at Cross Sections 1 and 2

	Cross Section 1	Cross Section 2
Reference Area (in ²)	33.3	33.6
Average Area (in ²)	34.6	34.7
Standard Deviation of Area (in ²)	0.842	0.673
COV	0.024	0.019
Minimum Area (in ²)	33.4	34.0
Maximum Area (in ²)	35.7	35.7
n	5	5

A one sample *t*-test was used to determine if the average areas estimated from the inspectors' measurements were statistically different from the reference areas. The *t*-statistic, *t*, can be calculated using Equation 6.1:

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} \quad (6.1)$$

where μ is the reference mean, \bar{X} is the sample mean, *s* is the sample standard deviation, and *n* is the sample size [58]. The *t*-statistic is assumed to have a *t* distribution with the degrees of freedom equal to the number of observations minus one. The probability value, or *p*-value, expressing the probability that the average area determined based on the inspection results is not different from the reference area can be calculated from the *t* distribution table [58]. A low *p*-value indicates that there is a low likelihood that the sample area is equal to the proposed mean. Table 6.13 shows the *p*-values for member areas at Cross Section 1 and Cross Section 2. The *p*-values less than 5% are

shown in bold since this was the threshold used to indicate significance. Based on these results, the average area estimated by this population of inspectors was statistically different from the reference area for the cover plates at two locations and the channel at one location. Additionally, the average area estimated for the truss chord was statistically different from the reference area at both cross sections.

Table 6.13 p -values for member areas at Cross Sections 1 and 2

Location	Cover Plate	Channel Web	Channel Top Flange	Channel Bottom Flange	Total Channel	Total Member
Cross Section 1	0.022					
Member A	0.045	0.021	0.188	0.012	0.001	0.005
Member B	0.433	0.458	0.010	0.059	0.054	0.132
Cross Section 2	0.025					
Member A	0.023	0.086	0.987	0.948	0.599	0.131
Member B	0.151	0.017	0.989	0.909	0.443	0.022

6.2.5.3 Task 3

In Task 3, the inspectors were asked to indicate the critical section for measuring section loss within the 28-inch region of the chord between Batten Plates C and D shown in Figure 6.12. The inspectors were not required to report any thickness measurements at this location, although they could take measurements to identify the section. A summary of the results for this task is shown in Table 6.14. The duration of this task ranged from 2 minutes to 11 minutes. All of the inspectors used an ultrasonic thickness gauge and a measuring device (folding ruler or tape measure) to complete this task. One inspector used a broom during this task.

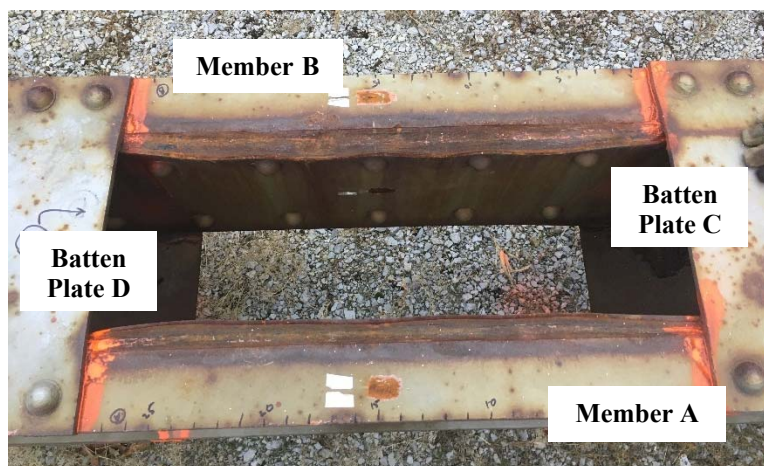


Figure 6.12 Limits of inspection for Task 3

Table 6.14 Summary of Task 3 Results

	Inspector 1	Inspector 2	Inspector 3	Inspector 4	Inspector 5
Location of Critical Section (in inches measured from Batten Plate C)	15	27	9	6	18.5
Duration (min.)	10	11	3	6	2

Each inspector identified a different location as the critical section, as shown in Figure 6.13. One inspector rationalized that the rivets appeared to be stretched and there was a gouge in the channel web at the location they identified as the critical section. Another inspector selected the critical section based on the amount of distortion at the top of the cover plate. A third inspector asserted that there was not much difference in thickness throughout the inspection region, so they selected the midpoint because it would be the worst case for buckling (note that the inspectors were told that this was a fracture critical bottom chord, implying that it is a tension member). The proctor observed that the majority of the inspectors only considered the thickness loss and/or distortion near the top of the cover plate. However, since this is fairly consistent through the region, it was not a particularly effective or efficient way to identify the section with the most section loss. Instead, the inspectors could have considered the thickness along the bottom of the cover plates as this varied between 0.28 inch and 0.375 inch: Since greater distortion generally indicates more section loss, this could have been identified without taking thickness measurements. The author determined the critical section to be located 25 inches from Batten Plate D similar to the section identified by Inspector 2.

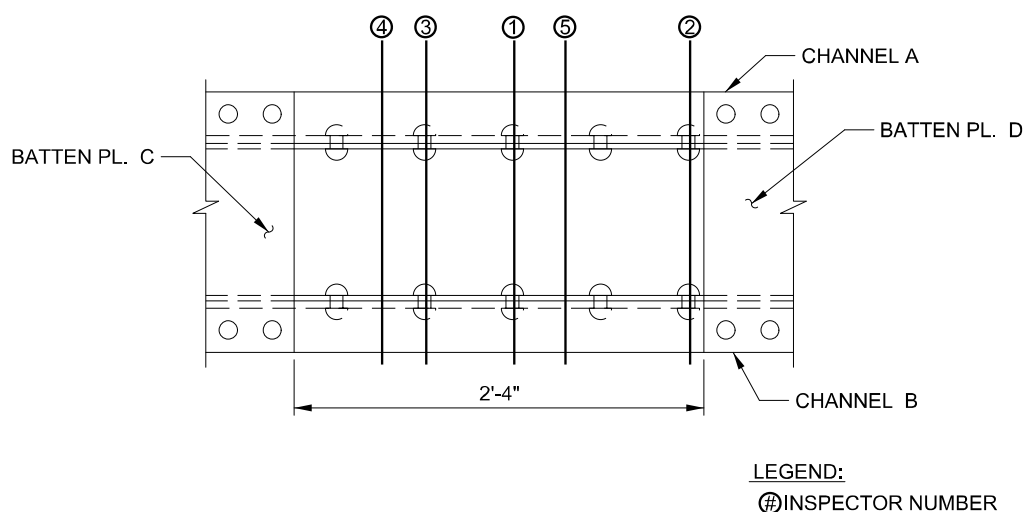


Figure 6.13 Critical section locations identified by the inspectors during Task 3

It is interesting to note that although this task was limited to 28 inches of the chord, the results do not reflect increasing agreement among the inspectors as compared to Task 1. In both tasks, three inspectors identified critical sections within a 10-inch region of the member. This suggests that inspectors applied similar broad reasoning to identify the critical region of the specimen, but their precise reasoning, used to identify the exact location, differed. For instance, in Task 1, the inspectors identified the critical region based mainly on the degree of section loss and/or distortion along the top edge of the web plates. However, in Task 3, the distortion and section loss was relatively uniform, and so this reasoning was less effective. Instead, inspectors used additional considerations, such as the likely failure mechanism or condition of the rivets and channel member, to identify the exact location of the critical section.

6.2.6 Recommendations

The following recommendations to improve the consistency and quality of visual inspections of corroded steel bridge members were developed based on the results from this round robin.

- Inspectors should receive training specific to corrosion inspection and evaluation. Instructions on how to properly calibrate and use an ultrasonic thickness gauge should be provided. Inspectors need improved instructions on what information is required by load rating engineers, and how these requirements vary based on bridge and member type.
- Thickness measurement of members with moderate deterioration should be taken at intervals not to exceed 3 inches, and no fewer than three measurements (outer edges and midpoint) should be recorded. Additional thickness measurements may be needed for critical members or members with more severe deterioration. Inspectors should clearly record the location of each measurement or the distance between measurements on their inspection forms.
- Inspectors should record dimensional measurements for all members to verify construction plans. Nominal (no deterioration) thickness measurements should also be recorded for reference.
- Before recording thickness measurements, inspectors should visually inspect the full length of the member. Inspectors should use typical inspection tools, including a flashlight, to perform this inspection. After the general inspection, they should focus

on critical regions of the member and areas that are prone to corrosion and pack rust or showing the most signs of deterioration.

6.3 Load Rating

A small pool of load rating engineers was invited to load rate the same portion of the truss chord that was used in the inspection round robin. A benchmark set of inspection data was provided to ensure that each load rater was working from the same information. The load raters were asked to determine the inventory level load rating factor for the Strength I limit state and the remaining fatigue life based on the information provided. The load raters' calculations were compared to the theoretical load rating determined by the author and to each other to evaluate the variability in interpreting inspection findings and applying code requirements.

6.3.1 Load Rating Scenario

The load ratings were completed at the participants' convenience between June 2018 and August 2018. The load raters were provided with the same load rating procedure, inspection report, construction plans, exit survey, and blank load rating summary sheet via e-mail. After completing the load rating, the load rater was asked to complete a written exit survey to gather information on their education, experience, and training. The load rating report, calculations, and exit form were returned via e-mail. The load rating procedures, inspection report, and exit survey are provided in Appendix D.

The evaluation was divided into two separate tasks that could be completed in any order. In Task 1, the load raters were asked to calculate the inventory level load rating factor for the Strength I limit state in the as-built (undamaged) condition and the as-inspected (damaged) condition. In Task 2, the load raters were asked to calculate the remaining fatigue life in the as-inspected condition. The load rating procedures specified that the load rating was to be completed using the load and resistance factor rating method in accordance with the 2nd Edition of the *AASHTO Manual for Bridge Evaluation* (MBE), including the 2016 interim revisions [2]. The load raters were allowed to use any other references or computer software available to them.

- These loads are unfactored. They include the distribution factor and dynamic load allowance.
- $(ADTT)_{SL} = 1500$ and it is assumed that $(ADTT)_{SL}$ is constant through the life of the bridge.

The loads were determined using a 2D SAP2000 model of Span 16, as shown in Figure 6.15. The total dead load was calculated from the construction plans and applied at the top chord joints. The predicted dead load in the member under consideration, L_2L_4 , was 320 kips. This compared well with the design dead load of 335 kips specified on the construction plans. For consistency, the design dead load was used in this exercise. Influence lines were used to determine the live load effects on member L_2L_4 . The predicted Strength I and Fatigue live loads for HL-93 loading were 322 kips and 104 kips, respectively. The fatigue live load was artificially increased to 134 kips for this exercise to produce a finite fatigue life. To validate the model, the live load effects from the H-20 loading were determined in accordance with the 1931 *AASHTO Standard Specifications* since this was the governing specification at the time of design [81]. The predicted live load was 150.7 kips and the predicted impact load was 30.1 kips. These values are nearly identical to the design live loads of 150.3 kips and the design impact load of 29.7 kips specified on the construction plans.

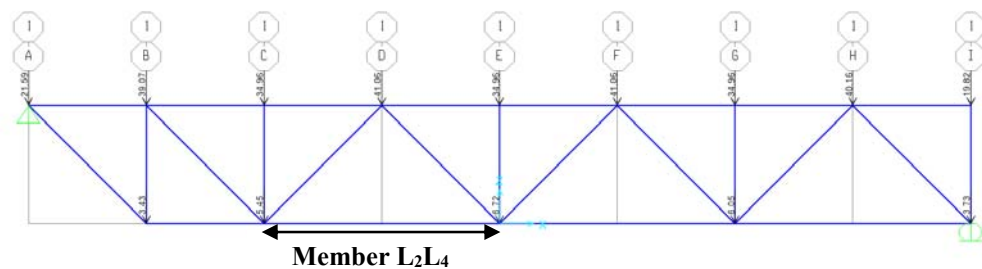


Figure 6.15 2D SAP model of Span 16

6.3.2 Engineer Demographics

Four engineers (3 males and 1 female) participated in this study. Two of the load raters were from federal agencies, one was from a state department of transportation, and one worked for a private engineering firm. Participating engineers were based in three states, Virginia (2), Indiana (1), and Minnesota (1), although these engineers may load rate bridges around the country. The inspectors ranged in age from 29 to 56 and their load rating experience varied from 0 years to 10 years. All four load raters were professional engineers with post-secondary degrees in civil engineering.

Three of the four load raters had completed the FHWA/NHI 4-day course *Fundamentals of LRFR and Applications of LRFR for Bridge Superstructures*. The engineers had performed between 0 and 50 load ratings during the 12 months prior to their participation in the study and the percentage of work time spent performing load ratings ranged from 0% to 90%. Three of the four load raters had previously load rated a steel truss bridge, although for one of the participants, this was the only bridge that they had ever load rated. In addition to load rating experience, two of the load raters also had inspection experience and/or experience designing steel bridges for new construction. Select load rater demographics are summarized in Table 6.15.

Table 6.15 Load rater demographics

	Load Rater 1	Load Rater 2	Load Rater 3	Load Rater 4
Age	40	42	56	29
Employment Sector	Federal Agency	Federal Agency	State DOT	Private Consultant
Years of Load Rating Experience	< 1	10	8	4
Professional Licensure	PE (Civil)	PE (Civil)	PE (Civil)	PE (Civil)
Education	Doctor of Philosophy	Master's Degree	Bachelor's Degree	Master's Degree
Number of load ratings in previous 12 months	0	5	30	50

6.3.3 Load Rating Results

The reference load rating, included in Appendix D, was developed in accordance with the 2nd Edition of the MBE and the 7th edition of the *AASHTO LRFD Bridge Design Specifications* (LRFD BDS) [54]. All four participants also used the MBE and the LRFD BDS to complete this load rating. One load rater used the 3rd edition of the MBE instead of the 2nd edition. Additionally, one load rater also referenced a state specific load rating manual and class notes. Three of the load raters used Microsoft Excel to complete the load rating and one used PTC Mathcad.

The load raters were asked to self-report the amount of time they spent on this evaluation. Three of the participants reported that it took them 8 hours or less to complete the evaluation. One load rater reported that the load rating took 50 hours. This load rater also reported that they used SAP2000 to verify the specified loads. Therefore, it is likely that the load rating portion of the

exercise required only a small percentage of the total reported time, although this is not known for certain.

6.3.3.1 Material Strength Assumptions

The load raters were required to make an assumption regarding the material strength (yield and tensile) of the truss chord members since it was not specified on the construction plans. It was expected that the load raters would use the year of construction to determine the minimum mechanical properties of the steel. This information is available from a variety of sources including the MBE, the *AISC Rehabilitation and Retrofit Guide* [82], and withdrawn ASTM specifications. The reference load rating assumed a yield strength of 33 ksi and an ultimate strength of 66 ksi based on Table 6A.6.2.1-1 in the MBE for construction between 1936 and 1963. (Note, the average yield and tensile strengths of the truss chord member were determined to be 39.5 ksi and 68 ksi, respectively, through material testing performed as part of a separate study [83].)

The material strengths assumed by the load raters are summarized in Table 6.16. All of the participants assumed a yield strength of 33 ksi, while the assumed tensile strength varied between 52 ksi and 66 ksi. Only two of the load raters provided a reference to support their assumption with one citing the MBE and the other citing the 1939 edition of ASTM A7. However, the engineer that referenced the MBE assumed a tensile strength of 52 ksi which corresponds to construction before 1905.

Table 6.16 Material strength assumptions

	Load Rater 1	Load Rater 2	Load Rater 3	Load Rater 4	Reference
Assumed Yield Strength (ksi)	33	33	33	33	33
Assumed Tensile Strength (ksi)	60	66	60	52	66
Reference	ASTM A7-39	Not cited	Not cited	MBE Table 6A.6.2.1-1	MBE Table 6A.6.2.1-1

6.3.3.2 Gross and Net Section Area Calculations

Before calculating the load rating factors and the remaining fatigue life, the load raters needed to determine the gross section and net section areas of the truss chord. In the as-built (undamaged) condition, this is relatively straightforward and can be calculated from the construction plans and

handbook properties. In the as-inspected (damaged) condition, the load raters were required to consider the inspection findings provided in the inspection report to determine the appropriate section areas.

The gross and net section areas in the as-built condition are summarized in Table 6.17. Estimates for the gross section area varied from 34.7 in² to 118.3 in² and estimates of the net section area varied from 27.3 in² to 33.2 in². The reference values were 34.7 in² and 29.0 in² for the gross section area and the net section area, respectively.

Table 6.17 As-built gross and net section area estimates

	Load Rater 1	Load Rater 2	Load Rater 3	Load Rater 4	Reference
Gross Section Area (in ²)	34.7	118	34.9	35.3	34.7
Net Section Area (in ²)	28.8	28.7	33.2	27.3	29.0

There is reasonable agreement among the areas calculated by the load raters, with the exception of the gross section area reported by Load Rater 2. The gross section area determined by Load Rater 2 is more than three times the reference value, but since Load Rater 2 did not provide any calculations to support this value, the cause of the discrepancy could not be identified. The following list summarizes some of the other sources of variability in these values. No distinction is made between assumptions or inaccuracies which yield a conservative estimate of member area and those which lead to an unconservative estimate.

