Nondestructive Evaluation of T1 Steel in the Sherman Minton Bridge

U.S. Department of Transportation
Federal Highway Administration

FOREWORD

This final report provides information to bridge owners regarding nondestructive testing of butt welds in fracture-critical members fabricated from American Association of State Highway and Transportation Officials M244 Grade 100 (ASTM A514/A517) steel, more commonly known as "T-1" steel.[1–4] This report documents an investigation undertaken to test T-1 steel in the Sherman Minton Bridge using a variety of different nondestructive methods, including, but not limited to, ultrasound, magnetic particle, and radiographic testing.

Jean A. Nehme, Ph.D., P.E.
Director, Office of Infrastructure
Research and Development

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. FHWA-HRT-23-045	2. Government Accession No.	3. Recipient's Catalog No.		
4. Title and Subtitle Nondestructive Evaluation of T-1 Steel in the Sherman Minton Bridge		5. Report Date March 2024		
		6. Performing Organization Code		
7. Author(s) Glenn Washer (ORCID: 0000-0002-2723-2296) and Robert Connor (ORCID: 0000-0002-6964-3317)		8. Performing Organization Report No.		
9. Performing Organization Name and Address Genex Systems 11848 Rock Landing Drive, Suite 303 Newport News, VA 23606		10. Work Unit No.		
		11. Contract or Grant No.		
12. Sponsoring Agency Name and Address Office of Infrastructure Research and Development Federal Highway Administration 6300 Georgetown Pike McLean, VA 22101-2296		13. Type of Report and Period Covered Final Report; April 2022—May 2022		
		14. Sponsoring Agency Code HRDI-20		
15. Supplementary Notes The contracting officer's representative was Hoda Azari (HRDI-20; ORCID: 0000-0002-7340-0035)				
16. Abstract The purpose of this document is to provide information to bridge owners on previous experience in conducting nondestructive testing (NDT) for a highway bridge constructed with American Association of State Highway and Transportation Officials M 270/M 244 (ASTM A514/517) 100 kilopounds per square inch steel.[1–4] These types of high-strength steel have been affected by hydrogen-assisted cracking, which may occur during the fabrication process because of uncontrolled hydrogen levels in the weld. The occurrence of hydrogen-assisted cracking in high-strength steels has led to several bridge closures to address cracking or fracture of nonredundant steel tension members (NSTMs), previously known as fracture critical members. The objective of this report is to present the results and an analysis of results from the NDT inspection of the Sherman-Minton Bridge, which was affected by hydrogen-assisted cracking in its NSTM tie girder. The report documents the results of performance testing by NDT inspectors working on the bridge. Researchers analyzed the correlation between NDT results and physical samples (cores) removed from the bridge and reported on the overall finding from the NDT inspection of butt welds in the bridge. These data are intended to inform bridge owners of the performance of NDT technologies applied in the field.				
17. Key Words Nondestructive testing, ASTM, steel tension members, bridge weld, butt welds, hydrogen-assisted cracking, Sherman Minton Bridge, S-M Bridge		18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161. https://www.ntis.gov		
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified		21. No. of Pages 49	22. Price N/A

Form DOT F 1700.7 (8-72)   Reproduction of completed pages authorized.

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
LENGTH				
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
AREA				
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
VOLUME				
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
NOTE: volumes greater than 1,000 L shall be shown in m^3				
MASS				
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2,000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
TEMPERATURE (exact degrees)				
°F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C
ILLUMINATION				
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
FORCE and PRESSURE or STRESS				
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
LENGTH				
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
AREA				
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
VOLUME				
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
MASS				
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2,000 lb)	T
TEMPERATURE (exact degrees)				
°C	Celsius	1.8C+32	Fahrenheit	°F
ILLUMINATION				
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m2	0.2919	foot-Lamberts	fl
FORCE and PRESSURE or STRESS				
N	newtons	2.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380.
(Revised March 2003)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AASHTO	American Association of State and Highway Transportation Officials
API	American Petroleum Institute
ASNT	American Society for Nondestructive Testing
ATS	Applied Technical Services
AWS	American Welding Society
CJP	complete joint penetration
COV	coefficient of variation
CVN	Charpy V-notch
DOT	department of transportation
FAD	failure assessment diagram
FFS	fitness for service
FHWA	Federal Highway Administration
FNR	false negative rate
INDOT	Indiana Department of Transportation
ksi	kilopounds per square inch
MT	magnetic particle testing
NDE	nondestructive evaluation
NDT	nondestructive testing
NSTMs	nonredundant steel tension members
Q/T	quenched and tempered
QC	quality control
RT	radiographic testing
SDH	side-drilled hole
S-M	Sherman-Minton
TA	Technical Advisory
TPR	true positive rate
UT	ultrasonic testing

INTRODUCTION

PURPOSE, GOALS, AND OBJECTIVES

The purpose of this document is to provide information for bridge owners based on previous experience in conducting nondestructive testing (NDT) for a highway bridge constructed of American Association of State and Highway Transportation Officials (AASHTO) M 270/M 244 (ASTM A514/517) 100 kilopounds per square inch (ksi) steel.[1–4] This high-strength steel has been affected by hydrogen-assisted cracking, which may occur during the fabrication process because of uncontrolled hydrogen levels in the weld. The occurrence of hydrogen-assisted cracking in high-strength steels has led to several bridge closures to address cracking or fracture of nonredundant steel tension members (NSTMs), previously known as fracture-critical members. For example, the Sherman-Minton (S-M) Bridge was closed for repairs on September 9, 2011, due to cracking in a vertical butt weld. The bridge closure led to the release of the Federal Highway Administration (FHWA) Technical Advisory (TA) 5140.32, which recommended verification of the soundness of butt-welds in NSTMs fabricated from AASHTO M 270/M 244 (ASTM A514/517) 100 ksi steel through visual inspection and NDT unless such verification had been previously completed.[5] More recently, the I–40 Hernando DeSoto Bridge closed unexpectedly for repair on May 11, 2021, due to a fracture in an NSTM from hydrogen-assisted cracking. The closure of the Hernando DeSoto Bridge led to the release of an FHWA memorandum[6] expanding on TA 5140.32 by requiring State departments of transportation (DOTs) to verify the soundness of butt welds on NSTMs fabricated from AASHTO M 270/M 244 (ASTM A514/517) 100 ksi steel using NDT, unless verification had been previously documented.

The objective of the report is to present results and an analysis of the results from the NDT inspection of the S-M Bridge, which was affected by hydrogen-assisted cracking in its NSTM tie girder. The report comprises the following actions:

- Documents the results of NDT performance testing by inspectors working on the bridge.

- Analyzes the correlation between NDT results and physical samples (cores) removed from the bridge.

- Presents the overall findings from NDT inspection of butt welds in the bridge.

These data are intended to inform bridge owners regarding the performance of NDT technologies applied in the field. The report begins with a description of the S-M Bridge design, materials used, and the different details affected by hydrogen-assisted cracking. A brief description of the NDT technologies used during the detailed inspection of the bridge is included. The NDT technologies used included magnetic particle testing (MT), ultrasonic testing (UT), and radiographic testing (RT). The report describes the results of NDT testing on the bridge, including quantitative data and the results of an analysis of the outcome of performance testing conducted by inspectors on the bridge, an analysis of NDT results for 21 cores removed from the bridge, and overall statistics from the inspection of the critical tie girder welds.

BACKGROUND

On September 9, 2011, the S-M Bridge closed because of a crack in the tie girder. The bridge remained closed for several months to enable completion of further inspections, to develop a repair plan for the bridge, and to execute repairs. This section of the report describes the bridge and details key types of cracking found, such as complete joint penetration (CJP) butt welds in the tie girder, welds in lateral plates attached to the web of the tie girder (i.e., shelf plates), corner welds, and lugs within the tie girder. Inspectors found that cracking affected each connection initiated at the time of fabrication.

BRIDGE DESCRIPTION

The S-M Bridge[7] carries I–64 over the Ohio River between New Albany, IN, and Louisville, KY. Construction began in 1960 and was completed in 1961. The interstate has six traffic lanes on the double-decked roadway, three westbound lanes on the upper deck, and three eastbound lanes on the lower deck. Figure 1 shows an elevation view of the bridge.

© 2011 Purdue University.

Figure 1. Photo. Elevation view of the S-M Bridge from the Indiana shoreline.

The main river crossing consists of two 800-ft-long simply supported tied arches. Each tie is composed of 100 ksi steel components that are bolted and welded together. The ties are supported by 2.5 inch-diameter steel cables or, in some cases, by vertical truss members. Rigid frames attached to the tie support the upper and lower roadway.

Engineers conducted an in-service inspection of the structure and identified cracking in the steel at several locations on the NSTM tie girders, such as transverse butt welds, longitudinal corner welds, welds attaching lateral gusset plates, and internal "lugs" at the diaphragms. Completion of a fitness-for-service (FFS) evaluation of the tie girder welds was one component of a comprehensive inspection conducted on the bridge. By using the results from material test data, residual stress profiles, and linear elastic fracture mechanics models, the engineers generated a failure assessment diagram (FAD) to determine whether cracks and crack-like flaws in certain details of the bridge were within an acceptable range, meaning the bridge would be expected to

remain stable. The engineers concluded from the FFS evaluation that corrective measures were required because the tolerable critical crack size would not likely be detected in the field. Additionally, inspection reports had already indicated surface crack lengths in excess of the critical crack length.