- Channel Area. Two load raters assumed an area of 11.8 in² for the C15x40 channel. This is the value given in current steel handbooks, but is slightly larger than the value given in steel handbooks at the time of construction. The other load raters and the reference value used an area of 11.7 in² as specified in the 3rd edition of the *AISC Steel Construction Manual*, originally released in 1937 [84].
- Number of holes in the net section. The number of rivet holes through the cross section varies along member L₂L₄, as shown in Figure 6.2. The majority of the member includes four rivets connecting the cover plates to the channel webs. At the batten plates, there are an additional four rivets connecting the batten plates to the channel flanges. Near the joints, there is a section with six holes through the webs, but the partially developed 1/2-inch

splice plates compensate for the two additional holes. In their net section calculations, two load raters assumed six holes through the web and four holes through the flange, but they did not include the developed portion of the splice plate. One load rater assumed a single hole through each web and no holes through the flange. The final load rater and the reference value calculated the net section based on eight holes through the cross section (four holes through the flanges and four holes through the webs).

- Rivet and hole diameter. The construction plans specify a 7/8-inch diameter rivet for connections in the main truss members. Still, one load rater assumed the rivet diameter was 3/4 inch. Two load raters and the reference value assumed the hole diameter was 1/16 inch greater than the rivet diameter. One load rater assumed the hole diameter matched the rivet diameter and another load rater assumed that the hole diameter was 1/8 inch greater than the rivet diameter.

The gross and net section areas in the as-inspected condition are summarized in Table 6.18. Estimates for the gross section area varied from 31.2 in² to 114.9 in² and estimates of the net section area varied from 25.3 in² to 28.7 in². The reference values were 32.5 in² and 27.9 in² for the gross section area and the net section area, respectively.

Table 6.18. Summary of as-inspected gross and net section area estimates

	Load Rater 1	Load Rater 2	Load Rater 3	Load Rater 4	Reference
Gross Section Area (in ²)	32.9	115	31.2	32.3	32.5
Governing cross section used to determine gross section area	Cross Section 1	Cross Section 1	Cross Section 1	Cross Section 1	Cross Section 1
Net Section Area (in ²)	27.3	25.3	Not Specified	28.7	27.9
Governing cross section used to determine net section area	Cross Section 1	Cross Section 1	Not Specified	Cross Section 1	Cross Section 3

Again, there is reasonable agreement among the values determined by the load raters, with the exception of the gross section area determined by Load Rater 2. The error in the as-built calculation was carried forward as the losses reported at Cross Section 1 were simply subtracted from the as-built gross section area. All four engineers used Cross Section 1 as the governing cross section, and two of the load raters provided calculations at all three locations to support this.

The following list summarizes some of the other sources of variability in these values. No distinction is made between assumptions or inaccuracies which yield a conservative estimate of area and those which lead to an unconservative estimate.

- Material losses in the cover plate. The load raters employed different methods to account for the material loss and distortion in the cover plates. One load rater used the simple average of the thickness measurements to determine the remaining area, and two load raters used a weighted average of the thickness measurements. The final load rater completely neglected the area of the cover plates that had been distorted by pack rust, but did not consider the thickness losses along the full depth of the plates. The reference areas were calculated using the middle Riemann sum approach discussed in Section 6.2.4.
- Material losses in the channel. The inspection report included thickness measurements for Channel B at Cross Section 1. In all other locations, no losses were reported. One load rater used the average of the thickness measurements to determine the remaining area of this channel, and one load rater used a weighted average of the thickness measurements. The other two load raters ignored the losses in the channel at Cross Section 1. The reference values accounted for the losses in this channel similar to the cover plates, although the losses in this member were very small and likely within the mill tolerance for the rolled shape.
- Number of holes in the net section. Although all three load raters that estimated the net section area in the as-inspected condition assumed that Cross Section 1 was the governing cross section, each assumed a different number of holes through the cross section. Load Rater 1 assumed eight holes (four flange, four web), Load Rater 2 assumed ten holes (four flange, six web), and Load Rater 4 assumed four holes (zero flange, four web). The reference values were calculated assuming four holes through the cross section (zero flange, four web) at Cross Sections 1 and 2 and eight holes through the cross section (four flange, four web) at Cross Section 3. With this assumption, the smallest gross section was calculated at Cross Section 1 and the smallest net section was calculated at Cross Section 3.
- Rivet and rivet hole diameter. As discussed above for the as-built condition.

6.3.3.3 Task 1

In Task 1, the load raters were asked to determine the member capacity and inventory load rating factor for the Strength I limit state in the as-built and as-inspected conditions. Using their previous assumptions about material strength and the calculated gross and net section areas, the load raters were able to determine the member's resistance to yielding and fracture. All of the load raters used the equations given in Section 6.8.2 of the LRFD BDS to determine the tensile resistance of the member and the equations from Section 6A.4.2 of the MBE to calculate the member capacity and load rating factor. The member capacity and rating factor results are summarized in Table 6.19 and Table 6.20 for the as-built and as-inspected conditions, respectively.

In the as-built condition, estimates for member capacity ranged from 923 kips to 1363 kips and estimates for the load rating factor varied from 0.9 to 1.68. The reference capacity was 978 kips and the corresponding rating factor was 0.99. Three of the four load raters determined that the capacity was governed by yielding on the gross section. Load Rater 2, who reported the greatest member capacity and highest rating factor, determined that fracture on the net section was the governing limit state. This was due to the previously discussed overestimation in the gross section area.

Table 6.19 Summary of Task 1 results (as-built condition)

	Load Rater 1	Load Rater 2	Load Rater 3	Load Rater 4	Reference
Governing limit state	Yielding on the gross section	Fracture on the net section	Yielding on the gross section	Yielding on the gross section	Yielding on the gross section
Member Capacity (kips)	923	1363	929	996	978
Inventory Rating Factor	0.9	1.68	0.9	1.02	0.99

In the as-inspected condition, estimates for member capacity ranged from 732 kips to 1022 kips and estimates for the rating factor varied from 0.56 to 1.07. The reference capacity was 866 kips and the corresponding rating factor was 0.79. Three of the four load raters determined that the capacity was governed by yielding on the gross section. Again, there is reasonable agreement among the values determined by the load raters, with the exception Load Rater 2. The MBE incorporates some of the recommendations from NCHRP Report 333 into the commentary for

Section 6A.6.5 titled “Effects of Deterioration on Load Rating” [2]. Specific to steel tension members that have experienced localized corrosion, the MBE allows that yielding on the reduced net section area may be considered the governing limit state for simplicity. However, this conservative approach was not adopted by any of the load raters.

Table 6.20 Summary of Task 1 results (as-inspected condition)

	Load Rater 1	Load Rater 2	Load Rater 3	Load Rater 4	Reference
Governing limit state	Yielding on the gross section	Fracture on the net section	Yielding on the gross section	Yielding on the gross section	Yielding on the gross section
Member Capacity (kips)	877	1022	831	732	866
Inventory Rating Factor	0.81	1.07	0.73	0.56	0.79

In addition to variability caused by the material strength and area estimates discussed above, the following list summarizes some of the other sources of variability in these values. No distinction is made between assumptions or inaccuracies which yield a conservative estimate of member capacity and those which lead to an unconservative estimate.

- Assumed value for R_p . The equation for tensile resistance to fracture on the net section in the LRFD BDS includes a hole reduction factor, R_p , to account for reduced fracture resistance in the vicinity of holes that were punched full size [78, Eq. 6.8.2.1-2]. Two of the load raters assumed R_p was equal to 0.9, as specified for holes punched full size. One inspector assumed that R_p was equal to 1.0, as specified for holes that are drilled full size or sub-punched and reamed. One inspector did not include this factor in the equation and so a value of 1.0 was used by default. The reference value assumed that R_p was equal to 1.0 based on a construction plan note which reads, “General reaming will be required as per M.H.D [Minnesota Highway Department] Specifications 2407 3E5a”.
- Assumed values for ϕ_c and ϕ_s . The equation for member capacity in the strength limit state from the MBE includes a condition state factor, ϕ_c , and a system factor, ϕ_s [2, Eq. 6A.4.2.1-2]. These factors are applied along with the strength resistance factor from the LRFD BDS to the nominal member capacity. The condition factor takes a value between 0.85 and 1.0 and is intended to account for the increased uncertainty in the capacity of deteriorated members and the high likelihood of additional deterioration before the next inspection.

The system factor also takes a value between 0.85 and 1.0 and is intended to account for reduced redundancy in specific superstructure types. Additionally, the MBE stipulates that the product of the condition factor and the system factor ($\phi_C \times \phi_S$) need not be taken as less than 0.85 [2, Eq. 6A.4.2.1-3]. The reference member capacity in the as-built condition was calculated based on a condition factor of 1.0, specified for members in good condition, and a system factor of 0.9, specified for riveted members in two-girder/truss/arch bridges. The reference member capacity in the as-inspected condition was calculated based on a condition factor of 0.85, specified for members in poor condition, and a system factor of 0.9. Since the product of 0.85 and 0.9 is 0.77, a factor of 0.85 was instead applied to the design capacity in the as-inspected condition. The values assumed by the load raters for these factors are shown in Table 6.21. The only factor that the load raters were in complete agreement on was the condition factor in the as-inspected condition. For all the other factors, at least two different values were used by the load raters. Additionally, two load raters did not apply the lower limit to the product of the condition and system factors even though it was warranted.

Table 6.21 Assumed condition and system factors

	Load Rater 1	Load Rater 2	Load Rater 3	Load Rater 4	Reference
Condition factor, as-built	0.85	1	0.85	1	1
System factor, as-built	1	0.9	0.85	0.9	0.9
Condition factor, as-inspected	0.85	0.85	0.85	0.85	0.85
System factor, as-inspected	1	0.9	0.85	0.85	0.9
Lower Limit Applied?	N/A	No	Yes	No	Yes

6.3.3.4 Task 2

In Task 2, the load raters were asked to record the governing fatigue category, the effective stress range, and the remaining fatigue life of the member. This evaluation was only performed for the as-inspected condition. The results for this task are summarized in Table 6.22.

Estimates for the effective stress range varied from 2.9 ksi to 3.7 ksi and the reference value was 3.6 ksi. One load rater assumed Category A as the governing fatigue category, two load raters

assumed Category C, and one load rater assumed Category D. The reference value was based on Category D. Finally, three of the four load raters determined that the member had infinite remaining fatigue life, while one load rater determined that the remaining fatigue life was just under 30 years. The reference remaining fatigue life was 32 years.

Table 6.22 Summary of Task 2 results

	Load Rater 1	Load Rater 2	Load Rater 3	Load Rater 4	Reference
Governing Fatigue Category	D	C	A	C	D
Effective Stress Range (ksi)	3.7	3.7	2.9	3.5	3.6
Remaining Fatigue Life (years)	29.6	Infinite	Infinite	Infinite	32

While Fatigue Category A is not appropriate for this member since it applies only to base metal with minimal surface roughness, the distinction between Category C and Category D is less clear. For design, the LRFD BDS assigns base metal at the net section of non-pretensioned, mechanically fastened joints to Category D [78, Table 6.6.1.2.3-1]. For evaluation, the MBE allows riveted connections to be considered as Category C details due to the internal redundancy of built-up members [2, Sections 7.2.1 and C7.2.1]. In the 2015 revisions to the 2nd edition of the MBE, AASHTO included an additional stipulation stating that the increase in fatigue life is not warranted for riveted members in “poor physical condition, such as with missing rivets or indications of punched holes” [2, Section 7.2.1]. This forces the load rater to make a judgement call between Category C and Category D based on the condition of the member. Although the MBE does not mention section loss or damage from pack rust, the member is in poor physical condition and so the increase in fatigue life may not be appropriate for this member. Based on this, the reference value assumed that Category D was the governing fatigue category for this member.

In addition to variability caused by the area estimates discussed above, the following list summarizes some of the other sources of variability in these values. No distinction is made between assumptions or inaccuracies which yield a conservative estimate of remaining fatigue life and those which lead to an unconservative estimate.

- Assumed value of R_R . The MBE equation for estimating total fatigue life includes a factor related to the probability of fatigue crack initiation, R_R [2, Eq. 7.2.5.1-1]. The MBE

includes four levels at which fatigue life can be estimated: minimum, Evaluation 1, Evaluation 2, and mean. The minimum expected fatigue life, used during design, provides the most conservative estimate while mean fatigue life yields the statistically most likely fatigue life. The Evaluation 1 and Evaluation 2 fatigue life estimates will fall between the minimum and the mean fatigue life estimates. The MBE provides only general guidance on selecting the appropriate fatigue life for evaluation, leaving the decision largely up to the engineer. In this study, Load Rater 1 reported the minimum fatigue life. Load Rater 2 elected to estimate the mean fatigue life, although they later determined the life to be infinite. The other two load raters did not provide any indication of which finite fatigue life level they would have used if they had found the member to have finite life. The reference value was calculated using the R_R factor for Evaluation 1 fatigue life. Although the chord may possess low toughness, by modern standards, the internal redundancy of the built-up member justifies accepting an increased probability of fatigue crack initiation. Additionally, the Evaluation 1 fatigue life corresponds to the evaluation life used in previous editions of the MBE and is the level most often recommended in state specific load rating manuals, if any recommendation is made [85]–[87]. The Ohio Department of Transportation recommends calculating the Evaluation 2 fatigue life [88].

- Application of R_p and R_s . The MBE equation for effective stress includes a partial load factor, R_s , and a multiple presence factor, R_p [2, Eq. 7.2.2-1]. Additionally, the partial load factor is included in the equation for maximum stress range [2, Section 7.2.4]. These factors are typically very close to 1.0, and so ignoring or misapplying them has negligible or no effect on the results. A few discrepancies in how these factors were applied were noted while reviewing the load raters' calculations and are noted here for completeness. One inspector had a minor typo in the equation for R_p , multiplying by the number of lanes instead of dividing. One load rater included both R_p and R_s in the calculation for maximum stress range. One load rater assumed R_p was equal to 1.0 and two load raters did not include R_p or R_s in any of the stress range calculations.
- Area used to determine live load stress range. One load rater used the gross section area instead of the net section area to determine live load stress range. The other three load raters and the reference value used the net section area to determine live load stress range.

- Typographical errors and different interpretations of the provided information. A number of other small errors or misinterpretations of the provided information were noted in the Task 2 calculations. There may have been ambiguity in how the background information was presented or conveyed to the participants, although there was no common mistake among all the load raters pointing to an obvious omission. In practice, these calculations would be subject to review from a higher level official, and it is likely that these errors would have been noticed and corrected. However, the fatigue life analysis is less straightforward than the strength evaluation, both in procedure and communication, and so minor mistakes are likely to occur more frequently. One load rater used a fatigue truck load of 135 kips, instead of 134 kips, and another reduced the given fatigue truck load by the multiple presence factor for a single lane (1.2) even though this factor was not included in the provided load. One load rater used the 3rd edition of the MBE which has adopted the larger load factors for the Fatigue I (0.8) and Fatigue II (1.75) limit states from the 8th edition of the LRFD BDS. One load rater used an area that did not match any of the previous calculations to determine the live load stress range. Another inspector assumed a traffic volume growth rate of 2%, even though the load rating procedures stated that the load raters were to assume a constant ADTT throughout the life of the bridge. This resulted in an underestimation of the number of cycles that had already been applied to the structure and an overestimation of the remaining fatigue life.

6.3.4 Recommendations

The following recommendations to improve the quality and consistency of load rating evaluations of corroded steel bridge members were developed based on the results from this round robin.

- Load raters should receive instruction specific to determining the gross section and net section areas in the as-inspected condition. A single method should be used by all load raters to ensure that results from one load rater can be compared to the results from another. This guidance should clearly state how the thickness measurements are used to determine remaining area (average, weighted average, etc.) and whether load raters should assume that the thickness measurements may apply anywhere along the member or at a specific location. The method used to calculate member area should inform how the inspectors collect and record measurements in the field.

- Load raters should be required to clearly document and cite all assumptions. If questions or discrepancies arise in the future, this information will make it easier to validate previous evaluation. Additionally, load raters should be encouraged to include sketches to support the area calculations. Sketches would be especially helpful in identifying where the net section area was calculated.
- State DOT manuals should provide clear guidance on issues not directly specified in the MBE, such as how to determine the appropriate fatigue category for riveted members or the recommended fatigue life evaluation level. Similarly, state DOT manuals should provide additional direction on addressing localized corrosion considering yielding on either the reduced gross section or the reduced net section. Examples for load rating members with moderate deterioration and section loss should be incorporated into the MBE or state specific load rating manuals.
- Load raters should have access to historical standards and specifications, both federal and local, and be required to use these when load rating older structures. Load raters should be encouraged to include the relevant pages from the historical documents, including the construction plans, in their calculation packages for future reference.
- The fatigue life evaluation method presented in the MBE is complicated. Small mistakes are common and so fatigue life evaluations should be carefully reviewed by a senior load rater. Load raters should be encouraged to familiarize themselves with the theory behind the equations and factors; simply “plugging and chugging” with the code equations may lead to mistakes.

6.4 Summary

To expand the findings from the visual inspections discussed in Chapter 3 and Chapter 5, a small inspection and load rating round robin was conducted using a bridge specimen extracted from a steel truss approach span in Winona, Minnesota. In addition to investigating the inspection techniques and variability in findings, this round robin considered the variability in engineering evaluations.

Five inspectors from diverse backgrounds were invited to perform a hands-on inspection of a portion of the lower truss chord to investigate the variability in thickness measurements recorded

during inspections of corroded steel bridge members. The variability was evaluated by comparing the inspectors' measurements to each other and to the reference measurements determined by the author. In Task 1, three of the five inspectors determined that the critical section was located within a 10-inch region of the 28 feet long specimen. This level of agreement indicates some consistency in the methods used by inspectors to identify where the most section loss has occurred along a bridge member. In Task 2, the inspectors provided thickness measurements at two predetermined locations. All five of the inspectors provided thickness measurements that overestimated the remaining member area and two of the inspectors did not record a sufficient number of measurements to provide a reasonably accurate estimate of the area (less than 5% error). In Task 3, all five inspectors identified a different critical section within a reduced region of the chord. Although this task was limited to 28 inches of the chord, the results do not reflect increasing agreement among the inspectors as compared to Task 1. In both tasks, three inspectors identified critical sections within a 10-inch region. Overall, the level of detail in the findings recorded by some of the inspectors suggests a general uncertainty about what information is needed from the field to support a load rating analysis. In practice, this uncertainty may result in the need for follow-up inspections to gather the necessary information.

Four engineers were invited to load rate the same portion of the truss chord used for the inspection round robin to investigate the variability in how inspection reports are interpreted and code requirements are applied for corroded steel bridge members. The variability in the results was evaluated by comparing them to each other and to the reference values calculated by the author. In Task 1, three of the four inspectors reported a load rating factor within 10% of the reference value for the truss chord in the as-built condition. In the as-inspected condition, two of the load raters reported a rating factor within 10% of the reference value. In Task 2, three of the four load raters reported an effective stress range within 5% of the reference value for the truss chord in the as-inspected condition. However, due to variability in determining the governing fatigue category, only one of the inspectors determined that the truss chord had a finite fatigue life.

Based on the results from these two studies, a number of recommendations were developed to improve the quality and consistency of the inspection and evaluation of corroded steel bridge members. Most critically, both inspectors and load raters should receive training specific to

inspecting and evaluating corroded steel bridge members. This training should provide a single method for determining the gross section and net section areas and discuss what information is needed from the inspection to support the engineering evaluation

7. SUMMARY AND CONCLUSIONS

This research had the objective of investigating the performance of bridge inspectors during visual inspection of steel bridge members and developing a benchmark measure of inspection capability. Although visual inspection has been the primary means of inspecting in-service bridges since 1968, only a small number of studies have attempted to quantify the reliability or accuracy of these inspections. This research aimed to address the knowledge gap surrounding visual inspection performance for steel bridges in order to support future advances in inspection and design procedures.

This research has shown that (1) there is significant variability in visual inspection performance, even in a tightly controlled environment, (2) it is difficult to predict inspection performance based on human and environmental factors expected to affect performance, (3) even long cracks (> 4 inches) may not be reliably detected under current inspection procedures, (4) UAS-assisted inspection with a general purpose UAS platform did not provide fatigue crack detection comparable to a hands-on inspection, and (5) there is variability in how engineers interpret inspection findings and apply code requirements while load rating in-service bridges.

Performance in these inspection tasks was similar in accuracy and consistency to other industrial inspection tasks and the results from this research confirm that visual inspection of a steel bridge is an inherently difficult task. It is not a matter of “trying harder”; the current training scheme, inspection procedures, and inspection equipment produce highly variable results and modifications are needed to improve performance. Strategies for improvement were developed based on the visual inspection literature and the quantitative and qualitative results from this research. Above all, establishment of a performance based qualification system for bridge inspectors is recommended to confirm that a satisfactory level of performance is consistently achieved in the field. Additionally, new workshops focused on the physical and mental aspects of visual inspection were created to address the disconnect between the theories taught in the classroom and their application in the field. Ultimately, the results from this research may serve as a benchmark against which proposed changes to inspection policies or technologies may be evaluated.

7.1 Summary of Principal Findings

The following summary provides the principal findings of this research.

7.1.1 Summary of Hands-on Inspection Findings

- Thirty (30) inspectors participated in the hands-on inspections of the POD specimens. The average detection rate for the 70 cracks located on the painted specimens was 65%. Detection rates ranged from 31% to 86%. Univariate analysis between detection rate and other factors, such as duration or temperature, revealed slight, but statistically significant, correlations. Detection rate increased with increasing inspection duration, temperature, and training, but decreased with increasing experience.
- The average number of false calls on the 147 painted specimens was 90. The number of false calls ranged from 14 to 268. Univariate analysis between the number of false calls and other factors, such as duration or temperature, did not reveal any statistically significant correlations. However, the multivariate analysis revealed that the number of false calls was related to the inspector's employment sector, the maximum wind speed on the day of the inspection, the use of a tape measure, and completion of the FWH *Introduction to Element Level Bridge Inspection* training course. Inspectors that used a tape measure made fewer false calls while inspectors that were employed by a private inspection/engineering firm, experienced higher wind speeds, and attended the element level inspection training course tended to make more false calls.
- For the majority of the inspectors, the likelihood of a crack being detected increased with crack length. However, this was not true for three of the inspectors.
- Probability of detection curves relating crack detection to crack length were generated in accordance with Military Handbook 1823a [17]. For all 30 inspectors and the full crack inventory, the 50% detection rate crack length was 1 inch and the 90% detection rate crack length was 5-1/2 inches. The data were too scattered to assign a 95% confidence bound.
- A random parameters binary logit model was used to identify variables beyond crack length that affected the probability that a crack would be detected. The analysis showed that crack type and length, inspection experience, inspection duration, and the elapsed time since the first hands-on inspection significantly impacted the likelihood of detection for an individual crack.

7.1.2 Summary of UAS-assisted Inspection Findings

- Four inspectors participated in the UAS-assisted field inspections of the POD specimens. They each inspected between 52 and 85 specimens. The average detection rate was 54% and the average number of false calls was 53. Based on the results from the 52 specimens that all four field inspectors inspected, the average detection rate during the UAS-assisted field inspections was found to be significantly lower than the average detection rate during the hands-on inspections, while there was no significant difference in the average number of false calls made during the two types of inspection. The greatest difference in performance was observed in the inspection of the overhead mounted welded cover plates.
- Over the four days of field inspection, the average flight time was five minutes and the maximum flight time was 18 minutes. Frequent breaks were required to change batteries, avoid gusting winds, restore GPS signal, etc. Cumulatively, these breaks caused significant delays during the UAS-assisted inspections, and none of the field inspectors had enough time to inspect all 147 specimens included in the hands-on inspections.
- Nineteen (19) inspectors participated in the UAS-assisted desk inspections of the POD specimens. The inspectors inspected 54 specimens using videos recorded during the field inspections. The average detection rate was 57% and the average number of false calls was 70. Detection rate varied significantly with the number of false calls. The average performance during the UAS-assisted desk inspections was not statistically different from the average performance during the field inspections, but it was significantly worse than the average performance during the hands-on inspections.

7.1.3 Summary of Truss Chord Inspection and Load Rating Findings

- Five inspectors participated in the hands-on inspection of the truss chord. There was little consistency among the inspectors in the quantity or location of thickness measurements. Even at the exact same location along the member, the number of recorded thickness readings ranged from 6 to 59. The difference between the reference area calculated by the author and the area estimated from the inspector's measurements decreased as the number of thickness measurements increased.
- The inspectors identified the location of the critical section with the same level of agreement in Task 1, which considered the full length of the specimen, and Task 3, which

was limited to 28 inches of the specimen. In both cases, the majority of the inspectors identified a critical section within a 10-inch region of the specimen. This suggests that the inspectors applied similar broad reasoning to identify the critical region, but their precise reasoning, used to identify the exact location of the critical section, differed.

- Four engineers participated in the load rating of the truss chord. There was little consistency among the load raters in the method used to calculate the gross section and net section areas. Three of the load raters used an average thickness (weighted or unweighted) to calculate the remaining area of the cover plates, while the fourth load rater ignored the thickness measurements and instead discounted the area of the cover plates that was distorted due to pack rust. Additionally, the inspectors applied different levels of conservatism in identifying where along the member the critical section occurred. Considering the same cross section from the inspection report, each load rater assumed a different number of holes when calculating the as-inspected net section area.
- The load raters demonstrated more uncertainty in determining the remaining fatigue life as compared to the strength evaluation. This uncertainty was reflected by an increase in the number of minor errors in the calculations and misinterpretations of the provided background information. Additionally, there were some inconsistencies in the application of the code, specifically the identification of the governing fatigue category and appropriate fatigue life level.

7.2 Recommendations for Future Research

While this research provided important findings and recommendations for visual inspection of steel bridge members, additional research is suggested to provide a more complete understanding of inspection performance for a variety of defect types and the effectiveness of suggested improvements to the current visual inspection procedures. Six topics for future research are proposed: (1) investigate the influence of specific variables on inspection performance, (2) investigate visual inspection performance for other defect types, (3) evaluate the influence of the Observational Skills training on inspection performance (4) collect additional data on visual inspection performance using inspection-specific UAS platforms, (5) establish reliability-based performance criteria and inspection intervals, and (6) collect additional data on the load rating of in-service steel bridge members.

The first research topic will investigate the influence of specific variables on inspection performance using smaller, focused test groups. Specifically, performance during nighttime inspections and warm weather inspections would be investigated. Additionally, in order to better understand the meaning of the false positives, a small subset of inspectors would be explicitly evaluated based on both their hit percentage and the number of false calls made. Finally, the effectiveness of specific flashlight types and other lesser user tools, such as telescoping inspection mirrors, head lamps, and magnifiers, would be investigated.

The second research topic will investigate visual inspection performance for other defect types. First, additional fatigue sensitive details, some with and some without fatigue cracks, would be included among the existing specimen inventory on the training structure. Second, the inspection course would be expanded to include other defect types, such as impact damage, coating failure, and loose or missing fasteners, using other bridge components from the S-BRITE Center. Ultimately, an attempt would be made to correlate inspection results on the POD test specimens with inspection performance on an actual bridge specimen.

The third research topic will deliver and evaluate the effectiveness of the Observational Skills training course. After delivering the training, the course attendees would be invited to perform a hands-on inspection of the POD specimens to assess the effect of the training on performance. Participants would include both inspectors that have already performed a hands-on inspection and those that have not. In this way, both overall performance and the change in performance could be evaluated.

The fourth research topic will investigate performance during UAS-assisted visual inspections using an inspection-specific UAS platform. Although the UAS platform used during this research was minimally acceptable, additional research would explore the improvement in fatigue crack detection, if any, using a high-end drone with enhanced navigation and imaging capabilities.

The fifth research topic involves the development of reliability-based inspection performance criteria and inspection intervals. Under the current fracture control plan (FCP), each aspect of fracture prevention – material, design, fabrication, and inspection – is governed independently by

a separate specification. Whereas, under the proposed integrated FCP, each aspect of fracture prevention would work cooperatively, thereby allowing one component to compensate for another to maintain an acceptable resistance to fracture [7]. For instance, low material toughness could be mitigated through more rigorous and/or frequent in-service inspections. The quantitative description of in-service inspection capability and reliability offered by the POD curves provides the information necessary to link material, design and fabrication requirements to inspection method and frequency. This research would involve identifying the performance metric(s) used to evaluate inspector performance, establishing the minimal acceptable level of performance, and determining the necessary crack types, quantities, and lengths to verify performance. The research would consider both quantitative and qualitative performance metrics. This research would also evaluate the change in performance after implementing performance testing for bridge inspectors.

The sixth research topic will extend and expand the load rating study. First, the current study would be extended by inviting additional engineers to complete the load rating evaluation discussed in Section 6.3. Then, the study would be expanded to include other members or connections in the truss and eventually to include members from other structure types. This research should focus on how engineers account for member deterioration in a load rating since none of the examples in the MBE address this issue.

REFERENCES

- [1] *National Bridge Inspection Standards*. Code of Federal Regulations Title 23 Part 650 Subpart C, 2009.
- [2] AASHTO, *The Manual for Bridge Evaluation*, 2nd ed. Washington, DC: American Association of State Highway and Transportation Officials, 2011.
- [3] FHWA, “National Bridge Inventory ASCII Files,” 2017. [Online]. Available: <https://www.fhwa.dot.gov/bridge/nbi/ascii2017.cfm>. [Accessed: 01-May-2018].
- [4] A. K. Agrawal and G. A. Washer, “Consistency of the New York State Bridge Inspection Program,” New York State Department of Transportation, Albany, NY, C-07-17, 2013.
- [5] M. Moore, B. Phares, B. Graybeal, D. Rolander, and G. Washer, “Reliability of Visual Inspection for Highway Bridges,” Federal Highway Administration, McLean, VA, FHWA-RD-01-020, 2001.
- [6] R. J. Connor, R. Dexter, and H. Mahmoud, “NCHRP Synthesis 354: Inspection and Management of Bridges with Fracture-Critical Details,” Transportation Research Board, Washington, DC, 2005.
- [7] R. Sherman, W. Collins, and R. Connor, “Towards an Integrated Fracture Control Plan,” in *Proceedings of the World Steel Bridge Symposium*, 2016.
- [8] H.-Y. Chung, L. Manuel, and K. H. Frank, “Optimal Inspection Scheduling of Steel Bridges Using Nondestructive Testing Techniques,” *Journal of Bridge Engineering*, vol. 11, no. 3, pp. 305–319, 2006.
- [9] NTSB, “Collapse of U.S. 35 Highway Bridge, Point Pleasant, West Virginia, December 15, 1967,” National Transportation Safety Board, Washington, DC, NTSB/HAR-71-1, 1970.
- [10] Federal-Aid Highway Act of 1968, Pub. L. No. 90-495, §26, 73 Stat. 145. 1968.
- [11] NTSB, “Collapse of a Suspended Span of Interstate Route 95 Highway Bridge over the Mianus River Greenwich, Connecticut, June 28, 1983,” National Transportation Safety Board, Washington, DC, NTSB/HAR-84/03, 1984.
- [12] Oregon DOT, *Bridge Inspection Program Manual*, Rev. 4. Oregon Department of Transportation, 2013.
- [13] Minnesota DOT, *Bridge and Structure Inspection Program Manual*. Minnesota Department of Transportation, 2014.

- [14] Indiana DOT, *Bridge Inspection Manual*. Indiana Department of Transportation, 2010.
- [15] G. A. Georgio, "Probability of Detection (PoD) Curves: Derivation, Applications and Limitations," Health and Safety Executive, Research Report 454, 2006.
- [16] R. G. Eastin, "A Critical Review of Strategies Used to Deal With Metal Fatigue," in *International Conference on Aeronautical Fatigue*, 2003.
- [17] Department of Defense, "MIL-HDBK-1823A: Nondestructive Evaluation System Reliability Assessment," 2009.
- [18] L. Gandossi and C. Annis, "Probability of Detection Curves: Statistical Best-Practices," European Network for Inspection and Qualification, ENIQ Report No. 41, 2010.
- [19] C. Annis, "Statistical Best-Practices for Building Probability of Detection (POD) Models," *R package mh1823, version 4.3.2*, 2016. [Online]. Available: <http://statisticalengineering.com/mh1823/index.html>. [Accessed: 01-Apr-2016].
- [20] J. E. See, "Visual Inspection: A Review of the Literature," Sandia National Laboratories, Albuquerque, NM, SAND2012-8590, 2012.
- [21] C. G. Drury and J. Watson, "Good Practices in Visual Inspection," 2002. [Online]. Available: https://www.faa.gov/about/initiatives/maintenance_hf/library/documents/#HumanFactorsMaintenance. [Accessed: 25-Jul-2017].
- [22] T. J. Gallwey, "Evaluation and Control of Industrial Inspection: Part I - Guidelines for the Practitioner," *International Journal of Industrial Ergonomics*, vol. 22, pp. 37–49, 1998.
- [23] C. G. Drury and J. L. Addison, "An Industrial Study of the Effects of Feedback and Fault Density on Inspection Performance," *Ergonomics*, vol. 16, no. 2, pp. 159–169, 1973.
- [24] E. D. Megaw, "Factors Affecting Visual Inspection Accuracy," *Applied Ergonomics*, vol. 10, no. 1, pp. 27–32, 1979.
- [25] T. J. Gallwey, "Evaluation and control of industrial inspection: Part II - The scientific basis for the guide," *International Journal of Industrial Ergonomics*, vol. 22, no. 1–2, pp. 51–65, 1998.
- [26] C. G. Drury, "The Effect of Speed of Working on Industrial Inspection Accuracy," *Applied Ergonomics*, vol. 4, no. 1, pp. 2–7, 1973.