Material and Fabrication

According to the design drawings, the tie girder is fabricated from a steel similar to T-1 steel, a quenched and tempered (Q/T) product that was the forerunner to ASTM A514 steel. T-1 is a trade name for a steel originally produced and developed by U.S. Steel Corporation and generally possesses yield strengths of 100 ksi with ultimate strengths ranging between 110 to 130 ksi. Today, the properties for all bridge steels must meet the requirements of AASHTO M 270, or its equivalent, ASTM A709.[8] The only grade of 100 ksi steel currently in those two material specifications is HPS (high-performance steel) 100W, which is different from the older grade 100 steel represented by ASTM A514.[1–4]

A review of the shop drawings for the S-M Bridge indicated the steel in the tie to be N-A-XTRA and not T-1, which was a trade name of a T-1 steel equivalent produced by the Great Lakes Steel Corporation. Although both steels appear to have been used in the bridge, it is not completely clear whether any T-1 steel was used in the tie because engineers only had limited sheets from the shop drawings available for review. Throughout this document, the term "T-1" is used to refer to any of the 100 ksi Q/T steels.

Of note is the date, 1960, on the design drawings, which predates fatigue design provisions. The specifications by the American Association of State Highway Officials, which is now AASHTO, did not include the first fatigue design provisions until 1965. These original provisions were limited; AASHTO introduced more complete updated provisions in 1973 and 1974. Consequently, the fatigue limit state was never considered in the original design since the design predated the modern provisions for fatigue design. Furthermore, the now common AASHTO/American Welding Society (AWS) D1.5 fracture control plan (FCP)—which governs material, fabrication, and shop inspection of welds in NSTMs—was not introduced until about 1978.[9] Regarding material, minimum Charpy V-notch (CVN) impact energies are now specified for all steel used in bridges.[8] Since the S-M Bridge was designed in 1960, no minimum CVN requirements were in place at the time of fabrication.

Engineers have long known that welding of T-1 steel is highly susceptible to hydrogen-assisted cracking. Observations of several bridges found similar cracking, a few of which are tied-arch bridges. Other terms for hydrogen cracking include "cold cracking" or "delayed cracking" because often the cracking occurs after the weld cools, generally within 48 h after welding. This cracking can be controlled using specialized procedures that attempt to eliminate hydrogen from entering the weld during the welding process. As mentioned, these cracks usually occur within 48 h of placing the weld, but often they occur immediately after completion and cooling of the weld pass. Unfortunately, the need to control hydrogen was not as well understood during the fabrication of the S-M Bridge. Today, AASHTO/AWS D1.5 requires much more rigorous control of hydrogen in all welding and has greatly reduced the likelihood of such issues in modern fabrication when using such high-strength steels.[9]

Cracking in the S-M Bridge

Regular inspections conducted during the service life of the bridge prior to 2011 found surface-breaking cracks on several different details on the bridge. The indepth inspection that led to the bridge closure found a significant number of additional cracks and other weld flaws. The following sections describe cracking in each different type of detail. The most significant cracking in terms of the safety of the bridge were cracks along the vertical butt welds used to join plates of varying thicknesses at the joints (figure 2).
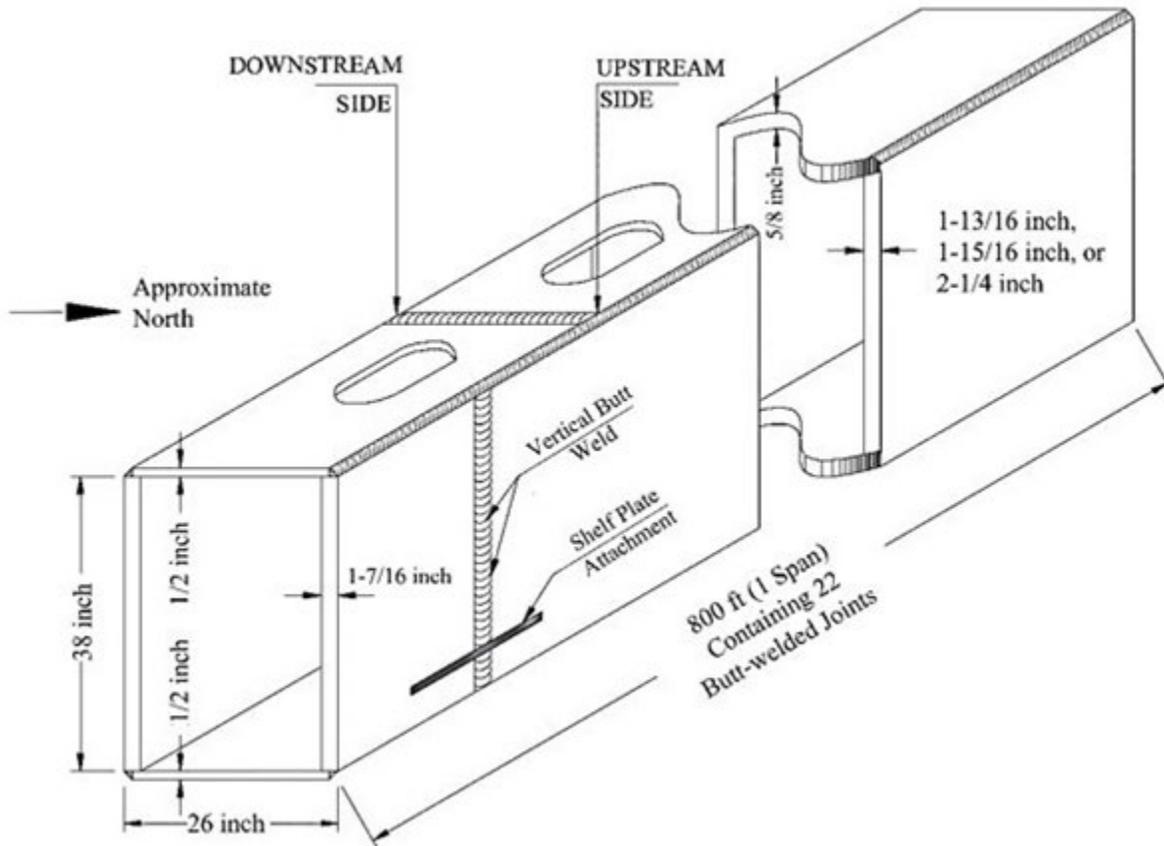
Figure 2. Schematic. Typical vertical butt weld at the transition of plate thickness.

Tie Girder Butt Welds

Figure 2 shows cracks observed in the tie girder at the butt welds, which are vertical welds that join plates of different thicknesses. Figure 3-A and figure 3-B show cracks that were oriented both parallel and perpendicular to the primary stress. Figure 3-A shows two lateral cracks in the weld (parallel to the axial stress), with closeup views of the MT indication for the crack. Figure 3-B shows a vertical crack (perpendicular to the axial stress) is approximately 3 inches in length with stop holes drilled at the top and bottom ends of the crack to arrest crack growth.

Based on the results of metallographic studies, the cracking observed on the S-M Bridge was due to hydrogen cracking. Since hydrogen cracking typically forms within 24 to 48 h after completion of the welding process, the cracks observed in the bridge may have been present since the bridge was fabricated. The quality control (QC) processes that the engineers used during the fabrication did not detect the cracking, which was ubiquitous among welds in the S-M Bridge.
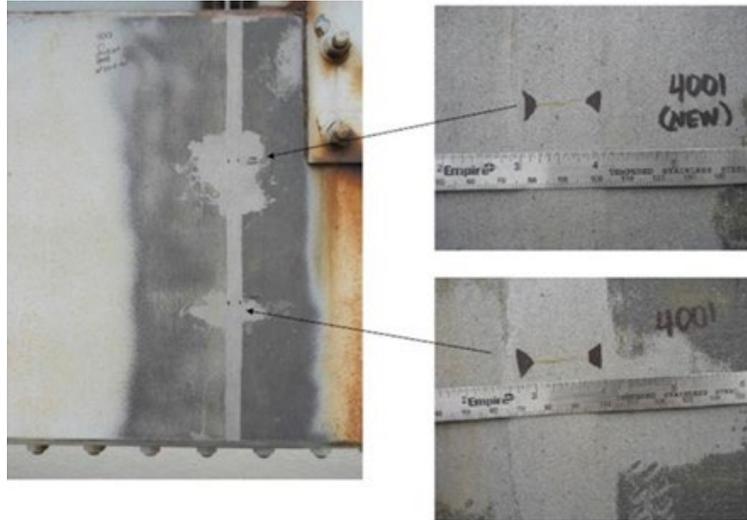
© 2011 Michael Baker International.

A. Typical transverse cracks in CJP butt weld in tie girder.

© 2011 Michael Baker International.

B. Typical longitudinal crack in CJP butt weld in tie girder.

Figure 3. Photos. Typical transverse and longitudinal cracks.

According to the design and shop drawings, the tie girder butt welds are double-V-groove CJP welds. The design drawings for the tie indicate the vertical plate thickness is 1-7/16 inches between floor beam frames. However, at the floor beam frames, the thickness of the vertical plate is either 1-13/16 inches, 1-15/16 inches, or 2-1/4 inches. The thickness of the top and bottom plates is half an inch between floor beam frames and increases to 5/8 inch in the region of the joint, as shown in figure 2.

Longitudinal Attachments to Tie Girder Vertical Plates at Floor Beam Frames

Inspectors found transverse cracks (perpendicular to axial stress) in the shelf plates used to attach gusset plates connecting the diagonal lateral braces. These plates were also made from T-1 steel and were attached with single bevel groove CJP welds. The shelf plates were welded to the tie girder splice plates and transverse to the vertical butt welds in the tie, also made of T-1 steel.

Cracks ran along the length and ends of these welds, as shown in figure 4, which shows an overall view of a shelf plate with three crack indications. Closeup views of each of the three cracks are shown with arrows to the crack location on the shelf plate. All cracks along the attachment in the weld were immediately adjacent to the base metal. Those at the ends were at the weld toes; hence, they extended into the base metal of the tie girder.

© 2011 Michael Baker International.

Figure 4. Photo. Typical cracking in fillet welds that attached longitudinal gusset plate to tie girder.