- [27] A. Schwaninger, D. Hardmeier, J. Riegelning, and M. Martin, "Use It and Still Lose It? The Influence of Age and Job Experience on Detection Performance in X-Ray Screening," *GeroPsych*, vol. 23, no. 3, pp. 169–175, 2010.
- [28] J. E. See, "Visual Inspection Reliability for Precision Manufactured Parts," *Human Factors*, vol. 57, no. 8, pp. 1427–1442, 2015.
- [29] E. D. Megaw and J. Richardson, "Eye Movements and Industrial Inspection," *Applied Ergonomics*, vol. 10, no. 3, pp. 145–154, 1979.
- [30] L. A. Demsetz, R. Cario, and R. Schulte-Strathaus, "Inspection of Marine Structures," Ship Structure Committee, Washington, DC, SSC-389, 1996.
- [31] J. Leach and P. E. Morris, "Cognitive Factors in the Close Visual and Magnetic Particle Inspection of Welds Underwater," *Human Factors*, vol. 40, no. 2, pp. 187–197, 1998.
- [32] T. J. Gallwey, "Selection Tests for Visual Inspection on a Multiple Fault Type Task," *Ergonomics*, vol. 25, no. 11, pp. 1077–1092, 1982.
- [33] P. A. Hancock, "Sustained Attention Under Thermal Stress," *Psychological Bulletin*, vol. 99, no. 2, pp. 263–281, 1986.
- [34] C. G. Drury, "Human Factors in Aircraft Inspection," in *Proceedings of Aging Aircraft Fleets: Structural and Other Subsystem Aspects*, 2001.
- [35] A. K. Gramopadhye, C. G. Drury, and J. Sharit, "Feedback strategies for visual search in airframe structural inspection," *International Journal of Industrial Ergonomics*, vol. 19, no. 5, pp. 333–344, 1997.
- [36] E. I. Wiener, "Vigilance and Inspection," in *Sustained Attention in Human Performance*, J. S. Warm, Ed. New York: Wiley, 1984, pp. 207–246.
- [37] B. J. Hillman, R. G. Swensson, S. J. Hessel, D. E. Gerson, and P. G. Herman, "The Value of Consultation among Radiologists," *American Journal of Roentgenology*, vol. 127, no. 5, pp. 807–809, 1976.
- [38] FAA, "Unmanned Aircraft Systems (UAS) Frequently Asked Questions," 2018. [Online]. Available: <https://www.faa.gov/uas/faqs/>. [Accessed: 18-Feb-2018].
- [39] J. Zink and B. Lovelace, *Unmanned Aerial Vehicle Bridge Inspection Demonstration Project, Phase I Report*. St. Paul, MN: Minnesota Department of Transportation, 2015.
- [40] J. Wells and B. Lovelace, *Unmanned Aircraft System Bridge Inspection Demonstration Project, Phase II Report*. St. Paul, MN: Minnesota Department of Transportation, 2017.

- [41] J. Wells and B. Lovelace, “Improving the Quality of Bridge Inspections Using Unmanned Aircraft Systems,” Minnesota Department of Transportation, St. Paul, MN, MN/RC 2018-26, 2018.
- [42] L. D. Otero, N. Gagliardo, D. Dalli, W. H. Huang, and P. Costentino, “Proof of Concept for using Unmanned Aerial Vehicles for High Mast Pole and Bridge Inspections,” Florida Department of Transportation, Tallahassee, FL, 2015.
- [43] C. Brooks *et al.*, “Evaluating the Use of Unmanned Aerial Vehicles for Transportation Purposes,” Michigan Department of Transportation, Lansing, MI, RC-1616, 2015.
- [44] C. Brooks *et al.*, “Implementation of Unmanned Aerial Vehicles (UAVs) for Assessment of Transportation Infrastructure – Phase II,” Michigan Department of Transportation, Lansing, MI, SPR-1674, 2018.
- [45] S. Dorafshan, R. Thomas, and M. Maguire, “Fatigue Crack Detection Using Unmanned Aerial Systems in Fracture Critical Inspection of Steel Bridges,” *Journal of Bridge Engineering*, vol. 23, no. 10, 2018.
- [46] J. Seo, L. Duque, and J. Wacker, “Drone-enabled Bridge Inspection Methodology and Application,” *Automation in Construction*, vol. 94, pp. 112–126, 2018.
- [47] D. Gillins, C. Parrish, M. Gillins, and C. Simpson, “Eyes in the Sky: Bridge Inspections With Unmanned Aerial Vehicles,” Oregon Department of Transportation, Salem, OR, FHWA-RD-RD-18-11, 2018.
- [48] K. Cunningham *et al.*, “UAS-Based Inspection of the Placer River Trail Bridge: A Data-Driven Approach,” in *Structures Congress*, 2015, pp. 1649–1660.
- [49] A. Khaloo, D. Lattanzi, K. Cunningham, R. Dell’Andrea, and M. Riley, “Unmanned Aerial Vehicle Inspection of the Placer River Trail Bridge through Image-based 3D Modelling,” *Structure and Infrastructure Engineering*, vol. 14, no. 1, pp. 124–136, 2018.
- [50] L. R. Snyder, “Procedures to Simulate Fatigue Cracks in Steel Bridge Specimens for Use in a Probability of Detection Study,” M.S. thesis, Lyles School of Civil Engineering, Purdue University, 2015.
- [51] J. M. Whitehead, “Probability of Detection Study for Visual Inspection of Steel Bridges,” M.S. thesis, Lyles School of Civil Engineering, Purdue University, 2015.
- [52] F. W. Spencer, “Visual Inspection Research Project Report on Benchmark Inspections,” Federal Aviation Administration, Washington, DC, DOT/FAA/AR-96/65, 1996.

- [53] J. M. Kulicki, Z. Prucz, D. F. Sorgenfrei, and D. R. Mertz, "NCHRP Report 333: Guidelines for Evaluating Corrosion Effects in Existing Steel Bridges," Transportation Research Board, Washington, DC, 1990.
- [54] AASHTO, *LRFD Bridge Design Specifications*, 7th ed. Washington, DC: American Association of State Highway and Transportation Officials, 2014.
- [55] J. W. Fisher, E. J. Kaufmann, and A. W. Pense, "Effect of Corrosion on Crack Development and Fatigue Life," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1624, pp. 110–117, 1998.
- [56] "Purdue Univ, IN History | Weather Underground." [Online]. Available: <https://www.wunderground.com/history/daily/us/in/lafayette/KLAF/date/>. [Accessed: 30-Jan-2017].
- [57] C. Annis and L. Gandossi, "Influence of Sample Size and Other Factors on Hit/Miss Probability of Detection Curves," European Network for Inspection and Qualification, ENIQ Report No. 47, 2012.
- [58] S. P. Washington, M. G. Karlaftis, and F. L. Mannering, *Statistical and Econometric Methods for Transportation Data Analysis*, 2nd Ed. New York: CRC Press, 2011.
- [59] F. L. Mannering, V. Shankar, and C. R. Bhat, "Unobserved Heterogeneity and the Statistical Analysis of Highway Accident Data," *Analytic Methods in Accident Research*, vol. 11, pp. 1–16, 2016.
- [60] K. M. Ghylin, C. G. Drury, R. Batta, and L. Lin, "Temporal Effects in a Security Inspection Task: Breakdown of Performance Components," in *Human Factors and Ergonomics Society 51st Annual Meeting*, 2007.
- [61] P. Fish, C. Schroeder, R. J. Connor, and P. Sauser, *Fatigue and Fracture Library for the Inspection, Evaluation, and Repair of Vehicular Steel Bridges*. West Lafayette, IN: Purdue University, 2015.
- [62] G. Washer, R. Connor, and D. Looten, "Performance Testing of Inspectors to Improve the Quality of Nondestructive Testing," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2408, pp. 107–113, 2014.
- [63] R. E. Shaw Jr., "Ultrasonic Testing Procedures, Technician Skills, and Qualifications," *Journal of Materials in Civil Engineering*, vol. 14, no. 1, pp. 62–67, 2002.

- [64] L. A. Demsetz, R. Cario, R. T. Huang, and R. Schulte-Strathaus, "Factors Affecting Marine Structural Inspection Performance," in *Ship Structure Symposium*, 1996.
- [65] J. L. Toquam, F. A. Morris, and J. R. Griggs, "Basic Visual Observation Skills Training Course: Final Report," International Atomic Energy Agency, PNL-SA-26411, 1995.
- [66] J. L. Toquam and F. A. Morris, "Recommended Observational Skills Training for IAEA Safeguards Inspections," International Atomic Energy Agency, PNL-10186, 1994.
- [67] S. Koenig, L. Nickles, D. Kimbler, B. Melloy, and A. Gramopadhye, "Visual Inspection Simulator (VisInS): A Computer-based Inspection Simulation Tool for Off-Line Experimentation," in *Proceedings of Industrial Engineering Research Conference*, 1995.
- [68] A. K. Gramopadhye and K. Wilson, "Noise, Feedback Training, and Visual Inspection Performance," *International Journal of Industrial Ergonomics*, vol. 20, pp. 223–230, 1997.
- [69] D.J. Simons and C. F. Chabris, "Gorillas in our midst: Sustained Inattentional Blindness for Dynamic Events," *Perception*, vol. 28, pp. 1059–1074, 1999.
- [70] K. J. Blacker, S. Negoita, J. B. Ewen, and S. M. Courtney, "N-back versus Complex Span Working Memory Training Predictors of Near Transfer," *Journal of Cognitive Enhancement*, vol. 52, 2017.
- [71] G. Ganis and R. Kievit, "A New Set of Three-Dimensional Shapes for Investigating Mental Rotation Processes: Validation Data and Stimulus Set," *Journal of Open Psychology Data*, vol. 3, 2015.
- [72] M. Bazerman and D. Moore, *Judgement in Managerial Decision Making*, 8th ed. New York: Wiley, 2013.
- [73] H. J. Einhorn and R. M. Hogarth, "Confidence in Judgment: Persistence of the Illusion of Validity," *Psychological Review*, vol. 85, no. 5, pp. 395–416, 1978.
- [74] M. Plotnikov, D. Ni, and J. Collura, "The State of the Practice of Unmanned Aircraft Systems in State Departments of Transportation," in *Proceedings of the TRB 2018 Annual Meeting*, 2018.
- [75] S. G. Hart, "NASA-Task Load Index (NASA-TLX); 20 Years Later," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2006.
- [76] S. G. Hart and L. E. Straveland, "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research," *Advances in Psychology*, vol. 52, pp. 139–183, 1988.

- [77] J. S. Warm, W. N. Dember, and P. A. Hancock, "Vigilance and Workload in Automated Systems," in *Automatation and Human Performance: Theory and Applications*, M. Mouloua and R. Parasuraman, Eds. Mahwah, NJ: Lawrence Earlbaum Associates, 1996, pp. 183–200.
- [78] G. Morgenthal and N. Hallermann, "Quality Assessment of Unmanned Aerial Vehicle (UAV) Based Visual Inspection of Structures," *Advances in Structural Engineering*, vol. 17, no. 3, pp. 289–302, 2014.
- [79] *DJI Mavic Pro User Manual*, V2.0. DJI, 2017.
- [80] VideoLAN, "VLC media player." [Online]. Available: <https://www.videolan.org/vlc/index.html>. [Accessed: 19-Mar-2018].
- [81] J. M. Kulicki and D. R. Mertz, "Evolution of Vehicular Live Load Models During the Interstate Design Era and Beyond," in *50 Years of Interstate Structures: Past, Present, and Future. Transportation Research Board Circular E-C104*, Washington, DC: Transportation Research Board, 2006.
- [82] R. L. Brockenbrough, *AISC Rehabilitation and Retrofit Guide*. American Institute of Steel Construction, Inc., 2002.
- [83] J. B. Lloyd, "Internal Redundancy of Mechanically-Fastened Steel Built-Up Axially-Loaded Members," Ph.D dissertation, Lyles School of Civil Engineering, Purdue University, 2018.
- [84] AISC, *Manual of Steel Construction*, 3rd ed. New York, NY: American Institute of Steel Construction, 1937.
- [85] Minnesota DOT, *Bridge Load Rating and Evaluation Manual*. Minnesota Department of Transportation, 2018.
- [86] Connecticut DOT, *Bridge Load Rating Manual*. Connecticut Department of Transportation, 2018.
- [87] New York State DOT, *Bridge Inspection Manual*. New York State Department of Transportation, 2017.
- [88] Ohio DOT, *Bridge Design Manual*. Ohio Department of Transportation, 2019.

APPENDIX A. HANDS-ON INSPECTION DOCUMENTS

Confidentiality Agreement

CONFIDENTIALITY AGREEMENT

Purdue University S-BRITE Center: Probability of Detection Study

Principal Investigator/ S-BRITE Center Director: Dr. Robert J Connor

S-BRITE Center Research Engineer: Jason B. Lloyd, PE

Graduate Students: Leslie E. Campbell, PE

As a participant in this research study I understand that I may have access to confidential information regarding the test procedures, study layout and test specimens. By signing this statement, I am indicating my understanding of my responsibilities to maintain confidentiality and agree to the following:

- I understand that details about study procedure and test specimens are completely confidential.
- I agree not to discuss, divulge, publish, or otherwise make known to unauthorized persons, including any other INDOT inspectors or colleagues, or to the public any information obtained in the course of this research study without written permission from the S-BRITE Director.
- I understand that sharing any information with other participants in the study invalidates the results and hinders INDOT from improving bridge inspection procedures and inspection quality. I agree not to divulge or otherwise make known to unauthorized persons any of this information, unless specifically authorized to do so by approved protocol or by the local principal investigator acting in response to applicable law or court order, or public health or clinical need.
- I agree to notify the principal investigator immediately should I become aware of an actual breach of confidentiality or a situation which could potentially result in a breach, whether this be on my part or on the part of another person.

Signature	Date	Printed name

Signature of Purdue Official	Date	Printed name

Inspection Procedure
Probability of Detection Study
Performance Testing Procedure
Visual Inspection Testing
Purdue University S-BRITE Center
2016

This document describes the guidance and procedures for inspectors participating in the Purdue University S-BRITE Probability of Detection Study. This procedure is for Visual Inspection only.

Proctor Instructions - VI:

Read these instructions to the inspector prior to the test.

You are about to participate in a probability of detection study for visual inspection of steel bridges. It is critical to the study that all procedures and findings are kept **CONFIDENTIAL**. Before participating in this study, a signed confidentiality agreement is required. Likewise, your identity will remain confidential, unless requested by INDOT. The results from your specific participation will not be shared or be published. If you do not wish to sign the confidentiality agreement, you may excuse yourself at this time.

The objectives of this portion of study are as follows:

1. Establish the effectiveness of current visual inspection techniques for identifying cracks while inspecting steel bridges;
2. Establish the probability of detection based on current inspection practices;
3. Establish methods for improving probability of detection of cracks in steel bridges.

You are permitted to use any tools and equipment that you typically utilize during a visual inspection. Lunch will be provided during your full day of inspection. You should conduct this inspection in the same manner and with the same care as normally implemented in the field. During your inspection, breaks are allowed for fatigue or other reasons as you need. If you normally take a break during a certain time period, you should take that break during the study. Please notify the Proctor when you wish to take your break.

Inspection Scenarios:

This will be a hands-on inspection and is to be performed in a fashion consistent with a typical “fracture critical” inspection. You may use any tools you normally use for such an inspection. The inspection is only focused on the visual detection of cracks. The Proctor will operate the manlift and guide you through the course in the correct order.

You will be inspecting a portion of the 108 specimens suspended from the support fixture. Each specimen has two faces with a corresponding inspection form. These forms are contained in the binder provided. Most specimens also include a portion of a cover plate. Your binder includes a corresponding inspection form for each cover plate. A completed form is required for each specimen face and cover plate inspected.

Inspection Procedure

Probability of Detection Study

Performance Testing Procedure

Visual Inspection Testing

Purdue University S-BRITE Center

2016

Assumptions Regarding Specimens

1. The specimens are intended to represent 1960's to 1970's welded fabrication and weld quality. While weld quality may not meet modern standards, it is not the focus of the study. You are not required to comment on weld quality.
2. Assume the pieces suspended from the frame are fracture critical members.
3. Treat all gusset plates and vertical stiffeners as if bracing, floorbeams, diaphragms, or cross-bracing was attached. For example, many specimens include gusset plates and vertical stiffeners welded to webs. Therefore, you are to assume there are lateral members attached to these components whether they are on the interior or exterior faces of the specimens or near the top or bottom flange. The bracing members have not been included to facilitate your access.
4. Assume all cover plate terminations are subject to tensile stress ranges.
5. You should assume both flanges could be tension flanges.
6. The location of the specimen on the frame should not be used to "infer" the loading or stress state in the specimen. In other words, specimens installed near the ends of the support frame should not be viewed as being near a bearing. All specimens should be viewed as being subjected to the same stress state.
7. Interior and exterior specimens should be treated the same.
8. Both faces of each specimen should be treated the same.
9. Any specimen could have any type of crack or even multiple cracks.

Inspection Notes

If you detect a crack, please sketch it on the form and record its length and location. If no crack is found, the inspection form must still be completed indicating such. The order of specimen inspection will follow the order of the inspection forms in your binder. As stated, the Proctor will operate the manlift and guide you through the course in the correct order. Ensure that the Specimen ID on the form matches the ID on the specimen. At the end of your inspection, please submit the completed inspection booklet to the Proctor. Upon completion of the entire inspection course, please complete the Inspector Exit Form, which will be given to you by the Proctor.