Tie Girder Cracks in Longitudinal Fillet Welds

The engineers used longitudinal fillet welds to join the four plates making up the rectangular tie girder box cross section. Cracks were reported at some locations in these welds, as shown in figure 5. In terms of volume, these fillet welds comprised most welding on the bridge. In total, about 2.4 mi of corner fillet welds are on the structure.

© 2011 Michael Baker International.

Figure 5. Photo. Cracks in fillet welds at the corner of a tie girder.

Cracked "Lugs" Within Tie Girder

Cracks also occurred within the tie girder at weld locations connecting internal diaphragms to the tie girder. These cracks appeared to be contained within the fillet welds themselves. However, due to the difficulty in examining these welds, the inspectors were uncertain whether some cracks extended into the base metal of the tie girder.

Overall, the inspectors found a significant number of cracks during typical inspections throughout the service life of the S-M Bridge. The Indiana DOT (INDOT) conducted a new round of inspections on the S-M Bridge beginning in April 2011. During this inspection period, the engineers found a 2.5-inch crack in one of the butt welds on the NSTM tie girder. The finding occurred while they were retrofitting the shelf plate connections (lateral bracing-to-tie girder connections). The process of retrofitting required removing 88 shelf plate connections to replace the welded detail with a bolted detail. Based on the FFS analysis, the engineers concluded that a crack with the same characteristics could cause the tie girder to fracture at a service temperature of 30 F (−1 ºC). As a result, INDOT closed the S-M Bridge to traffic on September 9, 2011, and it remained closed until February 17, 2012. During that time, the tie girder retrofit included the addition of more plates over the tie girder. The intent was that the additional plates could carry the bridge loads if one of the cracks in the bridge resulted in a fracture in the original steel of the tie girder.

Description of Nondestructive Evaluation (NDE) Technologies Used on the Bridge

This section of the report briefly describes the different NDT technologies implemented during the inspection of the S-M Bridge. The NDT technologies included MT, UT, and RT.

MT

MT consists of inducing a magnetic field in the surface of the steel and assessing leakage of that field that may occur if a crack is present within the induced field. The MT process begins by applying finely divided iron particles to the surface. The iron particles, which are attracted to the leaking field, collect at a surface-breaking crack or surface defect and create a visual "indication." The iron particles include a dye to make the particles red, yellow, or gray to create a color contrast with the material being tested.

Cracks detected by MT are typically surface breaking and, therefore, inspectors may detect them visually. MT improves the ability to detect these cracks due to the indications created on the surface by accumulating iron particles, which are observed by the inspector.

Field inspections with MT typically employ MT yokes that have a wire coil in the handle and two steel legs that couple the magnetic field into the steel member under inspection, as shown in figure 6. A crack in the steel material disrupts the flow of the magnetic field, resulting in flux leakage into the air. Cracks positioned orthogonally to the magnetic flux create the largest disturbance. When cracks are parallel to the direction of the magnetic field, the field disturbance is minimal, and the sensitivity reduced. For inspection, most MT procedures require orienting the yoke in two orthogonal positions for each area to be inspected.
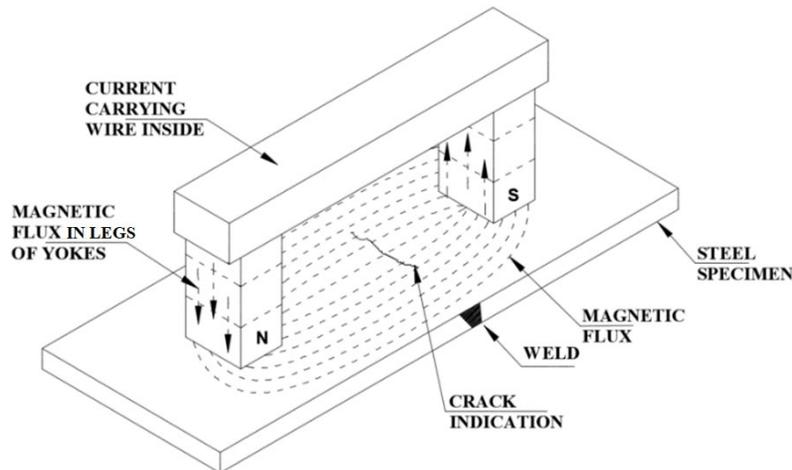
© 2022 University of Missouri.
N = north; S = south.

Figure 6. Schematic. MT for crack detection.

9

UT

Conventional UT consists of launching a refracted shear wave into the steel and assessing signals from internal discontinuities that reflect acoustic energy. Those who work in NDT commonly refer to reflections from internal discontinuities as "indications" because the source of the reflection may be a flaw or could be an innocuous reflector such as a geometric feature.

The transducer used for UT, shown in figure 7, consists of a piezoelectric crystal that launches a longitudinal wave at ultrasonic frequencies, typically 2.25 MHz for weld inspection. Inspectors fix an acrylic wedge to the transducer and couple it with a liquid, such as oil, water, or ultrasonic couplant. This process converts the longitudinal wave launched into the acrylic wedge to a shear wave through a mode conversion that occurs when the wave passes from one material to another. Following is the equation for the refraction of waves:

$$\theta_2 = \sin^{-1}\left(C_2 \frac{sin\theta_1}{C_1}\right)$$

(1)

Where:
θ_2 = refracted angle in steel.
C_2 = shear wave velocity in steel.
θ_1 = refracted angle in steel.
C_1 = longitudinal wave velocity in acrylic wedge.

To determine the depth of the indication in the weld using geometry and the known wave velocity in the steel, inspectors used the refracted angle of the beam. The refracted angles commonly used for inspection are specified angles with efficient mode conversion to introduce acoustic energy (waves) into the steel. Variations from the specified angles reduce the amount of energy introduced and, hence, may affect the amplitude of signals reflected from indications in the weld. Variation in the velocity of shear wave in the steel affects the refracted angle (figure 7), and, therefore, the amplitude of signal reflected from indication in the weld.

An indication is assessed to determine if it constitutes a "rejectable indication" by comparing the amplitude of a wave reflected from a 0.06-inch-(1.5 mm)-side-drilled hole (SDH), which provides a consistent edge reflection regardless of the incident angle. The amplitude of an indication in the weld is adjusted to account for the attenuation of the wave as it propagates through the steel. In the current AASHTO/AWS D1.5M/D1.5 *Bridge Welding Code*, the amplitude adjustment is 2 dB × (path length-1 inch).[9] The amplitude rating of a certain indication is the difference between the adjusted amplitude of the indication signal and the amplitude of the SDH. As a result, the decibel rating typically provided in a standard UT form represents the gain needed to make the indication amplitude the same as the SDH amplitude. If the adjusted indication signal is smaller than the SDH signal, then the gain needs to be added to the adjusted indication signal to match the SDH. If the adjusted indication signal is larger than that produced by the SDH, the gain needs to be subtracted from the indication signal to match the SDH signal. The amplitude of reflections reported from UT are large when the amplitude values are small, e.g., a −2 dB indication is larger in amplitude than with a +2 dB indication.
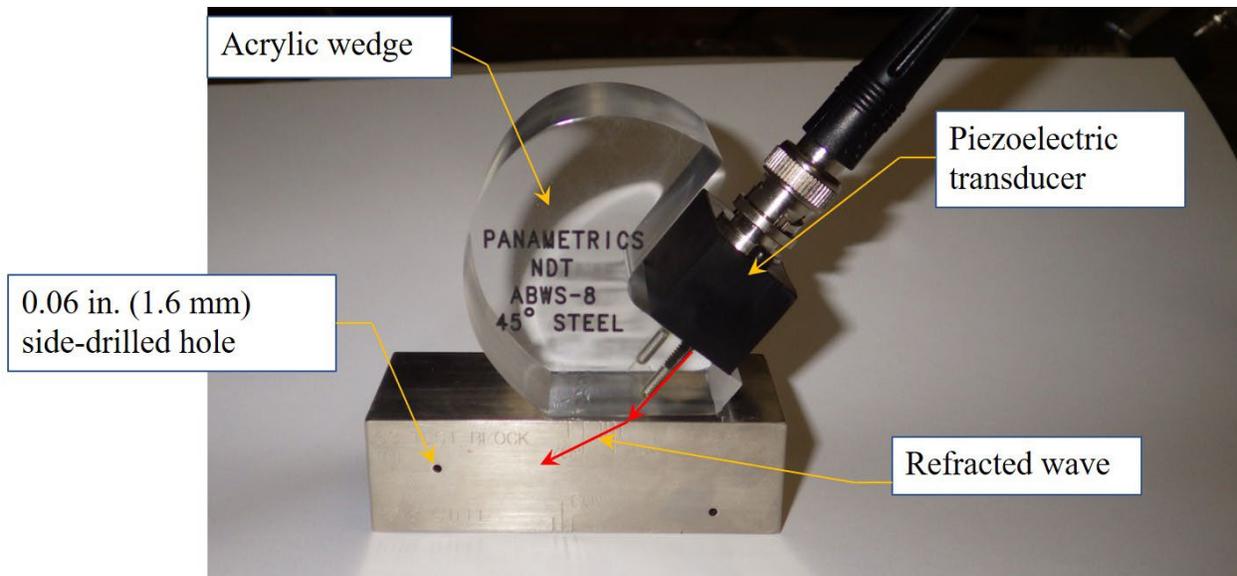
© 2022 University of Missouri.
In. = inch.

Figure 7. Photo. Ultrasonic transducer used to detect indications in welds.

The amplitude of the reflected waves is compared to the amplitude from the SDH to characterize an indication (presumably a flaw) in the weld, using the following scheme:[9]

- Class A large flaws—Rejectable.

- Class B—Rejectable when length is greater than 0.75 inch (20 mm).

- Class C—Rejectable when length is greater than 2 inches (50 mm) in the middle half or 0.75 inch (20 mm) in the top of the bottom quarter of the weld.

- Class D—Acceptable indication.

The acceptance criteria for indications detected with UT are illustrated in figure 8 for a 70 degree-shear wave in steel plates of different thicknesses. The figure shows the amplitude of the signals for each classification of indications from class A, "large indications," through class D, "acceptable" indications. Of note is that in the figure, the amplitudes are represented as directly compared with the SDH, e.g., a −8 dB signal is −8 dB less than the SDH signal and the −4 dB signal is greater in amplitude as compared with the −8 dB indication. The figure 8 graph provides an easier way to interpret acceptance criteria for indications detected than the conventional method of reporting in which the signs are inverted. For example, the threshold value in the AASHTO/AWS D1.5M/D1.5 *Bridge Welding Code* for a class A indication for plates 1.5–2.5 inches thick is +4 dB.[9]

Of note is that the difference in amplitude between a class A indication and a class D indication for plates 1.5–2.5 inches thick is only 5 dB. The difference between a class A, which is always rejectable, and a class B indication, which may be acceptable depending on the measured length of the indication, is only 1 dB. As a result, the amplitude of the signal reflected from a flaw in a

11

bridge weld is critical in terms of determining if the given indication is acceptable or rejectable. Therefore, the influence of the wave velocity and the attenuation in the steel are critical to the appropriate disposition of indications found in welds. Steps to mitigate these effects include installing reference blocks formed from material that is acoustically like the material being tested in terms of velocity and attenuation.
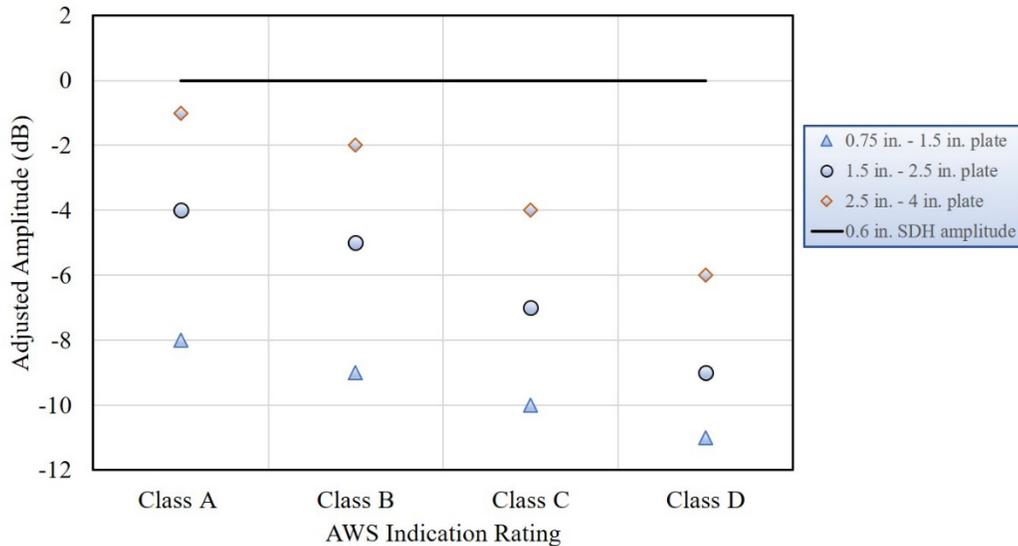
© 2022 University of Missouri.

Figure 8. Graph. AASHTO/AWS D1.5M/D1.5 indication rating requirement for 70-degree shear waves for plates ranging in thickness from 0.75 inch to 4 inches.[9]

RT

RT produces images depicting the density of the material being tested. Testing identifies indications according to contrast appearing on the image. An image quality indicator verifies that sufficient contrast exists on the image to identify indications.

The photons that penetrate the material can come from different sources. Decaying isotopes such as cobalt-60 or iridium-192 produce gamma rays. X-ray machines or linear accelerators also can produce photons. The general test arrangement for RT consists of placing the photon source on one side of the specimen and placing the film or detector on the other side of the specimen, as shown in figure 9. The film captures the legacy of the photons that penetrate the material. As photons propagate the material, they attenuate because of scattering from interactions with ion cores and absorption in the material. Low-density materials scatter and absorb photon energy less frequently than high-density materials. Materials with high atomic numbers (larger cores) result in more scattering losses than materials with low atomic numbers. For example, using lead as a shield is effective because of its high density and large atomic number.

Visual inspection of the resulting film depiction of the density of the materials is used to identify indications in RT. The detection of flaws, such as porosity or a lack of fusion, occurs because the material has a lower density where the porosity has formed. Because the attenuation of photons relies on scattering from ion cores in the material, vertical cracks aligned with the direction of

the photons penetrating through the material may not appear as detectable indications. Generally, RT is expected to be most sensitive to volumetric weld flaws, such as slag inclusions, porosity, and a lack of fusion in the weld. Many inspectors consider RT less effective for detecting cracks than UT, particularly if the crack is aligned parallel to the direction of the photon path from the source to the detector.
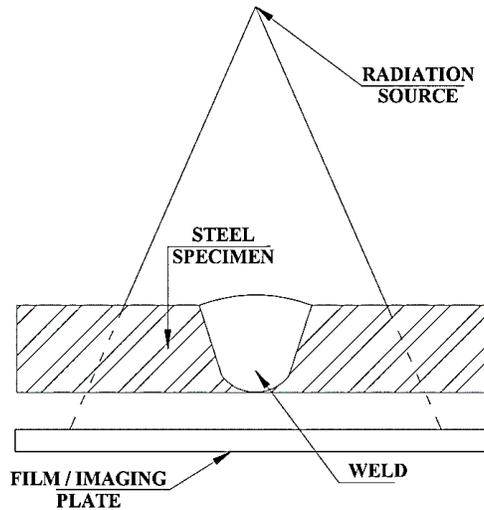
© 2022 University of Missouri.

Figure 9. Schematic. RT testing of welds.

The exposure time necessary to produce an image using RT varies based on the density and thickness of the material being tested and the intensity of the source. Exposure times typically range from several seconds to several minutes. As a result, the application of RT in the field for a large structure vibrating due to traffic loading or to wind can produce some loss of clarity in the image as compared with application to a static structure such as a plate in a fabrication shop.

Summary of NDT Techniques

Table 1 summarizes the capabilities of three NDT technologies used on the S-M Bridge. The table lists the types of flaws detected, the surface preparation typically required, and the access requirements for the technologies. For example, MT requires surface preparation to remove paint and clean the area to avoid iron particles being trapped in coating defects, to ensure adequate coupling of the magnetic field into the material to provide sensitivity, and to allow mobility of the iron participles to collect at an indication where the magnetic flux leaks into the air. The MT technology is sensitive only to surface-breaking cracks because the penetration of the magnetic flux is very close to the surface for a ferromagnetic material such as carbon steel. Consequently, MT can detect surface-breaking cracks only on the surface at the site of yoke application. The cracks appear as visual indications on the surface; hence, interpretation of the results is relatively simple and depends mostly on the inspector's vision capabilities.

Both UT and RT technologies provide through-thickness capabilities. However, the technologies are most sensitive to different types of flaws. UT most easily detects planar flaws (i.e., cracks) because the flat face of the crack reflects a significant amount of acoustic energy. Assuming the orientation of the crack is normal to the path of the wave, acoustic energy reflects back to the

transducer and, subsequently, the NDT technician assesses it. If the crack is not normal to the wave path, it is possible that the energy reflects away from the transducer and, consequently, the technician may not detect and assess it or may assess it as having a lower amplitude as compared with a crack that is normal to the wave path. For this reason, the scanning pattern used for conventional UT requires the technician to rotate the transducer ±10 to maximize the amplitude of the signal reflected from features in the weld such as flaws. UT is less sensitive than RT to volumetric defects such as porosity or slag inclusion because of the circular nature of the flaws. Those flaws tend to scatter the acoustic energy and, consequently, produce small reflections for assessment according to the threshold amplitudes illustrated in figure 8. The sensitivity of the amplitude of reflected waves to the morphology and orientation of the feature creates the indication and other factors such as the surface roughness and pressure applied to the transducer. The interpretation of UT results is complex compared to either MT or RT results.

RT is more sensitive to volumetric flaws than UT because the flaws represent areas of reduced density that appear clearly on the radiographic image produced. RT has the advantage of requiring very little surface preparation since the penetrating radiation passes easily through low-density coatings with very little attenuation. However, RT requires access to both sides of the specimen since the photons must pass through the thickness of the material and be detected on the opposite side, as illustrated in figure 9. As noted previously in the RT section of this document, RT is least sensitive to cracks aligned with the direction of the penetrating radiation because cracks oriented in this direction may produce very little attenuation and, consequently, may not appear clearly in the radiographic image. This situation is particularly true if the cracks are tight. Interpreting RT results is more complicated than MT, but less complex than UT, and relies on the interpreting contrast in the resulting image as compared to standards (length, size) for identifying rejectable indications.

Table 1. Summary of NDT technology capabilities.

NDT Technologies	Flaws Detected	Surface Prep	Access Requirements	Complexity of Interpretation
MT	Surface-breaking cracks only	Paint removal	Single-sided, inspection surface only	Low
UT	Cracks Volumetric flaws	Paint removal	Single-sided, through-thickness	High
RT	Volumetric flaws Cracks	Light cleaning to remove debris	Two-side access, through thickness	Moderate

Performance Testing on the S-M Bridge

Due to the critical need to ensure inspections on the S-M Bridge were effective, the project team verified the qualifications the of inspectors working on the bridge. The NDT had to comply with the AASHTO/AWS D1.5 in *Bridge Welding Code,*[9] which requires inspectors conducting inspections to be qualified as American Society for Nondestructive Testing (ASNT) Level II technicians.[10] The standard method for obtaining this qualification, according to SNT-TC-1A guidance, is an employer-based certification process in which employers develop and administer their own qualification examinations. The guide specifies general training and experience

requirements. Individuals qualified as Level III technicians can develop procedures for conducting some inspections and determining the qualifications of Level I and Level II technicians.[10] Individuals qualified as Level II technicians achieved 80 h of training, 840 h of experience with a specific method such as UT, and 1,600 h of experience with NDE.