Thank you for your participation and cooperation. Feel free to ask the Proctor any questions related to the study procedures at any time during this test.

Exit Survey

Probability of Detection Study Inspector Information



Please complete the following form. All information collected is used for research purposes only.

Inspector ID: _____ Inspection Date: _____
 Employer: _____ Years of Inspection Experience: _____
 Age: _____ Gender: _____
 Height: _____

- 1) Please list all tools you brought to complete this inspection.

- 2) Please list which tools you used during the inspection course.

- 3) Are there any additional tools you wish you would have had available?

- 4) Do you wear corrective lenses (glasses or contacts)? If yes, did you wear them during the inspection?

- 5) How many hands-on inspections have you completed in the last 12 months? How many of these inspections were performed on steel bridges?

- 6) Approximately how many routine inspections have you completed in the last 12 months? How many of these inspections were performed on steel bridges?

- 7) Please indicate which inspection training courses you have completed and the date of completion (estimates are acceptable for the dates):
 - Safety Inspection of In-Service Bridges (FHWA/NHI): _____
 - Bridge Inspection Refresher Training (FHWA/NHI): _____
 - Engineering Concepts for Bridge Inspectors (FHWA/NHI): _____
 - Underwater Bridge Inspection (FHWA/NHI): _____
 - Fracture Critical Inspection Techniques for Steel Bridges (FHWA/NHI): _____
 - Inspection and Maintenance of Ancillary Highway Structures (FHWA/NHI): _____
 - Introduction to Element Level Bridge Inspection (FHWA): _____
 - Inspecting Steel Bridges for Fatigue (Purdue/S-BRITE): _____

Others: _____

Exit Survey

Probability of Detection Study Inspector Information



8) Are you a certified welder or weld inspector?

Yes, I'm a CWI Yes, I'm a certified welder No Other (please specify)

9) Are you a registered professional engineer?

10) Approximately what percentage of your work time (on a per week basis) is spent performing bridge inspections?

11) What is the highest educational level that you have completed?

- _____ High School degree or equivalent
- _____ Trade School Degree
- _____ Associate's Degree
- _____ Bachelor's Degree in CIVIL ENGINEERING? Or OTHER? (circle one)
- _____ Master's Degree in CIVIL ENGINEERING? Or OTHER? (circle one)
- _____ PhD in CIVIL ENGINEERING? Or OTHER? (circle one)

12) How rushed did you feel during the inspection?

Not Rushed Moderately Rushed Very Rushed

13) How accessible were the bridge specimens?

Inaccessible Mostly Accessible Very Accessible

14) Did the presences of the observer affect your inspection practice?

Not affected Somewhat affected Very affected

Exit Survey

Probability of Detection Study Inspector Information



15) How did your effort level during this task compare to your effort level during a typical bridge inspection?

More effort

Similar effort

Less effort

16) How did your focus level during this inspection compare to your focus level during a typical bridge inspection?

More focused

Similar focus

Less focused

17) Overall, how did you feel (physically) during the inspection today?

Poor

Below Average

Average

Above Average

Great

18) Has anyone discussed anything related to this study prior to your participation?

Sample Inspection Forms

Specimen ID: 2TSP11-B

Inspector ID:

North

weld defects
off weld faces & toes

root cracks

1" long
crack from toe
of weld

weld stop/start
toe of weld

1 1/8" long crack

1" long
crack from toe
of weld

1 1/8" long crack

1 of 1

Sample Inspection Forms

Specimen ID: 2TSP20-B		Inspector ID:														
<div style="border: 1px solid black; width: 80%; margin: 0 auto; height: 300px; position: relative;"><div style="position: absolute; left: 50%; top: 50%; transform: translate(-50%, -50%); font-family: cursive; font-size: 2em;">NO CRACKS</div></div>																
<div style="border: 1px solid black; padding: 5px;">BOWEN LABORATORY PURDUE UNIVERSITY</div> <p>1040 SOUTH RIVER ROAD WEST LAFAYETTE, IN 47906 P: 765-494-2227 F: 765-494-9886</p>		<div>PROJECT: PROBABILITY OF DETECTION</div> <div>SHEET NOTES:</div>		<div>REVISIONS:</div> <table border="1" style="width: 100%; border-collapse: collapse;"><thead><tr><th style="text-align: center;">NO.</th><th style="text-align: center;">DATE</th><th style="text-align: center;">BY</th></tr></thead><tbody><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr></tbody></table> <div>DESIGNED BY:</div> <div>DRAWN BY:</div> <div>CHECKED BY:</div> <div>DATE:</div> <div>PROJECT NO.: SPR 3820</div>	NO.	DATE	BY									
NO.	DATE	BY														
		<div>SHEET TITLE:</div> <div style="text-align: right;">SHEET NO.: 1 of 1</div>														

APPENDIX B. UAS-ASSISTED DESK INSPECTION DOCUMENTS

Inspection Procedure

Probability of Detection Study

UAS Inspections: Instructions for Post-Inspection Review
Utah State University and Purdue University
2018

This document describes the procedures for inspectors participating in the Utah State University Probability of Detection Study at Purdue University's S-BRITE Center. This procedure is for Post-Inspection Review only.

Read these instructions before beginning the post-inspection review. If you have any questions, do not hesitate to contact any members of the research team (contact information is provided at the end of this document).

Introduction:

You are about to participate in a probability of detection study for visual inspection of steel bridges. It is critical to the study that all procedures and findings are kept confidential. Before participating in this study, please sign the provided confidentiality agreement and upload to the shared folder. Please do not share the videos or discuss your findings with other inspectors or personnel in your office. Likewise, your identity will remain confidential; the results from your specific participation will not be shared or be published.

The objectives of this portion of study are as follows:

1. Establish the effectiveness of unmanned aerial system (UAS) inspection for identifying cracks in steel bridges;
2. Evaluate the practicality of performing bridge inspections from the office using video captured with a UAS;
3. Establish methods for improving UAS inspection of steel bridges;
4. Identify the strengths and weaknesses of UAS inspection of steel bridges.

Inspection Scenario:

This inspection consists of the review of video footage of a steel bridge recorded with a UAS. The inspection is only focused on the visual detection of cracks. Even though the inspection will be performed from your office, it should be performed with the level of care and attention to detail normally implemented in the field. You will be responsible for reviewing the videos and completing the accompanying inspection forms.

The time limit for this inspection is 8 hours. Please monitor your time and at the end of 8 hours, return the inspection forms that have been completed. Please do your best to review all of the videos within this time frame, but it is acceptable if some videos are not reviewed.

(Note for inspectors that also performed the real-time inspection in the field:

For the majority of the specimens, the videos provided are the same ones you recorded during your inspection. In the case that a video was not captured, a video from another inspection has been provided. Your completed inspection binder from the field inspection is available for your reference while performing the post-inspection review. These forms are provided for reference only; please be sure to indicate all findings from the post-inspection review on the blank set of forms provided.)

Inspection Procedure

Probability of Detection Study UAS Inspections: Instructions for Post-Inspection Review Utah State University and Purdue University 2018

Inspection Notes

You will be inspecting a portion of the 108 girder specimens suspended from the support fixture shown in Figures 1 and 2 below. Each specimen has two faces with a corresponding inspection form. Most specimens also include a portion of a welded cover plate attached to the bottom flange and there are a series of riveted plates mounted to the vertical columns of the frame. A form has been provided for each of these plates.

The order of specimen inspection will follow the order of the inspection forms in your binder. The layout of the specimens is also provided for your reference. Each specimen has at least one accompanying video, and some specimens have multiple associated videos.

If you detect a crack, please sketch it on the form and record its length and location. If no crack is found, the inspection form must still be completed indicating such. Dimensioned drawings of the specimens are provided for your reference in estimating crack length. Also, please record the starting and ending time for each specimen in the spaces provided on the form. Times can be rounded to the nearest minute.

After completing the inspection of Specimen 2LSP9-B, please complete the NASA Task Load Index worksheet included with your inspection forms. This assessment is intended to provide a quantitative measure of workload.

Pre- and Post- Inspection Notes

Before beginning the review, a brief teleconference will be scheduled to go over the inspection procedure and answer questions. Please review the material provided prior to this call. A series of three "vision tests" will also be administered during this pre-inspection meeting. The videos for these three tests will be provided.

After finishing the inspection, please complete the Inspector Exit Form. Both the inspection booklet and the exit form should be uploaded to the shared folder or emailed to the Research Team.

Specimen Background:

1. The specimens are intended to represent 1960's to 1970's welded fabrication and weld quality. While weld quality may not meet modern standards, it is not the focus of the study. You are not required to comment on weld quality.
2. Assume the pieces suspended from the frame are fracture critical members.
3. Treat all gusset plates and vertical stiffeners as if bracing, floorbeams, diaphragms, or cross-bracing was attached. For example, many specimens include gusset plates and vertical stiffeners welded to webs. Therefore, you are to assume there are lateral members attached to these components whether they are on the interior or exterior faces of the specimens or near the top or bottom flange.
4. Assume all cover plate terminations are subject to tensile stress ranges.
5. You should assume both flanges could be tension flanges.

Inspection Procedure

Probability of Detection Study

UAS Inspections: Instructions for Post-Inspection Review

Utah State University and Purdue University

2018

6. The location of the specimen on the frame should not be used to “infer” the loading or stress state in the specimen. In other words, specimens installed near the ends of the support frame should not be viewed as being near a bearing. All specimens should be viewed as being subjected to the same stress state.
7. Interior and exterior specimens should be treated the same.
8. Both faces of each specimen should be treated the same.
9. Regardless of the orientation of the riveted plates, you should assume these are tension members. They are intended to represent a component in a built-up truss chord or a riveted cover plate with the load parallel to the longitudinal axis.
10. Any specimen could have any type of crack or even multiple cracks.

Thank you for your participation and cooperation. Please contact us with any questions.

Dr. Marc Maguire

[REDACTED]

Dr. Robert J. Connor

[REDACTED]

Sattar Dorafshan

[REDACTED]

Leslie Campbell

[REDACTED]

Inspection Procedure

Probability of Detection Study UAS Inspections: Instructions for Post-Inspection Review Utah State University and Purdue University 2018



Figure 1. POD Frame, looking east



Figure 2. Underside of POD Frame

Exit Survey



Probability of Detection Inspection Inspector Information



Please complete the following form. All information collected is used for research purposes only.

Inspector ID: _____ Inspection Date: _____
 Employer: _____ Years of Inspection Experience: _____
 Age: _____ Gender: _____

Please fill out the table below with the technical specifications of the screen(s) used to review the inspection footage. When applicable, information provided should reflect the current settings of your display, not the maximum capability of the system. If only a single display was used, please enter N/A in the fields for Display 2.

	Display 1 (Primary Display)	Display 2 (Secondary Display)
Screen Make and Model		
Video Card Make and Model		
Size (diagonal and aspect ratio)		
Resolution		
Refresh Rate		
Color Depth		
Scaling		
Brightness (-60 to 100)		
Contrast (40 to 100)		
Gamma (0.4 to 5.0)		
Hue (0 to 260)		
Saturation (-100 to 100)		

Please answer the following questions regarding the media player used to review the inspection footage.

- 1) What media player was used to review the inspection footage? Indicate the name and version (if available).

- 2) Is this software freely available for download on the internet? If not, how did you obtain it?

- 3) How satisfied were you with the quality of the playback?

Very Satisfied
Satisfied
Not Satisfied

Exit Survey



Probability of Detection Inspection Inspector Information



- 4) Place a check mark next to the features available in your media player.

Ability to rewind, fast-forward, and pause the video: _____

Ability to adjust the speed of playback: _____

Ability to zoom in and out on the video: _____

Ability to adjust the brightness of the video: _____

Ability to adjust the contrast of the video: _____

Ability to adjust the color of the video: _____

Ability to adjust the saturation of the video: _____

Ability to capture a still photo from the video: _____

Others: _____

- 5) Indicate the usefulness of the playback features for the review of the inspection footage.

Description of Activity	Usefulness Rating (1 - Not very useful, 5 – Very useful)					
Pause the video	1	2	3	4	5	Not Used
Rewind the video	1	2	3	4	5	Not Used
Fast forward the video	1	2	3	4	5	Not Used
Decrease the speed of playback	1	2	3	4	5	Not Used
Increase the speed of playback	1	2	3	4	5	Not Used
Zoom in or out on the video	1	2	3	4	5	Not Used
Adjust the brightness of the video	1	2	3	4	5	Not Used
Adjust the contrast of the video	1	2	3	4	5	Not Used
Adjust the color of the video	1	2	3	4	5	Not Used
Adjust the saturation of the video	1	2	3	4	5	Not Used
Capture a still photo from the video	1	2	3	4	5	Not Used

- 6) Are there any technical capabilities that your software lacked that you wish you would have had available during the post-inspection review?

Exit Survey



Probability of Detection Inspection Inspector Information



7) Are you a registered professional engineer or engineer-in-training?

Yes

No

If yes, please indicate your license below.

SE

PE (civil)

PE (other)

FE (civil)

FE (other)

Other

8) What is the highest educational level that you have completed?

_____ High School degree or equivalent

_____ Trade School Degree

_____ Associate's Degree

_____ Bachelor's Degree in CIVIL ENGINEERING? Or OTHER? (circle one)

_____ Master's Degree in CIVIL ENGINEERING? Or OTHER? (circle one)

_____ PhD in CIVIL ENGINEERING? Or OTHER? (circle one)

9) How did your effort level during this task compare to your effort level during a typical bridge inspection?

More effort

Similar effort

Less effort

10) How did your focus level during this task compare to your focus level during a typical bridge inspection?

More focused

Similar focus

Less focused

11) Do you think this inspection provided _____ quality as compared to a UAS inspection performed live in the field?

Worse

Better

Similar

Please briefly explain the reason(s) for this selection.

Exit Survey



Probability of Detection Inspection Inspector Information



Please answer the following questions regarding your background and inspection experience.

- 1) Have you previously participated in a UAS inspection of a steel bridge?

Yes

No

If yes, approximately how many?

- 2) Have you previously reviewed video footage captured during a UAS inspection of a steel bridge?

Yes

No

If yes, approximately how many?

- 3) Approximately how many hands-on, fracture critical inspections have you completed in the last 12 months?

- 4) Approximately how many routine inspections have you completed in the last 12 months?
What percentage of these were performed on steel bridges?

- 5) Approximately what percentage of your work time (on a per week basis) is spent performing bridge inspections?

- 6) Please indicate which inspection training courses you have completed and the date of completion (estimates are acceptable for the dates):

Safety Inspection of In-Service Bridges (FHWA/NHI): _____

Bridge Inspection Refresher Training (FHWA/NHI): _____

Engineering Concepts for Bridge Inspectors (FHWA/NHI): _____

Underwater Bridge Inspection (FHWA/NHI): _____

Fracture Critical Inspection Techniques for Steel Bridges (FHWA/NHI): _____

Inspection and Maintenance of Ancillary Highway Structures (FHWA/NHI): _____

Introduction to Element Level Bridge Inspection (FHWA): _____

Inspecting Steel Bridges for Fatigue (Purdue/S-BRITE): _____

Others: _____

Exit Survey



Probability of Detection Inspection Inspector Information



12) Do you think this inspection provided _____ quality as compared to an arm's length inspection?

Worse Better Similar

Please briefly explain the reason(s) for this selection.

13) Do you have any suggestions or recommendations to improve performance during a UAS inspection?

14) Has anyone discussed anything related to this study prior to your participation? Please indicate if you have previously completed a hands-on inspection or UAS inspection of the POD specimens.

Sample Inspection Forms

START TIME: <u>5:32</u>		END TIME: <u>5:37</u>																			
Specimen ID: 2TSP11-B		Inspector ID:																			
<div style="display: flex; align-items: center;"><div><p>BOWEN LABORATORY PURDUE UNIVERSITY</p><p>1040 SOUTH RIVER ROAD WEST LAFAYETTE, IN 47906 P: 765-494-2227 F: 765-494-6886</p></div></div>		<p>PROJECT: PROBABILITY OF DETECTION</p> <p>SHEET NOTES:</p>																			
<table border="1" style="width: 100%; border-collapse: collapse;"><thead><tr><th colspan="3">REVISIONS:</th></tr><tr><th>NO.</th><th>DATE</th><th>BY</th></tr></thead><tbody><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr></tbody></table> <p>DESIGNED BY: _____</p> <p>DRAWN BY: _____</p> <p>CHECKED BY: _____</p> <p>DATE: _____</p> <p>PROJECT NO.: <u>SPR 8830</u></p>		REVISIONS:			NO.	DATE	BY													<p>SHEET TITLE:</p> <p>SHEET NO.: 1 of 1</p>	
REVISIONS:																					
NO.	DATE	BY																			

Sample Inspection Forms

START TIME: 4:35PMEND TIME: 4:37:PM

Specimen ID: 2TSP20-B

Inspector ID:

No cracks observed.