The purpose of performance testing is to verify that the training and experience documented through the required certification have resulted in adequate performance in the field. In this case, the desired outcome was to reliably detect internal flaws in the welds of the S-M Bridge. Performance testing assessed two NDE technologies used to detect defects in the welds, UT and MT. RT also occurred but was not part of the performance tests.

High variability in test results can occur among certified inspectors who may implement the procedures differently, have different levels of experience and expertise, and may interpret responses differently.[11–14] As a result, one cannot assume the reliability of the implementation of the technology among any group of certified inspectors. Performance testing ensures inspectors have a minimum level of performance to provide acceptably consistent results when implementing the procedures. Within a performance test, an inspector implements the required procedures on certain test specimens containing known flaws. The evaluation focuses on the inspector's ability to identify and assess these flaws to determine if the inspector is properly implementing the procedure and interpreting results. Inspectors unable to implement the test procedures successfully within the performance test become disqualified, at least until such time as they can demonstrate sufficient capabilities.

The goal of the performance testing was to improve the quality of NDT results from inspections conducted on the S-M Bridge. The objective of the testing was to ensure individual inspectors conducting NDT were capable of successfully implementing the specified procedures to detect and assess cracks in welds.

Specimens Used for the Performance Testing

Performance testing typically utilizes test plates of similar material and geometry to the specimens being inspected. The test plates used for the performance test have discontinuities within the specimens of similar size and shape to those anticipated or required for the inspection to meet its objectives. Machining, such as cutting a notch in the steel, electrical discharge machining a notch, or drilling an SDH produces artificial discontinuities in the test plate. Implanted cracks or volumetric defects can provide a more representative discontinuity for flaws in welded specimens. Performance testing on the S-M Bridge used specimens with implanted cracks.

Researchers at FHWA's NDE Laboratory provided the test plates used in the performance testing on the S-M Bridge. Figure 10 shows a drawing of a typical test plate with two implanted flaws. The test plate shown in figure 10 contains two crack-like flaws that lie along the fusion line of the weld. The steel test plates typically contained implanted flaws of varying length and depth within a groove weld placed along the centerline of the plate. The performance test plates had both subsurface flaws and surface-breaking flaws. The MT analysis used only specimens with surface-breaking flaws because this technology detects only surface-breaking cracks or very near-surface cracks. The thickness of the test plates used was smaller than the welded plates in

the S-M Bridge. Although obtaining plates of similar geometry to the S-M Bridge members would have been preferable, time constraints precluded obtaining new specimens that matched the S-M Bridge plate geometry.
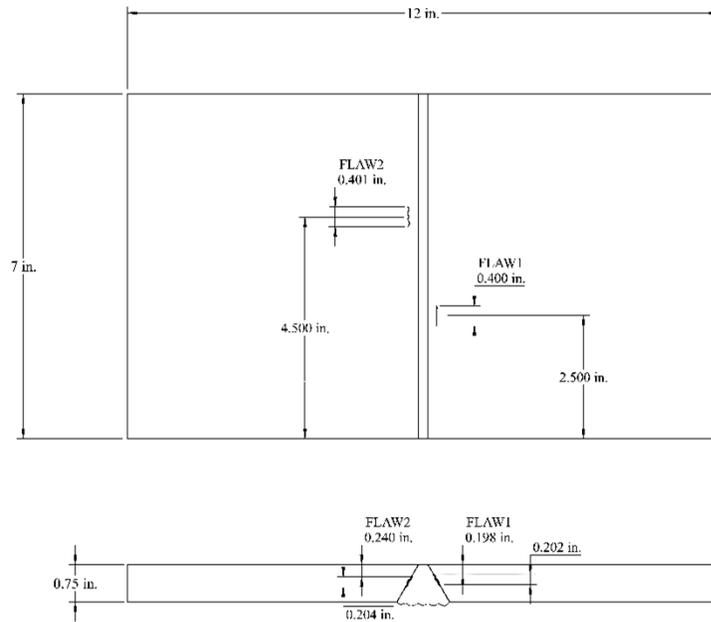
© 2022 University of Missouri.

Figure 10. Schematic. Typical test plate used for performance testing showing the plan for implanted flaws.

Performance testing conducted on the S-M Bridge used different specimens to assess MT and UT. Table 2 lists the dimensions and numbers of flaws in the plates used for performance testing. Notes in the table indicate if the test plate was used only for UT, only for MT, or for both. To evaluate MT, the inspectors used only test plates SMB-4 and SMB-6. The plate SMB-4 had one longitudinal flaw on the surface of the test specimen that was tested using MT. MT could not detect a second flaw in the plate because it was subsurface. The UT performance test assessed this flaw. Plate SMB-6 had two flaws on the surface tested using MT, one longitudinal crack along the centerline of the weld, and one crack transverse to the weld axis.

The performance tests of UT inspectors used test plates with the objective of including a range of crack sizes from small (about 0.4 inch) to large (about 2+ inches) at different depths in the test plates. These test plates included flaws ranging in length from 0.4 to 2.4 inches located at the toe of the weld (on the inspection surface), the root (opposite surface), and middepth (subsurface) along the fusion line. The performance testing occurred at two different times with the initial performance testing conducted in May 2011 and additional performance tests conducted in September 2011. Table 2 lists all the test plates used for performance testing. The initial tests that occurred in May 2011 used only test plates SMB-2, SMB-5, and SMB-10. The second performance test did not use these plates for various reasons. SMB-2 included very large flaws created by a fabricator such that no records of the actual dimension of the flaws were made, making analysis difficult. SMB-5 had an additional indication not documented in the certification sheets for the specimen, so the second set of performance tests did not use it.

16

Finally, SMB-10 contained a flaw similar to SMB-3, thus the second test omitted it due to time constraints. The results reported here for UT include only those test plates—SMB-3, -4, -8, and -9—used in both performance tests.

Table 2. Description of specimens used for performance testing on the S-M Bridge.

Specimen Designation	Overall Specimen Dimensions (inches)	Length (inches)	Location
SMB-2[a]	15 × 1 × 1	Undetermined[c]	Subsurface
SMB-3[b]	7 × 12 × 3/4	1.2	Surface breaking at toe, longitudinally oriented
SMB-4	12 × 12 × 1/2	0.94[b]	Surface breaking at root, longitudinally oriented
		2.4	Surface breaking at toe, longitudinally oriented
SMB-5[a]	7 × 12 × 3/4	0.4	Subsurface, longitudinally oriented
		0.4	Subsurface, longitudinally oriented
SMB-6[d]	12 × 12 × 1/2	1.45	Surface breaking, longitudinally oriented
		0.43	Surface breaking, transversely oriented.
SMB-8[b]	7 × 12 × 3/4	—	—
SMB-9[b]	7 × 12 × 3/4	0.4	Subsurface, longitudinally oriented
SMB-10[a]	7 × 12 × 3/4	1.2	Surface breaking at toe, longitudinally oriented

—Indicates the specimen had no flaws, thus, no length or location to report.
[a]Used only for initial UT performance test, May 2011.
[b]Used only for UT performance tests.
[c]Used to indicate multiple flaws in the plate, each with an undetermined length.
[d]Used only for MT performance testing.

Criteria for Performance Testing

Performance testing of inspectors requires criteria to judge the acceptance of an inspector's performance during the test. Performance testing criteria for the implementation of the AASHTO/AWS D1.5M/D1.5 standard are not available because performance testing has not been a normal practice for implementing the code requirements.[10] Criteria for the evaluation of UT are available for different weld inspection codes.[14,15] For example, AWS D1.8 has criteria for performance testing of inspectors that requires at least 20 flaws of relevant size distributed in test plates.[15] Successful completion of the test includes detecting at least 87 percent of the flaws, 15 percent or less false calls, and an overall inspector rating derived from a combination of true positives (correct flaw detection) and false calls. The American Petroleum Institute (API) also has criteria available for ultrasonic examination of offshore structural fabrication.[16] The API criteria derives inspection scoring based on the length of flaws detected as compared with the actual flaw lengths in the test plates used for the testing. The length of flaws detected in excess of the actual flaw length in the test plates is used as a measure of false calls.

17

The project team reviewed and considered the procedures for assessing inspector qualifications in developing the performance testing for the S-M Bridge. However, the criteria were not used for the S-M Bridge performance for several reasons. First, the urgency of the testing did not provide sufficient time to develop the necessary test plates to meet the required population of flaws to implement the criteria. Second, because the inspectors had no experience using these criteria for bridge weld inspection, those in charge of the inspection believed that the requirements might be too stringent for typical inspectors to be successful. Additionally, the focus of the testing was the detection of flaws and accurate length measurements that did not significantly undersize a flaw length. Therefore, implementation of successful testing in the field required criteria that focused on the detection of flaws within a small population of flaws mandated by AWS D1.8.[15] The developed performance testing criteria included key elements of the AWS D1.8 inspector qualification requirements, such as specified accuracy in flaw length, number of missed flaws, and number of false calls, but the team customized these requirements for the task of inspecting the S-M Bridge.

Because the bridge had NSTMs, the team emphasized the detection of flaws in the test plates to ensure that inspectors did not miss flaws in the bridge during the NDT and did not underestimate the length of the flaw. For both MT and UT, the criteria required detecting and reporting all known flaws in the test plates. For MT, the criteria included detecting each of the known flaws and estimating the length of the flaw within the interval (−0.25, +0.75 inch) relative to the known length of the implanted flaw.

The length estimate provided by the AWS UT procedure utilized during the inspections was known to have limitations, particularly when the flaw length is smaller than the dimension of the ultrasonic transducer used. As a result, the required length accuracy for UT varied according to the length of the flaw, as shown in table 3. The smallest flaw (0.4 inch) had length accuracy requirements. For medium-sized flaws greater than half an inch and less than 1 inch, the expected accuracy was −0.25 inch, +1.0 inch. For large flaws, the expected accuracy was −0.25 inch, +0.75 inch. During the test, each inspector could have two false calls and one exception to these limits, providing the exception did not exceed twice the identified limits. The definition of a false call is a reported flaw where no known flaw in the area existed. The tolerance on the location of the indication was ±0.75 inch along the longitudinal axis of the weld.

Table 3. Table showing length accuracy requirements for performance testing on the S-M Bridge.

Category	Flaw Size (inch)	Length Accuracy (inch)
Small flaws	≤0.5	No requirement
Medium-sized flaws	>0.5 and <1.0	−0.25, +1.0
Large flaws	>1.0	−0.25, +0.75

Results From MT Performance Testing

The performance testing results of the testing of inspectors for the MT methods resulted in 2 failed tests and 14 passed tests for the 15 inspectors tested. One inspector failed the test due to a missed flaw, under sizing a large flaw, and not following reporting procedures; this inspector was retrained and retested and passed the subsequent performance test. As a result, 16 total

performance tests were conducted. The second inspector failed the performance test due to grossly overestimating the size of a flaw by approximately equal to 70 percent; this inspector chose not to retest and, subsequently, did not perform inspections using MT on the bridge.

Figure 7 shows the length measurements provided by the inspectors during the performance tests. The horizontal axis in the figure is the actual length of the flaw, and the vertical axis shows the measured values reported by the inspectors. A solid triangle represents the average length measurements, and an open circle indicates each individual measurement. A 1:1 line drawn on the figure shows the exact correlation of measured and actual flaw length for illustration purposes. As shown in the figure, the reported results both overestimated and underestimated the length of the flaw as compared to its actual length.

The team analyzed these data to quantify the accuracy of the MT technique as applied by the inspectors successfully completing the performance tests. These data provide insight regarding how accurate a typical MT inspection would be in the field based on the sample of 15 inspectors' testing assessed on the S-M Bridge. Table 4 shows the results from the successful performance tests. The data in the table include the average error, which is the average of the difference between the actual length of the implanted flaw and the reported length from each inspector. As shown in figure 11, reported flaw lengths are sometimes less than and sometimes greater than the actual length.
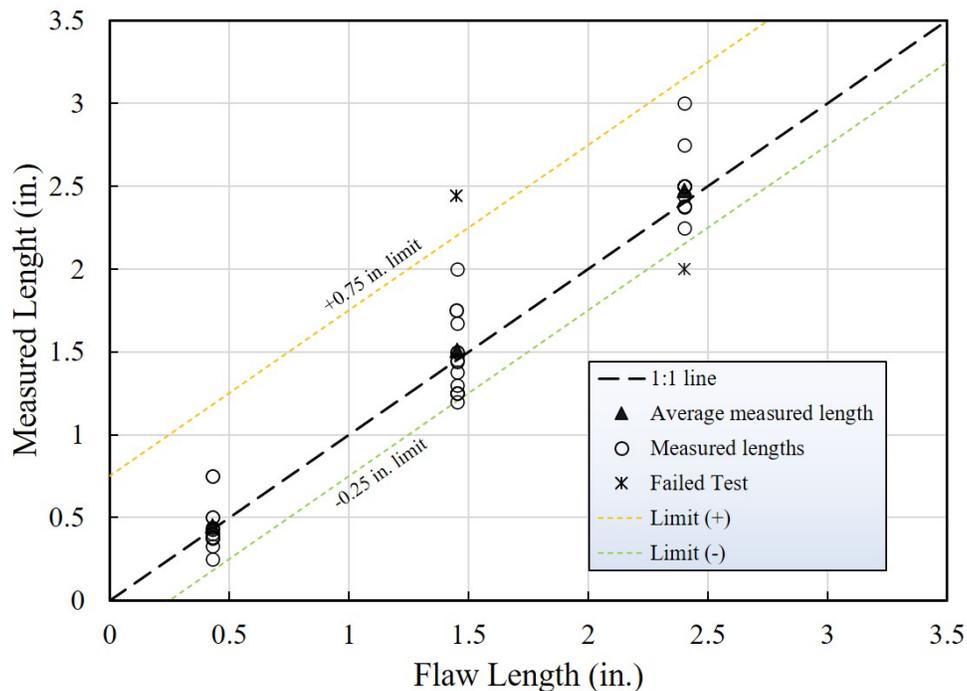
© 2022 University of Missouri.

Figure 11. Graph. Measured flaw lengths versus actual flaw length for MT performance test.

The maximum standard deviation of length measurement error was 0.23 inch. In other words, for the inspectors successful in completing the performance test, the maximum standard deviation of length measurement error was less than 0.25 inch. The table also shows the coefficient of variation (COV), which is the standard deviation divided by the mean value. The COV illustrates the dispersion of the results from the mean value, which decreases as the crack length increases, as shown in table 4.

Table 4. Statistics for MT crack length measurement results for the successful performance tests.

Actual Length (inch)	**0.43**	**1.45**	**2.40**
Average measured length (inch)	0.43	1.46	2.49
Average error (inch)	0.00	0.01	0.09
Standard deviation (inch)	0.12	0.23	0.17
COV (percent)	29	16	7

Results From UT Performance Testing

Seven inspectors completed UT performance tests. Six inspectors were successful in completing the performance test. One inspector was not successful due to significantly underestimating the length of a long flaw and reporting four false calls. The unsuccessful inspector reported a flaw length of 0.875 inch for the actual flaw length of 2.4 inches. The inspector did not request a retest.

The team conducted an analysis of the results of the performance tests to assess the capability of the inspectors to accurately determine the length of a flaw. Figure 12 presents the results of reported flaw lengths from the performance testing. A 1:1 line illustrates where measured data exactly matched the known length of the flaw. The solid triangle shows the average length measurements, and open circles show the individual measurements from inspectors. As these data illustrate, measured length exceeds the actual length in all cases for the smallest flaw. Most of the reported crack lengths for the longer flaws were also longer than the actual length of the flaw. Significant flaw scatter occurred in the results from different inspectors for the same flaw, as illustrated in the figure.
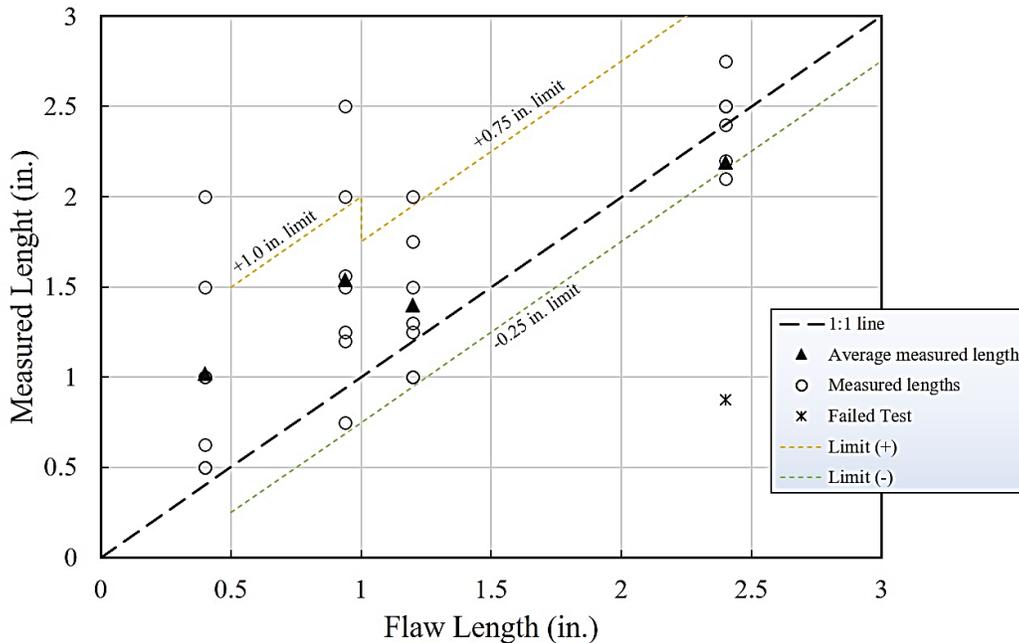
© 2022 University of Missouri.

Figure 12. Graph. Flaw length measurement data from the UT performance tests.

To assess the variation in the results from inspectors, the team analyzed the data from the performance tests to calculate the error and the scatter in results as illustrated by the COV. Table 5 shows the results of the analysis, which included the results only from successful performance tests. These data demonstrate that the accuracy of the length measurement varies as a function of the length of the flaw; for very small flaws, the lengths are greatly overestimated, whereas for larger flaws, the estimated lengths are much closer to the actual lengths of the flaws. For example, the average error for the smallest flaw, a 0.40-inch flaw, was 0.68 inches, or about 170 percent of the actual flaw length. For the largest flaw, a 2.4-inch flaw, the average error was only 0.01 inch, or about 0.4 percent of the actual flaw length. The absolute error for the 2.4-inch flaw was 0.18 inch, or about 7.5 percent of the flaw length. Of note is that the COV decreased as the flaw size increased. Overall, the accuracy of the measurement increased, and the scatter in the data decreased as the flaw size increased.

Table 5. Statistics from UT length measurements of four flaws in test plates.