1040 SOUTH RIVER ROAD
WEST LAFAYETTE, IN 47906
P: 765-494-2227 F: 765-494-8888

PROJECT:
**PROBABILITY OF
DETECTION**

SHEET NOTES:

REVISIONS:
NO. DATE BY

DESIGNED BY:
DRAWN BY:
CHECKED BY:
DATE:
PROJECT NO.: SPR 3820

SHEET TITLE:

SHEET NO.: 1 of 1

NASA-TLX Worksheet

At this time, please take a break from your inspection and complete the NASA Task Load Index worksheets.

NASA Task Load Index (NASA-TLX)¹

Instructions:

The evaluation you're about to perform is a technique that has been developed by NASA to assess the relative importance of size factors in determine how much workload you experienced while performing a task that you recently completed.

These six factors are defined below. Read through them to make sure you understand what each factor means.

Definitions:

Mental Demand (low/high)

How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, forgiving or exacting?

Physical Demand (low/high)

How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

Temporal Demand (low/high)

How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

Performance (good/poor)

How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?

Effort (low/high)

How hard did you have to work (mentally and physically) to accomplish your level of performance?

Frustration Level (low/high)

How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?

¹NASA TLX: Task Load Index. NASA, 15 Feb. 2017, <https://humansystems.arc.nasa.gov/groups/TLX/>.

NASA-TLX Worksheet

Pairwise Comparisons:

For each pair, circle the factor that was more important to your experience of the workload in the task that you recently performed.

Mental Demand	Performance
or	or
Effort	Temporal Demand
Effort	Mental Demand
or	or
Performance	Physical Demand
Frustration	Frustration
or	or
Mental Demand	Effort
Physical Demand	Physical Demand
or	or
Temporal Demand	Performance

NASA-TLX Worksheet

Performance	Temporal Demand
or	or
Mental Demand	Frustration
Effort	Temporal Demand
or	or
Physical Demand	Mental Demand
Temporal Demand	Performance
or	or
Effort	Frustration
Physical Demand	
or	
Frustration	

NASA-TLX Worksheet

Rating Scales:

For each of the six scales, evaluate the task you recently performed by indicating your experience at the accurate location on the scale. Each line has two endpoint descriptors that describe the scale.

Consider your responses carefully in distinguishing among the different task conditions, and consider each scale individually.

Mental Demand

How much mental and perceptual activity did you spend for this task?



Physical Demand

How much physical activity did you spend for this task?



Temporary Demand

How much time pressure did you feel in order to complete this task?



Performance

How successful do you think you were in accomplishing the goals of the task?



Effort

How hard did you have to work to accomplish your level of performance?



Frustration

How insecure, discouraged, irritated, stressed, and annoyed were you during this task?



Vision Test Worksheet

Jaeger Near Visual Acuity Test

What is the smallest paragraph you can read in the Jaeger Vision Test_1.jpg file? No. _____

Please type/write the paragraph here:

What is the smallest paragraph you can read in the Jaeger Vision Test_2.mp4 file? No. _____

Please type/write the paragraph here:

APPENDIX C. TRUSS CHORD INSPECTION DOCUMENTS

Inspection Procedure

Truss Chord Inspection Round Robin Visual Inspection Testing Procedure Purdue University 2018

This document provides the procedures and instructions for inspectors participating in the Truss Chord Inspection Round Robin at Purdue University.

Read these instructions to the inspector prior to the inspection -

Introduction:

You are about to participate in a study for visual inspection of steel bridges. It is critical to the study that all procedures and findings are kept **CONFIDENTIAL**. Likewise, your identity will remain confidential; the results from your specific participation will not be shared or be published.

The objectives of this portion of study are as follows:

1. Establish the variability in identifying the critical section and estimating section loss during hands-on, visual inspection of steel bridges
2. Establish methods for improving consistency in these measurements

Inspection Scenario:

This inspection is focused solely on the evaluation of corrosion and section loss in steel bridge members, NOT crack detection, although cracks may be present. You should conduct this inspection in the same manner and with the same care as normally implemented in the field. All inspection activities will be completed from the ground beneath or adjacent to the specimen. You may take breaks as needed throughout this inspection activity.

This inspection is divided into three separate tasks. The tasks must be completed in the specified order. Please notify the proctor when you complete each task so that the time may be recorded. At that time, the instructions for the next task will be provided.

At the end of your inspection, please submit the completed inspection forms to the proctor. Upon completion of the entire inspection course, please complete the Inspector Exit Form. Feel free to ask the proctor any questions related to the study procedures at any time during the inspection.

Member Background:

1. This member was removed from a bridge constructed in the early 1940s in the upper Midwest region of the US.
2. This member was a part of the bottom chord of the deck truss.
3. This member is fracture critical.

Inspection Procedure

Truss Chord Inspection Round Robin

Visual Inspection Testing Procedure
Purdue University
2018

Inspection Tasks:

Task 1:

In Task 1, you are asked to identify the MOST critical location for determining remaining capacity along the full length of the member. At the critical location, you are asked to provide sufficient information (measurements, notes, sketches, photos, etc.) such that a load rating engineer could determine the load rating for this member. You may use any tools that you brought with you to complete this task. The location of the critical section and your inspection findings should be recorded on the forms provided. Photos may be taken with the camera provided. The time limit for this task is 30 minutes.

Task 2:


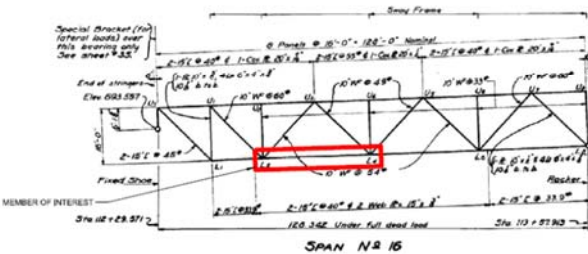

In Task 2, you are asked to estimate the remaining thickness of the member at the two (2) locations identified on your inspection form. You may use any tools that you brought with you to complete this task. If needed, please use the provided ultrasonic thickness gauge. Your measurements should be recorded in the tables provided. Please label the location of the measurements on the cross section sketches. There is no time limit for this task.

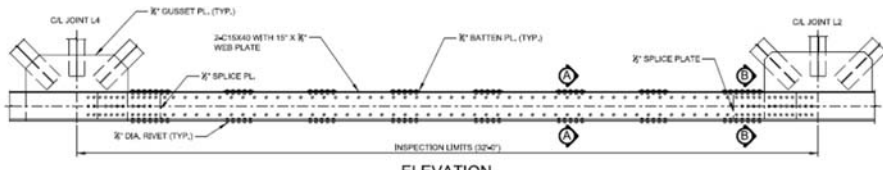
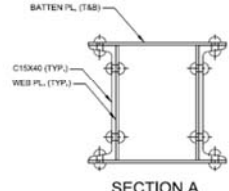
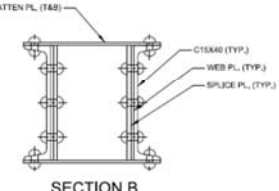

Task 3:

In Task 3, you are asked to identify the MOST critical location for determining remaining capacity within the region identified on your inspection form. The location of the critical section should be recorded on the inspection form. You do NOT need to record the remaining thickness at this location. You may use any tools that you brought with you to complete this task. An ultrasonic thickness gauge is available for your use, if needed. There is no time limit for this task.

Thank you for your participation and cooperation.

Inspection Forms

 <p>SPAN OF INTEREST</p>			
<p align="center">ELEVATION OF FULL BRIDGE</p>			
 <p align="center">ELEVATION OF SPAN OF INTEREST</p>			
 <p>1040 SOUTH REVER ROAD WEST LAFAYETTE, IN 47906 P: 765-464-2227 F: 765-464-6986</p>	<p>PROJECT:</p> <p align="center">TRUSS INSPECTION ROUND ROBIN</p>		<p>SHEET TITLE:</p>
	<p>INSPECTOR ID: _____</p> <p>INSPECTION DATE: _____</p> <p>PROCTOR INITIALS: _____</p>		<p>SHEET NO.:</p>

 <p align="center">ELEVATION SCALE: 1" = 4'</p>			
 <p align="center">SECTION A SCALE: 1" = 1'-0"</p>		 <p align="center">SECTION B SCALE: 1" = 1'-0"</p>	
 <p>1040 SOUTH REVER ROAD WEST LAFAYETTE, IN 47906 P: 765-464-2227 F: 765-464-6986</p>	<p>PROJECT:</p> <p align="center">TRUSS INSPECTION ROUND ROBIN</p>		<p>SHEET NOTES:</p> <ol style="list-style-type: none"> 1. INDICATE THE LOCATION OF THE CRITICAL SECTION ON THE ELEVATION VIEW OF THE TRUSS CHORD (THIS SHEET). 2. PROVIDE MEASUREMENTS TO SUPPORT SECTION LOSS CALCULATIONS AT THIS LOCATION. SHEETS NO. THROUGH NO. MAY BE USED FOR RECORDING MEASUREMENTS, ADDITIONAL SKETCHES, ETC. 3. THE TIME LIMIT FOR THIS TASK IS 30 MINS.
	<p>INSPECTOR ID: _____</p> <p>INSPECTION DATE: _____</p> <p>PROCTOR INITIALS: _____</p>		<p>SHEET TITLE:</p> <p align="center">TASK 1 SKETCHES</p> <p>SHEET NO. 1-1</p>

Inspection Forms

<p style="text-align: center;">ELEVATION SCALE: 1" = 4'</p>			
<p style="text-align: center;">CROSS SECTION 1 SCALE: 1 1/2" = 1'-0"</p>	<p style="text-align: center;">CROSS SECTION 2 SCALE: 1 1/2" = 1'-0"</p>	<p>PROJECT: TRUSS INSPECTION ROUND ROBIN</p> <p>SHEET NOTES:</p> <ol style="list-style-type: none"> 1. PROVIDE THICKNESS MEASUREMENTS AT CROSS SECTION 1 AND 2. RECORD YOUR FINDINGS ON SHEETS 2-1 AND 2-1 (IF NEEDED). 2. INDICATE THE LOCATION OF THE MEASUREMENTS ON THE CROSS SECTION SKETCHES (THIS SHEET). <p>SHEET TITLE: TASK 2 SKETCHES</p> <p>SHEET NO.: 2-1</p>	
<p>BOWEN LABORATORY PURDUE UNIVERSITY 1040 SOUTH RIVER ROAD WEST LAFAYETTE, IN 47906 P: 765-494-2227 F: 765-494-6886</p>			

<p style="text-align: center;">ELEVATION SCALE: 1" = 4'</p>			
<p style="text-align: center;">PLAN VIEW OF REGION 3 SCALE: 1/4" = 1'-0"</p>			
<p>PROJECT: TRUSS INSPECTION ROUND ROBIN</p> <p>SHEET NOTES:</p> <ol style="list-style-type: none"> 1. INDICATE THE LOCATION OF THE CRITICAL SECTION WITHIN REGION 3 ON THE PLAN VIEW SKETCH. PROVIDE THE DISTANCE BETWEEN THE FACE OF BATTEN PLATE C AND THE CRITICAL SECTION (X = ____). 2. THICKNESS MEASUREMENTS DO NOT NEED TO BE PROVIDED AT THIS LOCATION. <p>SHEET TITLE: TASK 3 SKETCHES</p> <p>SHEET NO.: 3-1</p>			
<p>BOWEN LABORATORY PURDUE UNIVERSITY 1040 SOUTH RIVER ROAD WEST LAFAYETTE, IN 47906 P: 765-494-2227 F: 765-494-6886</p>			

Exit Form

Truss Inspection Round Robin
Inspector Information

Please complete the following form. All information collected is used for research purposes only.

Inspector ID: _____ Inspection Date: _____
 Employer: _____ Years of Inspection Experience: _____
 Age: _____ Gender: _____

- 1) Please list all tools you brought to complete this inspection.

- 2) Please list which tools you used during the inspection course.

- 3) Are there any additional tools you wish you would have had available?

- 3) Approximately how many hands-on, fracture critical inspections have you completed in the last 12 months?

- 4) Approximately how many routine inspections have you completed in the last 12 months?
 What percentage of these were performed on steel bridges?

- 5) Approximately what percentage of your work time (on a per week basis) is spent performing bridge inspections?

- 6) Please indicate which inspection training courses you have completed and the date of completion (estimates are acceptable for the dates):
 - Safety Inspection of In-Service Bridges (FHWA/NHI): _____
 - Bridge Inspection Refresher Training (FHWA/NHI): _____
 - Engineering Concepts for Bridge Inspectors (FHWA/NHI): _____
 - Underwater Bridge Inspection (FHWA/NHI): _____
 - Fracture Critical Inspection Techniques for Steel Bridges (FHWA/NHI): _____
 - Inspection and Maintenance of Ancillary Highway Structures (FHWA/NHI): _____
 - Introduction to Element Level Bridge Inspection (FHWA): _____
 - Inspecting Steel Bridges for Fatigue (Purdue/S-BRITE): _____
 - State Specific Bridge Inspection Refresher Training: _____
 - Others: _____

Exit Form

Truss Inspection Round Robin
Inspector Information

7) Are you qualified as a Team Leader?

Yes

No

8) Have you received any training specific to estimating section loss in steel members? If yes, please provide a brief description.

Yes

No

9) Are you a registered professional engineer or engineer-in-training?

Yes

No

If yes, please indicate your license below.

SE

PE (civil)

PE (other)

FE (civil)

FE (other)

Other

10) What is the highest educational level that you have completed?

_____ High School degree or equivalent

_____ Trade School Degree

_____ Associate's Degree

_____ Bachelor's Degree in CIVIL ENGINEERING? Or OTHER? (circle one)

_____ Master's Degree in CIVIL ENGINEERING? Or OTHER? (circle one)

_____ PhD in CIVIL ENGINEERING? Or OTHER? (circle one)

11) How did your effort level during this task compare to your effort level during a typical bridge inspection?

More effort

Similar effort

Less effort

12) How did your focus level during this task compare to your focus level during a typical bridge inspection?

More focused

Similar focus

Less focused

APPENDIX D. TRUSS CHORD LOAD RATING DOCUMENTS

Load Rating Procedure

Truss Chord Load Rating Round Robin

Evaluation Instructions
Purdue University
2018

This document provides the procedures and instructions for engineers participating in the Truss Chord Load Rating Round Robin at Purdue University.

Introduction:

You are about to participate in a study on corrosion evaluation of steel bridges. It is critical to the study that all procedures and findings are kept **CONFIDENTIAL**. Please do not share the inspection information or discuss your evaluation with other engineers or personnel in your office. Likewise, your identity will remain confidential; the results from your specific participation will not be shared or be published.

The objectives of this portion of study are as follows:

1. Establish the variability in load rating evaluations of steel bridge tension members with corrosion and section loss.
2. Investigate the various approaches load rating engineers use to interpret and apply inspection findings.
3. Identify methods for improving consistency in how the effects of corrosion are accounted for in load rating.

Load Rating Scenario:

The evaluation is focused on the load rating of the lower chord in Span 16 between joints L2 and L4 (Ref. Construction Plans, Sheet 33). The joints and gusset plates do not need to be load rated. The load rating should be completed using the load and resistance factor rating method in accordance with the 2nd Edition of the *AASHTO Manual for Bridge Evaluation*, including the 2016 interim revisions. The evaluation is divided into two separate tasks and may be completed in any order.

The unfactored dead and live loads are provided. The live load corresponds to the HL-93 loading. You will need to make an assumption regarding the material strength (yield and ultimate) and identify the governing fatigue category. The inspection report and relevant construction plans have been provided for your reference.

Bridge Data:

- The bridge was constructed in 1941 in the upper Midwest region of the United States.
- The member under evaluation is the bottom truss chord in Span 16 between joints L2 and L4.
- The member is fracture critical.
- $P_{DC} = 335$ kips, $P_{DW} = 0$ kips, $P_{LL+IM} = 322$ kips (design truck with lane load),
 $P_{LL+IM} = 134$ kips (fatigue truck)
- These loads are unfactored. They include the distribution factor and dynamic load allowance.
- $(ADTT)_{SL} = 1500$ and it is assumed that $(ADTT)_{SL}$ is constant through the life of the bridge.

Load Rating Procedure
Truss Chord Load Rating Round Robin
Evaluation Instructions
Purdue University
2018

Evaluation Tasks:

Task 1:

Calculate the inventory level load rating factor for the Strength I limit state in the as-built (undamaged) condition AND the as-inspected (damaged) condition.

Task 2:

Calculate the remaining fatigue life in the as-inspected (damaged) condition.

A blank load rating report has been provided for recording your results. Please show all work to support your calculation of load rating, remaining fatigue life, member capacity, and effective stress range, including appropriate load and resistance factors. Reference applicable equations and provisions from the *AASHTO Manual for Bridge Evaluation*. After finishing the evaluation, please complete the Load Rater Exit Form. The load rating report, calculations, and exit form may be emailed to the Research Team (campb270@purdue.edu).

Thank you for your participation and cooperation. Please contact us with any questions.