Actual Length (inch)	0.40	0.94	1.2	2.4
Average measured length (inch)	1.08	1.53	1.43	2.41
Average error (inch)	0.68	0.59	0.23	0.01
Standard deviation (inch)	0.58	0.63	0.40	0.23
Average absolute error (inch)	0.68	0.66	0.36	0.18
COV (percent)	54	41	28	10

Inspectors estimated the flaw lengths using the −6 dB method described in the AASHTO/(AWS) D1.5M/D1.5 procedures, which consists of determining the position where the amplitude of the indication is 50 percent below the decibel rating for the applicable flaw classification.[9] The next step was to determine the center-point location of the transducer at the beginning and end of the indication with the flaw length characterized as the distance between these two center points. When the flaw length is smaller than the transducer, the beam profile and beam spreading effects result in an overestimate of the flaw length. The results from the performance test illustrated this effect on the reported flaw lengths, which were much longer than the actual flaw length when the flaw length was less than 1 inch.

Amplitude Variations From UT

The amplitude of indications detected using UT is the key criteria used to determine if the indication is rejectable or not. Consequently, variations in amplitude between different inspectors leads to different interpretations or resolution of an indication. Although the specific reported amplitude of an indication was not part of the criteria for the performance test, the team used the amplitudes reported on the UT inspection reports to characterize the variation between different inspectors examining the same flaw. Figure 13 shows the individual measurements reported by inspectors during the performance test for the four different flaws ranging in size from 0.40 inch to 2.4 inches. The figure also shows the average of the reported amplitudes and the standard deviation. As the figure shows, large scatter was in the reported amplitude measurements used to classify the flaw and determine if it is rejectable. It should be noted that the amplitude data are plotted as a function of flaw length for convenience but are not necessarily dependent on the flaw length.
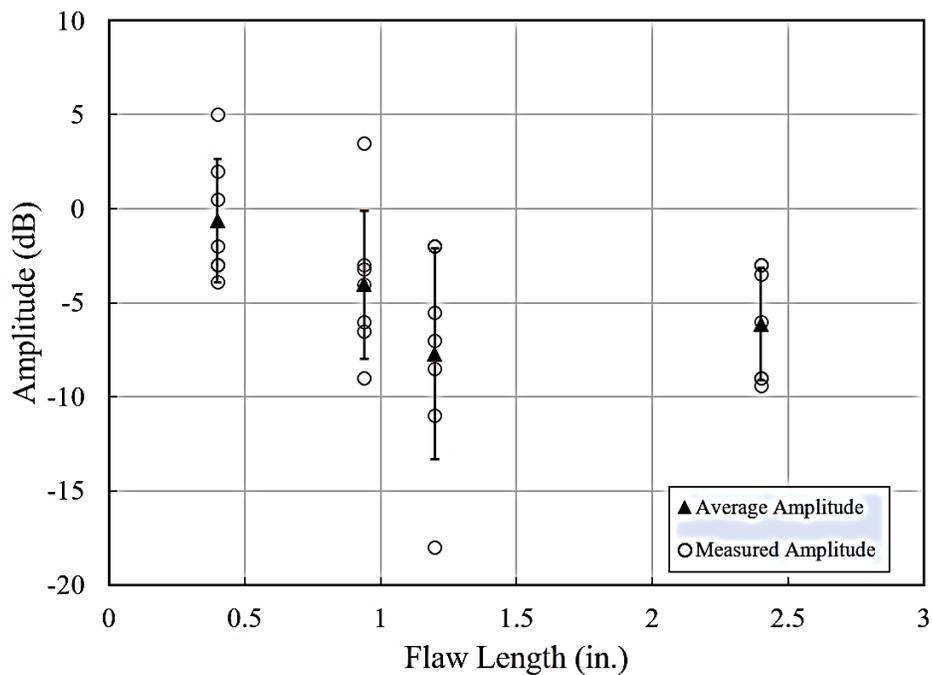
© 2022 University of Missouri.

Figure 13. Graph. Reported amplitudes for flaws in test plates.

As shown in figure 13, in some cases, the measurement amplitudes varied by as much as 16 dB for the same flaw. Table 6 shows statistics for the amplitude measurements reported from the performance testing shown in figure 13. The data show that COV decreases significantly as the flaw length increases, with the shortest flaw (0.4 inch) having a COV of more than 300 percent. For the longest flaw (2.4 inches), the COV is still significant at more than 50 percent. Given the relatively small differences between rejectable and acceptable flaw amplitude levels (figure 8), these data indicate that the classification of flaws as rejectable or not rejectable will vary significantly among different inspectors examining the same flaw.

Table 6. Statistics from UT amplitude measurements of four flaws in test plates.

Actual Length (inch)	0.40	0.94	1.2	2.4
Average amplitude (dB)	−1.07	−3.2	−8.67	−6.15
Standard deviation (dB)	3.34	3.59	5.47	3.28
COV percent	313	112	63	53

In Situ Testing

During the inspections of the S-M Bridge, the inspectors removed several cores from the bridge to assess the weld quality, perform forensic analysis of flaws detected through NDE, and characterize the material in terms of grain size and metallurgy. The data from 21 cores were collected from available reports, including the laboratory analysis of the cores.[17,18] The team removed two small cores to examine the detected cracks (cores no. 2 and no. 3), but the location of these cores was not available to correlate with NDT results from the bridge. Core no. 1 came from a location where four stop holes had been drilled previously to arrest existing cracks. As a result, the NDT results for this location are somewhat ambiguous because these holes may have interfered with applying NDT or analyzing results. For the 18 other cores, the team removed 9 of the cores at or near the shelf plate attachment that traversed the vertical butt weld. For these cores, the shelf plate may have affected access to the surface for UT and MT inspections. The analysis results described in the following paragraph reflect this possibility.

Table 7 summarizes the results from field NDT inspections and the core results. The core results for five cores signaled that no indications were found in the core, even though at least one NDT technology reported an indication at the core location. These cores are not shown in the table. Nine of the cores indicated surface-breaking cracks. The inspectors found subsurface flaws in 5 cores out of 21 cores. Among these cores, core no. 5 had one major subsurface hydrogen-induced flaw (at/after fabrication time). The other cores had minor weld flaws, according to the metallurgical analysis of the cores.[18]

The analysis of MT field inspections showed that out of nine cores containing surface flaws, MT identified only three cores with surface flaws. The remaining six cores were accepted by MT; however, the location of these six cores were either close to or behind the shelf plates. Therefore, MT may possibly not have had access to detect surface-breaking flaws at these core locations. Despite the shelf plates being located close to or at the cores no. 1, no. 8, and no. 11 locations, MT detected the surface flaws of these cores. Note that core no. 1 is not shown in the table due to the presence of four stop holes in the core which prevented meaningful UT and RT results.

The data shown in table 7 include the results of the metallurgical analysis showing the length and the depth of the flaw. The table shows the surface-breaking flaws in the top portion of the table and the subsurface flaws in the bottom section of the table. For subsurface flaws, the depths represent the depth of the top and bottom of the flaw. The table also shows the results for the NDT technologies of UT and RT. If the NDT technology showed an indication at the same position vertically in the weld as the core location, the numbers marked in bold. If the NDT technology reported an indication near the core, but not actually in the core (according to the report vertical location of the indication), values are in italics. The team determined the *overlap* based on the vertical location of the flaw in the core as compared with the indication reported by NDT. Consequently, if an indication was near the core but not in the core, the overlap is 0 inch. In other words, the flaw characterized in the metallurgical analysis was not at the location of the reported NDT indication. For these cases, the testers assumed that NDT did not detect the flaw found in the core removed from the bridge.

Table 7. Data from 14 cores removed from the S-M Bridge.

Flaw Indications (Metallurgy Lab)				UT Indications at Core Locations (inch)			RT Indications at Core Locations (inch)			Method Indicated Correctly
Core No.	Depth (inch)	Length (inch)	Depth from Surface (inch)	Depth (inch)	Length (inch)	Overlap (inch)	Length (inch)	Overlap (inch)	Type	
4[1]	0.31	1.125	NA	*0.82*	*2.5*	*0*	**0.5**	**0.5**	IF	RT
8[1]	0.11	0.68	NA	Missed by UT (above shelf plate location)			Missed by RT			MT 2/3 cracks
	0.1	0.26	NA							
	0.14	0.49	NA							
9[1]	0.16	0.63	NA	Missed by UT (partially behind shelf plate)			**1**	**0.63**	C	RT
10[1]	0.26	0.94	NA	*0.61*	*1*	*0*	Accepted Cracks b/w holes, 0.1 inch at core (acceptable)			RT
11[1]	0.38	1.44	NA	**0.74**	**2.5**	**1.44**	*0.6*	*0*	*SI*	UT, MT
12[1]	0.22	1.44	NA	Missed by UT (near shelf plate location)			Missed by RT			NONE
13[1]	0.35	1.5	NA	Missed by UT (near shelf plate location)			Missed by RT			NONE
14[1]	0.27	1.5	NA	**0.65**	**8.5**	**0.8**	**1.17**	**1.17**	C	UT and RT[3]
	0.24	1					**0.3**	**0.3**	C	
5[2]	0.97	1.2	0.23-I 0.36-E	**0.90**	**2.5**	**1.2**	**1**	**1**	C	UT and RT

Flaw Indications (Metallurgy Lab)				UT Indications at Core Locations (inch)			RT Indications at Core Locations (inch)			Method Indicated Correctly
Core No.	Depth (inch)	Length (inch)	Depth from Surface (inch)	Depth (inch)	Length (inch)	Overlap (inch)	Length (inch)	Overlap (inch)	Type	
15[2]	0.008	0.086	0.78-I 0.78-E	**0.71**	**1.1**	**0.086**	**0.35**	**0.086**	IF	UT and RT
16[2]	0.007	0.1	0.90-I 0.68-E	Missed by UT			**0.2**	**0.1**	IF	RT
19[2]	0.011	0.08	0.89-I 0.71-E	**0.6**	**1.25**	**0.08**	Missed by RT			UT
21[2]	0.002	0.01	0.28-I 1.28-E	**1.34**	**1**	**0.01**	Missed by RT			UT

[1]Surface-breaking flaws.
[2]Subsurface flaws.
[3]Four cracks reported by RT (1.17, 0.22, 0.23, 0.3).
I = from interior surface of the web plate; E = exterior surface of the web plate; NA = not applicable; NIF = no indications found; P = porosity; IR = incomplete fusion; C = crack; SI = slag inclusion.