Dr. Robert J. Connor
rconnor@purdue.edu

Leslie Campbell, PE
campb270@purdue.edu
(916) 803-2428

Blank Load Rating Report

BRIDGE RATING REPORT

(Note: Load rater to complete all YELLOW fields)

Load Rater:
 Date:
 Rating Method: Load and resistance factor

IDENTIFICATION

Structure Type: Steel deck truss (approach span)
 Description: Two (2) parallel trusses/simple span
 Year Built: 1941

GEOMETRY

Span Number:	16 (approach span)	O-to-O Coving:	35'-2"
Span Length:	128'-0"	Clear Roadway:	27'-0"
Truss Spacing:	32'-2"	Skew:	0°

STEEL SUPERSTRUCTURE

Structural Steel:
 Steel Fy: ksi
 Steel Fu: ksi

DESIGN LOADS (unfactored)

Dead Load:
 P_{DC} 335 kips
 P_{DW} 0 kips
 Live Load: Strength I
 P_{LL+IM} 322 kips
 Live Load: Fatigue
 P_{LL+IM} 134 kips

STRENGTH I RATING (as-built condition)

Vehicle Configuration: HL-93
 Member Capacity: kips
 Inventory Rating Factor:

STRENGTH I RATING (as-inspected condition)

Vehicle Configuration: HL-93
 Member Capacity: kips
 Inventory Rating Factor:

FATIGUE ANALYSIS (as-inspected condition)

Vehicle Configuration: HL-93 (Fatigue Truck)
 Governing Fatigue Category:
 Effective Stress Range: ksi
 Remaining Fatigue Life: years *If infinite fatigue life remains, indicate such here.

Mock Inspection Report

Inspector: Bridge Inspection Report			
Inspection Date: 5/2018			
STRUCTURE TYPE AND MATERIAL			
(43) STRUCTURE TYPE, MAIN:		(45) NUMBER OF SPANS IN MAIN UNIT:	3
(A) KIND OF MATERIAL/DESIGN: 4 - Steel continuous		(46) NUMBER OF APPROACH SPANS:	21
(B) TYPE OF DESIGN/CONSTR: 10 - Truss (thru)		(107) DECK STRUCTURE TYPE:	1 - Concrete Cast-in-Place
(44) STRUCTURE TYPE, APPROACH SPANS:		(108) WEARING SURFACE/PROT SYS:	
(A) KIND OF MATERIAL/DESIGN: 3 - Steel		(A) WEARING SURFACE:	1 - Monolithic Concrete
(B) TYPE OF DESIGN/CONSTR: 9 - Truss (deck)		(B) DECK MEMBRANE:	0 - None
		(C) DECK PROTECTION:	1 - Epoxy Coated Reinforcing
AGE OF SERVICE			
(27) YEAR BUILT: 1941		(28) LANES:	
(106) YEAR RECONSTRUCTED: 1985		(A) ON BRIDGE:	2
(42) TYPE OF SERVICE:		(B) UNDER BRIDGE:	6
(A) ON BRIDGE:		(29) AVERAGE DAILY TRAFFIC:	17500
(B) UNDER BRIDGE:		(30) YEAR OF AVERAGE DAILY TRAFFIC:	2004
		(109) AVERAGE DAILY TRUCK TRAFFIC:	10%
		(19) BYPASS DETOUR LENGTH:	65 MI.
GEOMETRIC DATA			
(48) LENGTH OF MAX SPAN: 450 FT		(35) STRUCTURE FLARED:	0 - No flare
(49) STRUCTURE LENGTH: 2281 FT		(10) INV RTE, MIN VERT CLEARANCE:	18 FT
(50) CURB/SIDEWALK WIDTHS:		(47) TOT HORIZ CLEARANCE:	9.4 FT
(A) LEFT: 4.9 FT		(53) VERT CLEAR OVER BR RDWY:	17.9 FT
(B) RIGHT: 4.6 FT		(54) MIN VERTICAL UNDERCLEARANCE:	
(51) BRDG RDWY WIDTH CURB-TO-CURB: 27 FT		(A) REFERENCE FEATURE:	H
(52) DECK WIDTH, OUT-TO-OUT: 35.2 FT		(B) MIN VERT UNDERCLEAR:	13.9 FT
(32) APPROACH ROADWAY: 9.4 FT.		(55) LATERAL UNDERCLEARANCE RIGHT:	
(33) BRIDGE MEDIAN: 0 - No median		(A) REFERENCE FEATURE:	H
(34) SKEW: 0 DEG		(B) MIN LATERAL UNDERCLEAR:	8.9 FT
		(56) MIN LATERAL UNDERCLEAR ON LEFT:	0 FT

1 of 14



Inspector: Bridge Inspection Report			
Inspection Date: 5/2018			
INSPECTIONS			
(90) INSPECTION DATE:	6/16	(91) DESIGNATED INSPECTION FREQUENCY:	12
(92) CRITICAL FEATURE INSPECTION:		(93) CRITICAL FEATURE INSPECTION DATE:	
A) FRACTURE CRITICAL REQUIRED/FREQUENCY:	Y24	A) FRACTURE CRITICAL DATE:	6/16
B) UNDERWATER INSPECTION REQUIRED/FREQUENCY:	Y60	B) UNDERWATER INSP DATE:	8/12
C) OTHER SPECIAL INSPECTION REQUIRED/FREQUENCY:	N	C) OTHER SPECIAL INSP DATE:	
CONDITION			
(58) DECK:	5 - Fair Condition	(60) SUBSTRUCTURE:	5 - Fair Condition
(59) SUPERSTRUCTURE:	4 - Poor Condition	(61) CHANNEL/CHANNEL PROTECTION:	4 - Bank and embankment protection is severely undermined
		(62) CULVERTS:	N - Not Applicable
CONDITION COMMENTS			
(58) DECK: 5 - Fair Condition			
Comments:			
There are transverse cracks with efflorescence with a 10' to 20' spacing on the underside of the deck.			
There are large longitudinal spalls up to 2" deep in spans 17 and 23. The spall in span 17 is about 8' x 5' in area with 1 exposed rebar and the spall in span 23 is 12' x 10' in area with 3 exposed rebars.			
(58.01) WEARING SURFACE: 4 - Poor Condition			
Comments:			
The wearing surface is rough. There are areas of map cracking, spalling, and patching. Spalling deeper than 2" in spans 15 (1.5'x1') and 16 (1'x1').			
(59) SUPERSTRUCTURE: 4 - Poor Condition			
Comments:			
Crevice corrosion is common at gusset plates and in seams of built-up members in deck trusses, with section loss and distortion in primary members. See attached sheets for thickness measurements. Section loss has led to crack formation in a few locations. Areas of pitting and minor section loss along the flanges of the truss members.			
(60) SUBSTRUCTURE: 5 - Fair Condition			
Comments:			
There are some cracks and spalls on the piers.			
Deteriorated concrete with exposed rebars on Piers 14 and 15. Pier 22 has cracking and spalling with exposed rebar.			

2 of 14

Mock Inspection Report

Inspector:		Bridge Inspection Report	
Inspection Date: 5/2018			
(61) CHANNEL/CHANNEL PROTECTION:	4 - Bank and embankment protection is severely undermined		
Comments: Bank and embankment protection is severely undermined. River control devices have severe damage.			
(62) CULVERTS:	N - Not Applicable		
Comments:			
CLASSIFICATION			
(20) TOLL:	3 - On Free Road	(21) MAINT. RESPONSIBILITY:	1 - State Highway Agency
(22) OWNER:	1 - State Highway Agency	(26) FUNCTION CLASS OF INVENTORY RTE:	14 - Other Principal Arterial
(37) HISTORICAL SIGNIFICANCE:	2 - Eligible for the National Register of Historic Places	(100) STRAHNET HIGHWAY:	0 - Not a STRAHNET route
(101) PARALLEL STRUCTURE:	N - No parallel structure	(102) DIRECTION OF TRAFFIC:	2 - 2-way traffic
(103) TEMPORARY STRUCTURE:		(104) HIGHWAY SYSTEM OF INVENTORY BRIDGE:	1 - Structure/route is on the NHS
(105) FEDERAL LANDS HIGHWAYS:	0 - Not applicable	(110) DESIGNATED NATIONAL NETWORK:	0 - Inventory route not on network
(112) NBIS BRIDGE LENGTH:	Y		
NAVIGATION DATA			
(38) NAVIGATION CONTROL:	1 - Navigation control on waterway	(39) NAVIAGATION VERTICAL CLEAR:	63 FT
(111) PIER OF ABUTMENT PROTECTION:	1 - Navigation protection not required	(116) MINIMUM NAVIGATION VERT. CLEARANCE, VERT. LIFT BRIDGE:	N/A
		(40) NAV HORIZONTAL CLEARANCE:	438 FT

3 of 14

Inspector:		Bridge Inspection Report	
Inspection Date: 5/2018			
			
Photo 1: Elevation View of Approach Span 16			
			
Photo 2: Bottom Truss Chord (Member L2-L4)			

4 of 14

Mock Inspection Report

Inspector:
Inspection Date: 5/2018

Bridge Inspection Report



Photo 3: General condition of web plate and channels (looking from L2 to L4)



Photo 4: General condition of exterior face of channels (Channel A)

5 of 14

Inspector:
Inspection Date: 5/2018

Bridge Inspection Report



Photo 5: General condition of batten plates



Photo 6: Crack in web plate near top (similar in a few locations)

6 of 14

Mock Inspection Report

Inspector:
Inspection Date: 5/2018

Bridge Inspection Report



Photo 7: Cross Section 1 (129" from L2)



Photo 8: Cross Section 1 (129" from L2)

7 of 14

Inspector:
Inspection Date: 5/2018

Bridge Inspection Report



Photo 9: Channel A at Cross Section 1 (top)



Photo 10: Channel B at Cross Section 1 (top)

8 of 14

Mock Inspection Report

Inspector:
Inspection Date: 5/2018

Bridge Inspection Report



Photo 11: Small crack in web plate of Channel B at Cross Section 1



Photo 12: Channel A at Cross Section 1 (bottom)

9 of 14

Inspector:
Inspection Date: 5/2018

Bridge Inspection Report



Photo 13: Cross Section 2 (113" from L4)



Photo 14: Channel A at Cross Section 2 (top)

10 of 14

Mock Inspection Report

Inspector:
Inspection Date: 5/2018

Bridge Inspection Report



Photo 15: Channel B at Cross Section 2 (top)



Photo 16: Channel A at Cross Section 2 (bottom)

11 of 14

Inspector:
Inspection Date: 5/2018

Bridge Inspection Report



Photo 17: Channel B at Cross Section 2 (bottom)



Photo 18: Small holes in web plate of Channel B near Cross Section 2

12 of 14

Mock Inspection Report

Inspector:
Inspection Date: 5/2018

Bridge Inspection Report



Photo 19: Cross Section 3 (18" from L2)



Photo 20: Cross Section 3 (18" from L2)

13 of 14

Inspector:
Inspection Date: 5/2018

Bridge Inspection Report



Photo 21: Bottom batten plate at Cross Section 3



Photo 22: Web and splice plates on Channel B at Cross Section 3

14 of 14

Mock Inspection Report

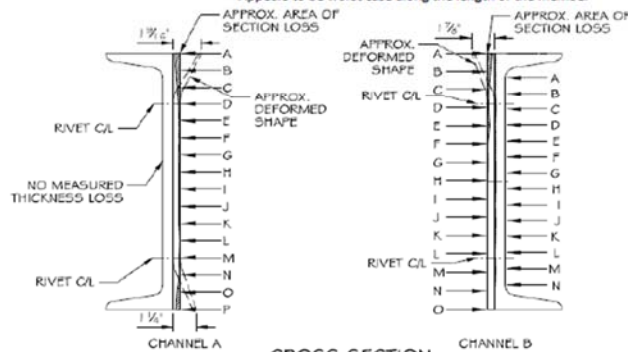
Inspector:

Inspection Date: 5/2018

Bridge Inspection Report - METALWORK LOSSES

CROSS SECTION NO.	SPAN/ MEMBER	SIDE	DESCRIPTION
1	16/L2 - L4	A	157" from C/L of L2 - crevice corrosion between channel and web pl. 1-9/16" of distortion at top, 1-1/4" of distortion at bottom Appears to be worst case along the length of the member

1	16/L2 - L4	B	157" from C/L of L2 - crevice corrosion between channel and web pl. 1-7/8" of distortion at top, 0" of distortion at bottom Small crack forming due to section loss at top of web plate Appears to be worst case along the length of the member
---	------------	---	--



CROSS SECTION
(LOOKING FROM L4 TO L2)

- NOTES:
1. MEASUREMENTS EQUALLY SPACED.
2. MEASUREMENTS TAKEN 1.29" FROM EDGE OF GUSSET PLATE AT JOINT L2.
3. MEASURED THICKNESS OF WEB PLATE IN AN UNDETERIORATED REGION IS 0.39\"

CHANNEL A, WEB PLATE	
LOCATION	MEASUREMENT
A	0.11"
B	0.15"
C	0.25"
D	0.3"
E	0.33"
F	0.34"
G	0.31"
H	0.32"
I	0.32"
J	0.32"
K	0.32"
L	0.37"
M	0.33"
N	0.35"
O	0.24"
P	0.19"

CHANNEL B, WEB PLATE	
LOCATION	MEASUREMENT
A	0.16"
B	0.24"
C	0.31"
D	0.33"
E	0.27"
F	0.37"
G	0.35"
H	0.37"
I	
J	
K	
L	
M	
N	
O	

CHANNEL B, CHANNEL	
LOCATION	MEASUREMENT
A	0.5"
B	0.5"
C	0.46"
D	0.5"
E	0.49"
F	0.49"
G	0.47"
H	0.48"
I	0.5"
J	0.48"
K	0.49"
L	0.52"
M	0.49"
N	0.52"

Page 1 of 3

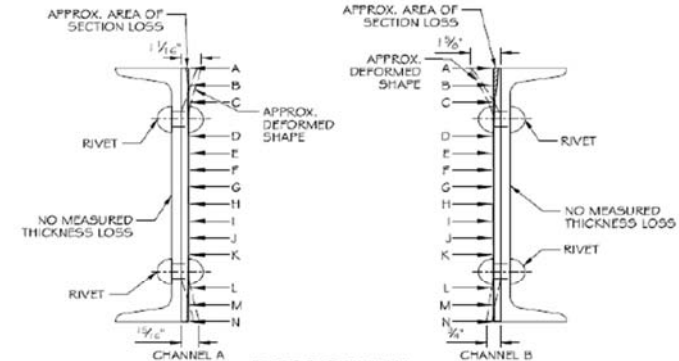
Inspector:

Inspection Date: 5/2018

Bridge Inspection Report - METALWORK LOSSES

CROSS SECTION NO.	SPAN/ MEMBER	SIDE	DESCRIPTION
2	16/L2 - L4	A	140" from C/L of L4 - crevice corrosion between channel and web pl. 1-1/16" of distortion at top, 15/16" of distortion at bottom Section occurs at rivet holes

2	16/L2 - L4	B	140" from C/L of L4 - crevice corrosion between channel and web pl. 1-5/8" of distortion at top, 3/4" of distortion at bottom Section occurs at rivet holes, small holes in web plate near top edge
---	------------	---	---



CROSS SECTION
(LOOKING FROM L4 TO L2)

- NOTES:
1. MEASUREMENTS EQUALLY SPACED.
2. MEASUREMENTS TAKEN 1.13" FROM EDGE OF GUSSET PLATE AT JOINT L4.
3. MEASURED THICKNESS OF WEB PLATE IN AN UNDETERIORATED REGION IS 0.39\"

CHANNEL A, WEB PLATE	
LOCATION	MEASUREMENT
A	0.23"
B	0.24"
C	0.34"
D	0.39"
E	
F	
G	
H	
I	
J	
K	
L	0.36"
M	0.32"
N	0.26"

CHANNEL B, WEB PLATE	
LOCATION	MEASUREMENT
A	0.11"
B	0.17"
C	0.29"
D	0.39"
E	0.39"
F	0.38"
G	0.39"
H	0.39"
I	0.38"
J	0.38"
K	0.39"
L	0.36"
M	0.34"
N	0.31"

Page 2 of 3

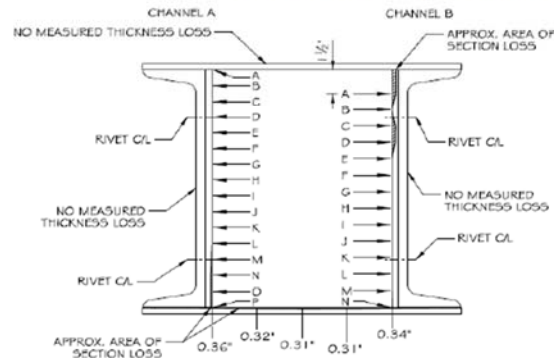
Mock Inspection Report

Inspector:

Inspection Date: 5/2018

Bridge Inspection Report - METALWORK LOSSES

CROSS SECTION NO.	SPAN/ MEMBER	SIDE	DESCRIPTION
3	16/L2 - L4	A	45.5" from C/L of L2, at outside edge of L2 joint (before splice plate) Web plate in good condition, minor thickness loss along bottom batten plate
3	16/L2 - L4	B	45.5" from C/L of L2, at outside edge of L2 joint (before splice plate) Near complete loss of top 1.5" of web plate Distortion/thickness loss in top batten plate at edge of Channel B



CROSS SECTION
(LOOKING FROM L4 TO L2)

NOTES:

1. MEASUREMENTS EQUALLY SPACED.
2. MEASUREMENTS TAKEN 17.5" FROM EDGE OF GUSSET PLATE AT JOINT L2.
3. MEASURED THICKNESS OF WEB PLATE IN AN UNDETERIORATED REGION IS 0.39".

LOCATION	MEASUREMENT
A	0.37"
B	
C	
D	
E	
F	
G	
H	
I	
J	
K	
L	
M	0.34"
N	0.31"
O	
P	

LOCATION	MEASUREMENT
A	0.16"
B	0.37"
C	0.16"
D	0.19"
E	0.37"
F	
G	
H	
I	
J	
K	
L	
M	
N	0.34"

Exit Form

Load Rating Round Robin Load Rater Information



Please complete the following form. All information collected is used for research purposes only.

Name: _____ Date of Load Rating: _____
 Employer: _____ Years of Load Rating Experience: _____
 Age: _____

1) Please list the references, if any, you consulted while performing this load rating.

2) Please list the computer programs, if any, you used to perform this load rating.

3) Have you previously load rated a steel truss bridge?

Yes

No

4) Approximately how long did it take you to complete this load rating? Round to the nearest quarter hour.

5) Is there any additional information regarding the bridge structure that you wish had been provided?

6) Approximately how many bridge load ratings have you completed in the last 12 months? What percentage of these were performed on steel bridges?

7) Approximately what percentage of your work time (on a per week basis) is spent performing bridge load ratings?

Exit Form

Load Rating Round Robin
Load Rater Information



8) Please indicate which inspection training courses you have completed and the date of completion (estimates are acceptable for the dates):

Fundamentals of LRFR and Applications of LRFR for Bridge Superstructures (FHWA/NHI): _____

Load and Resistance Factor Rating for Highway Bridges (FHWA/NHI): _____

Load Rating of Steel Truss Bridges (FHWA Webinar): _____

Implementation of the LRFR Method (FHWA Webinar): _____

State Specific Bridge Load Rating Training: _____

Other:

9) Have you received any training specific to evaluating corrosion or section loss in steel members?

Yes

No

If yes, please provide a brief description.

10) Are you a registered professional engineer or engineer-in-training?

Yes

No

If yes, please indicate your license below.

SE

PE (civil)

PE (other)

FE (civil)

FE (other)

Other

11) What is the highest level of education that you have completed?

_____ High School degree or equivalent

_____ Trade School Degree

_____ Associate's Degree

_____ Bachelor's Degree in CIVIL ENGINEERING? Or OTHER? (circle one)

_____ Master's Degree in CIVIL ENGINEERING? Or OTHER? (circle one)

_____ PhD in CIVIL ENGINEERING? Or OTHER? (circle one)

Exit Form

Load Rating Round Robin
Load Rater Information



12) Do you have experience performing routine or hands-on inspections of steel bridges?

Yes

No

If so, how many years of inspection experience do you have?

13) Do you have experience designing steel bridges for new construction?

Yes

No

If yes, have you designed a steel truss bridge?

Load Rating Reference Calculations

BRIDGE RATING REPORT

(Note: Load rater to complete all YELLOW fields)

Load Rater: LEC
 Date: June 2018
 Rating Method: Load and resistance factor

IDENTIFICATION

Structure Type: Steel deck truss (approach span)
 Description: Two (2) parallel trusses/simple span
 Year Built: 1941

GEOMETRY

Span Number: 16 (approach span) O-to-O Coping: 35'-2"
 Span Length: 128'-0" Clear Roadway: 27'-0"
 Truss Spacing: 32'-2" Skew: 0°

STEEL SUPERSTRUCTURE

Structural Steel:
 Steel Fy: 33 ksi (Actual material properties not known, assumed values based on MBE
 Steel Fu: 66 ksi Table 6A.6.2.1-1 for construction in 1941)

DESIGN LOADS (unfactored)

Dead Load:
 P_{DC} 335 kips
 P_{DW} 0 kips
 Live Load: Strength I
 P_{LL+IM} 322 kips
 Live Load: Fatigue
 P_{LL+IM} 134 kips

STRENGTH I RATING (as-built condition)

Vehicle Configuration: HL-93
 Member Capacity: 978 kips
 Inventory Rating Factor: 0.99

STRENGTH I RATING (as-inspected condition)

Vehicle Configuration: HL-93
 Member Capacity: 866 kips
 Inventory Rating Factor: 0.79

FATIGUE ANALYSIS (as-inspected condition)

Vehicle Configuration: HL-93 (Fatigue Truck)
 Governing Fatigue Category: D
 Effective Stress Range: 3.6 ksi
 Remaining Fatigue Life: 32 years *If infinite fatigue life remains, indicate such here.

Load Rating Reference Calculations

Bridge Rating Details -
Calculations

By: LEC
Date: June 2018

References:

1. AASHTO LRFD Bridge Design Specifications, 7th Edition
2. AASHTO Manual for Bridge Evaluation, 2nd Edition (with 2016 revisions)

Step 1: Determine the net area and the gross area of the member in the as-built and as-inspected conditions (see page 4)

$$A_{g_asbuilt} := 34.65 \text{ in}^2$$

$$A_{n_asbuilt} := 28.95 \text{ in}^2$$

$$A_{g_asinspected} := 32.5 \text{ in}^2 \quad (\text{at cross section 1})$$

$$A_{n_asinspected} := 27.91 \text{ in}^2 \quad (\text{at cross section 3})$$

Step 2: Determine member capacity considering yielding on the gross section (AASHTO LRFD BDS 6.8.2.1)

$$P_{n_yield} = \phi_y \cdot A_g \cdot F_y$$

$$\phi_y := 0.95 \quad (\text{AASHTO LRFD BDS Section 6.5.4.2})$$

$$F_y := 33 \text{ ksi} \quad (\text{Assumed yield strength based on AASHTO MBE Table 6A.6.2.1-1})$$

As-built condition:

$$P_{n_yield_asbuilt} := \phi_y \cdot A_{g_asbuilt} \cdot F_y = 1086.28 \text{ kip}$$

As-inspected condition:

$$P_{n_yield_asinspected} := \phi_y \cdot A_{g_asinspected} \cdot F_y = 1018.88 \text{ kip}$$

Step 3: Determine member capacity considering fracture on the net section (AASHTO LRFD BDS 6.8.2.1)

$$P_{n_fracture} = \phi_u \cdot A_n \cdot U \cdot R_p \cdot F_u$$

$$\phi_u := 0.8 \quad (\text{AASHTO LRFD BDS Section 6.5.4.2})$$

$$U := 1 \quad (\text{not end connected})$$

$$R_p := 1 \quad (\text{holes were reamed per note on Sheet 33 of the construction plans})$$

$$F_u := 66 \text{ ksi} \quad (\text{assumed ultimate strength based on AASHTO MBE Table 6A.6.2.1-1})$$

As-built condition:

$$P_{n_fracture_asbuilt} := \phi_u \cdot A_{n_asbuilt} \cdot U \cdot R_p \cdot F_u = 1528.56 \text{ kip}$$

As-inspected condition:

$$P_{n_fracture_asinspected} := \phi_u \cdot A_{n_asinspected} \cdot U \cdot R_p \cdot F_u = 1473.65 \text{ kip}$$

Load Rating Reference Calculations

Bridge Rating Details -
Calculations

By: LEC
Date: June 2018

Step 4: Determine inventory load rating factor for Strength I limit state (AASHTO MBE 6A.4.2)

$$RF = \frac{Capacity - (\gamma_{DC})(DC) - (\gamma_{DW})(DW)}{(\gamma_{LL})(LL + IM)}$$

Dead Load:	$DC := 335 \text{ kip}$	$\gamma_{DC} := 1.25$	(AASHTO MBE Table 6A.4.2.2-1)
Wearing Surface:	$DW := 0 \text{ kip}$	$\gamma_{DW} := 1.5$	
Live Load (incl. impact):	$LL := 322 \text{ kip}$	$\gamma_{LL} := 1.75$	

$$Capacity = \phi_c \cdot \phi_s \cdot \phi \cdot R_n \quad (\text{AASHTO MBE Eq. 6A.4.2.1-2})$$

$\phi_c := 0.85$ (Poor condition, AASHTO MBE Section 6A.4.2.3)

$\phi_s := 0.9$ (Riveted members in two girder truss bridge, AASHTO MBE Section 6A.4.2.4)

Note: $\phi_c \cdot \phi_s > 0.85$ per Section AASHTO MBE Section 6A.4.2.1-3, therefore $\phi_c \phi_s := 0.85$

$$C_{asbuilt} := \phi_s \cdot \min(P_{n_yield_asbuilt}, P_{n_fracture_asbuilt}) = 977.65 \text{ kip} \quad (\text{Note, condition factor not included in as-built rating})$$

$$C_{asinspected} := \phi_c \phi_s \cdot \min(P_{n_yield_asinspected}, P_{n_fracture_asinspected}) = 866.04 \text{ kip}$$

Note: yielding on the gross section governs in both cases

As-built condition:

$$RF_{asbuilt} := \frac{C_{asbuilt} - (\gamma_{DC})(DC) - (\gamma_{DW})(DW)}{(\gamma_{LL})(LL)} = 0.99$$

As-inspected condition:

$$RF_{asinspected} := \frac{C_{asinspected} - (\gamma_{DC})(DC) - (\gamma_{DW})(DW)}{(\gamma_{LL})(LL)} = 0.79$$

Step 5: Estimate effective stress range for determining fatigue life (AASHTO MBE 7.2.2, 2015 revision)

$$\Delta f_{eff} = R_p \cdot R_s \cdot \Delta f$$

Live load from fatigue truck (incl. impact):	$LL_{fatigue} := 134 \text{ kip}$	$\gamma_{fatigueI} := 1.5$ $\gamma_{fatigueII} := 0.75$	(AASHTO LRFD BDS Table 3.4.1-1)
--	-----------------------------------	--	---------------------------------

Span length:	$L := 128 \text{ ft}$
Number of lanes:	$n_L := 2$
Single lane average daily truck traffic:	$ADTT_{SL} := 1500$
Structure age:	$age := 2018 - 1941 = 77$

Load Rating Reference Calculations

Bridge Rating Details -
Calculations

By: LEC
Date: June 2018

$$R_p := 0.988 + 6.87 \cdot 10^{-5} \cdot \left(\frac{L}{ft} \right) + 4.01 \cdot 10^{-6} \cdot (ADTT_{SL}) + \frac{0.0107}{n_L}$$

$$R_p = 1.01 \quad (\text{Multiple presence factor, AASHTO MBE Eq. 7.2.2.1-1})$$

$$R_s = 1.0 \quad (\text{Stress-range estimate partial load factor, AASHTO MBE Section 7.2.2.1.1})$$

$$\Delta f := \frac{LL_{fatigue}}{A_{n_{\text{inspected}}}} \cdot \gamma_{fatigueII} = 3.6 \frac{\text{kip}}{\text{in}^2} \quad (\text{factored calculated stress range for Fatigue II load combination})$$

$$\Delta f_{eff} := R_p \cdot R_s \cdot \Delta f = 3.63 \frac{\text{kip}}{\text{in}^2}$$

Step 6: Determine remaining fatigue life

$$\Delta F_{TH} := 7 \text{ ksi} \quad (\text{Constant amplitude fatigue threshold for Fatigue Category D, AASHTO LRFD BDS Table 6.6.1.2.5-3 - reasonable due to deteriorated condition})$$

- Infinite Life Check (AASHTO MBE 7.2.4, 2015 revisions)

$$\Delta f_{max} := \frac{\gamma_{fatigueI}}{\gamma_{fatigueII}} \cdot \Delta f \cdot R_p = 7.26 \frac{\text{kip}}{\text{in}^2}$$

Since $\Delta f_{max} > \Delta F_{TH}$, finite fatigue life remains

- Finite Life Check (AASHTO MBE 7.2.5, 2015 revisions)

$$Years_{remaining} = \frac{N_{available} - N_{used}}{365 \cdot n \cdot ADTT_{SL}}$$

$$A := 22 \cdot 10^8 \text{ ksi}^3 \quad (\text{Detail category constant, AASHTO LRFD BDS Table 6.6.1.2.5-1})$$

$$n := 1 \quad (\text{Cycles per truck passage, AASHTO LRFD BDS Table 6.6.1.2.5-2})$$

$$R_R := 1.3 \quad (\text{Resistance factor for evaluation, AASHTO MBE Table 7.2.5.1-1 for Category D Evaluation 1 Life})$$

$$N_{available} := \frac{R_R \cdot A}{(\Delta f_{eff})^3} = 59780670 \quad (\text{AASHTO MBE Eq. 7.2.5.1-1})$$

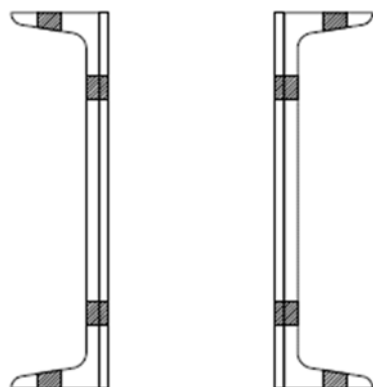
$$N_{used} := 365 \cdot n \cdot ADTT_{SL} \cdot age = 42157500$$

$$N_{remaining} := N_{available} - N_{used} = 17623170.23$$

$$Y_{remaining} := \frac{N_{remaining}}{365 \cdot n \cdot ADTT_{SL}} = 32.19 \quad (\text{AASHTO MBE Eq. 7.2.5.1-1})$$

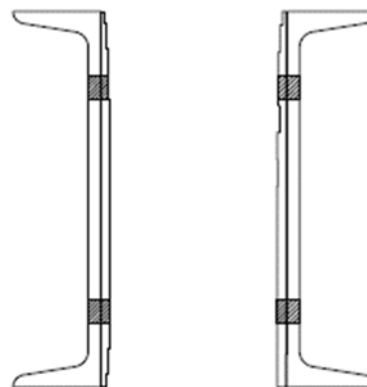
Load Rating Reference Calculations

ELEMENT	GROSS AREA (SQ. IN.)	NET AREA (SQ. IN.)
CHANNEL A, WEB PLATE	5.63	4.92
CHANNEL A, CHANNEL	11.7	9.55
CHANNEL B, WEB PLATE	5.63	4.92
CHANNEL B, CHANNEL	11.7	9.55
TOTAL	34.65	28.95



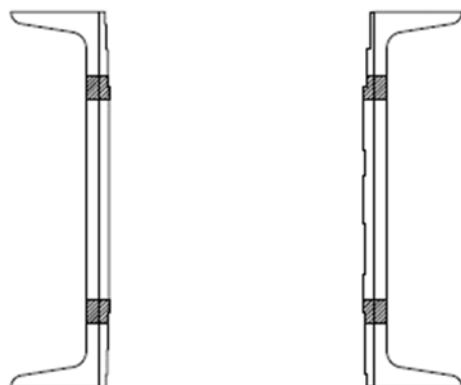
AS-BUILT

ELEMENT	GROSS AREA (SQ. IN.)	NET AREA (SQ. IN.)
CHANNEL A, WEB PLATE	4.43	3.84
CHANNEL A, CHANNEL	11.7	10.73
CHANNEL B, WEB PLATE	5.06	4.41
CHANNEL B, CHANNEL	11.31	10.39
TOTAL	32.5	29.36
% LOSS	6.2%	6.1%



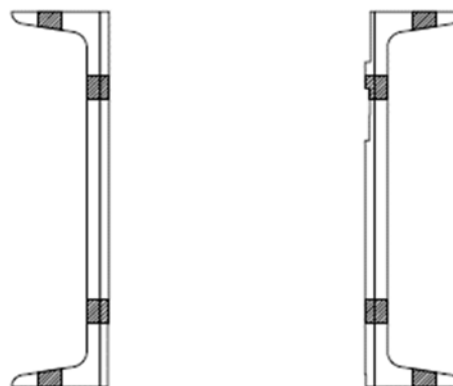
AS-INSPECTED (CROSS SECTION 1)

ELEMENT	GROSS AREA (SQ. IN.)	NET AREA (SQ. IN.)
CHANNEL A, WEB PLATE	5.41	4.71
CHANNEL A, CHANNEL	11.7	10.73
CHANNEL B, WEB PLATE	5.18	4.50
CHANNEL B, CHANNEL	11.7	10.73
TOTAL	33.98	30.67
% LOSS	1.9%	2.0%



AS-INSPECTED (CROSS SECTION 2)

ELEMENT	GROSS AREA (SQ. IN.)	NET AREA (SQ. IN.)
CHANNEL A, WEB PLATE	5.37	4.70
CHANNEL A, CHANNEL	11.7	9.55
CHANNEL B, WEB PLATE	4.73	4.11
CHANNEL B, CHANNEL	11.7	9.55
TOTAL	33.5	27.91
% LOSS	3.3%	3.6%



AS-INSPECTED (CROSS SECTION 3)