Surface-Breaking Flaws

The engineers removed nine cores from the S-M Bridge that had surface-breaking flaws (cracks). Typically, MT, UT, and, possibly, RT detect surface-breaking flaws. Table 8 summarizes the results. In the table, an "X" indicates the NDT technology reported an indication at the core location, a "—" signifies no indication reported. The note "SP" indicates the core came from a location partially or fully behind a splice plate, a factor that could affect access for testing. As shown in the table, the location of six of these cores were partially or fully behind a shelf plate attachment. The detection of these flaws in the field was inconsistent, although the proximity of several of the cores to the shelf plate attachment may have affected the ability of the MT and UT.

Table 8. Summary of NDT results for cores with surface-breaking flaws (cracks).

Core No.	MT	UT	RT	Notes
1	X	X	—	Above SP
4	—	—	X	Above SP
8	2/3 cracks	—	—	Partially behind SP
9	—	—	X	SP
10	—	—	X	SP
11	X	X	—	Partially behind SP
12	—	—	—	SP
13	—	—	—	SP
14	—	X	X	Above SP

— = Not detected; SP = splice plate; X = detected.

Exemplar results for surface-breaking flaws can be used to illustrate NDT performance for detecting in situ surface-breaking flaws. For example, locations of cores no. 4 and no. 11 were largely above the shelf plate and appeared to have surface-breaking flaws that extended into the vertical butt weld. For core no. 4, a surface-breaking crack was present as shown in figure 14-A, figure 14-B, and 14-C. Figure 14-A shows a photograph of a 1-inch MT indication on the surface

of the core. The crack extends to some depth into the core as shown in figure 14-B. Corrosion on the crack surface (figure 14-C) indicates the crack had existed for some period before the core's removal from the bridge. Only RT detected the flaw in core no. 4; neither UT nor MT detected this flaw in situ on the bridge.

A. MT indication on the surface.
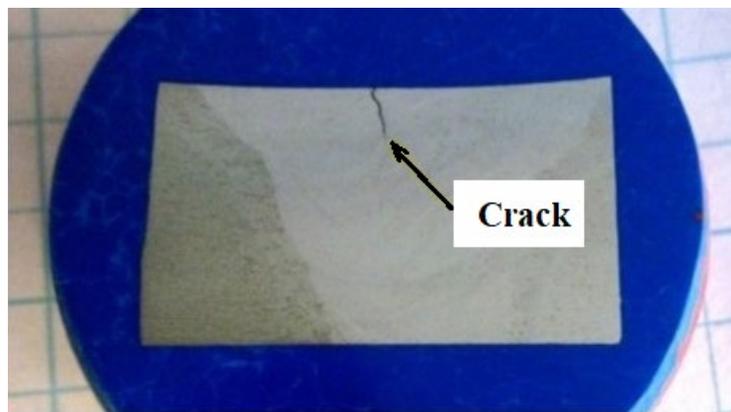
B. Polished and etched crack extending into core no. 4.

© 2012 ATS.

C. Crack surface after core was broken open at the crack.

Figure 14. Photos. Core no. 4.

In contrast, figure 15 shows a crack in core no. 11 that is about 1.44 inches long and was detected by both UT and MT. When in situ, this crack initiated in the shelf plate weld and extended into the vertical butt weld. The shelf plate weld, located transverse to the vertical butt weld, was removed prior to the core being removed from the bridge. The MT report documented a transverse crack in the shelf plate weld that extended into the vertical butt weld. The UT report showed a 2.5-inch indication in that area. The same inspector completed the MT and UT inspections on the same day. However, the RT report for that portion of the vertical butt weld did not state indications near the core. As shown in table 7, the RT report showed a small indication near the core, but the location of the indication was several inches below the location of the core so no overlap occurred between this indication and the actual crack. Consequently, because no overlap existed, the project team assumed is that RT did not detect this flaw in the field.
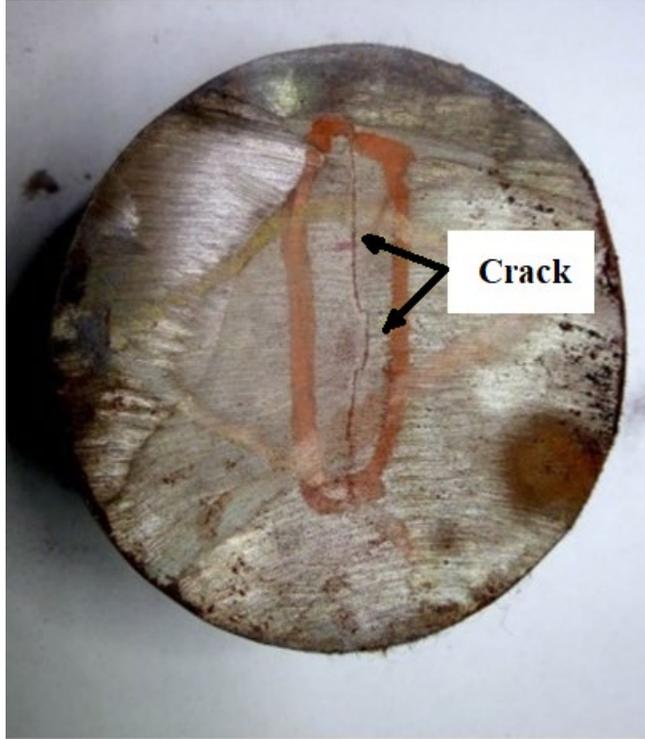
A. MT indication on the surface.
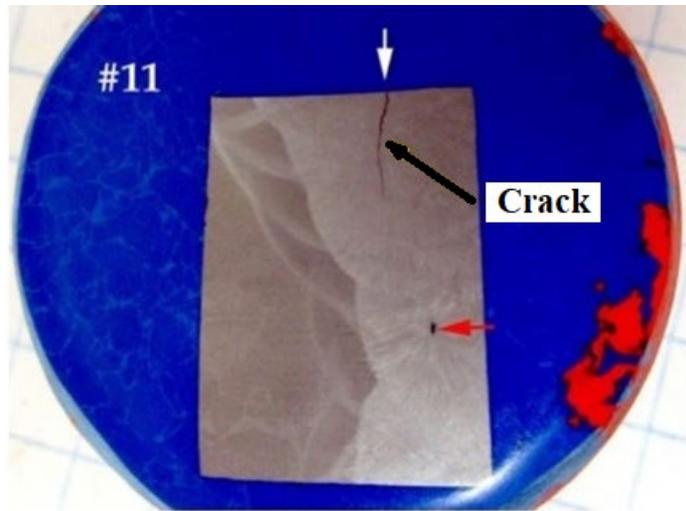
B. Polished and etched crack extending into core no. 11.
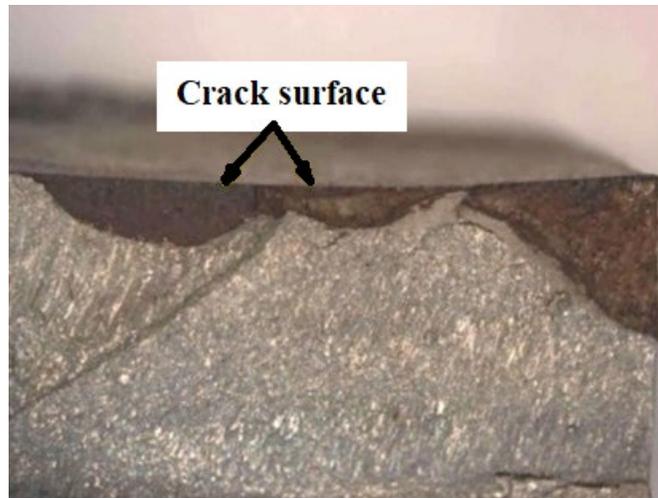
©2012 ATS.

C. Crack surface after core was broken open at the crack.

Figure 15. Photos. Core no. 11.

Subsurface Flaws

The bottom of table 7 summarizes the results for five cores containing subsurface flaws. The MT method is not capable of detecting subsurface flaws, so the inspectors did not expect to detect this type of flaw. The locations for these cores were along the vertical butt weld in areas away from the shelf plate. These cores had no geometrical features to affect the NDT technologies implemented at the location of the core. For these five cores, RT detected three of the five flaws, and UT detected four of the five flaws.

Core no. 5 included a large subsurface flaw, as shown in figure 16. This flaw was the only one of the five subsurface flaws that inspectors determined to be a crack. The inspectors determined that the other four subsurface flaws were weld flaws, such as lack of fusion.[17] Core no. 5 showed evidence of additional welding near the vertical weld, probably associated with a repair at the time of the fabrication.

The metallurgical analysis revealed that the crack in core no. 5 was about 1.2 inches long, although the engineers determined that the crack extended outside the core. Both RT and UT reported flaws at this location. RT reported a 1-inch flaw, and UT reported a 2.5-inch crack at the location of the flaw. The results presented in the final field inspection report indicated that UT had missed a major crack in core no. 5. However, the inspectors discovered that UT did detect this flaw, but the labeling of the original UT inspection form was incomplete.[17]

The subsurface flaw shows an irregular cracking pattern as compared with a fatigue-type crack, which commonly is a crack with a smooth surface. The inspectors found a crack in core no. 5, which they assessed to be a hydrogen-assisted crack with a very irregular surface, as shown in figure 15-B, which shows the crack at a magnification of 9x, and in figure 15-C at a magnification of 100x.
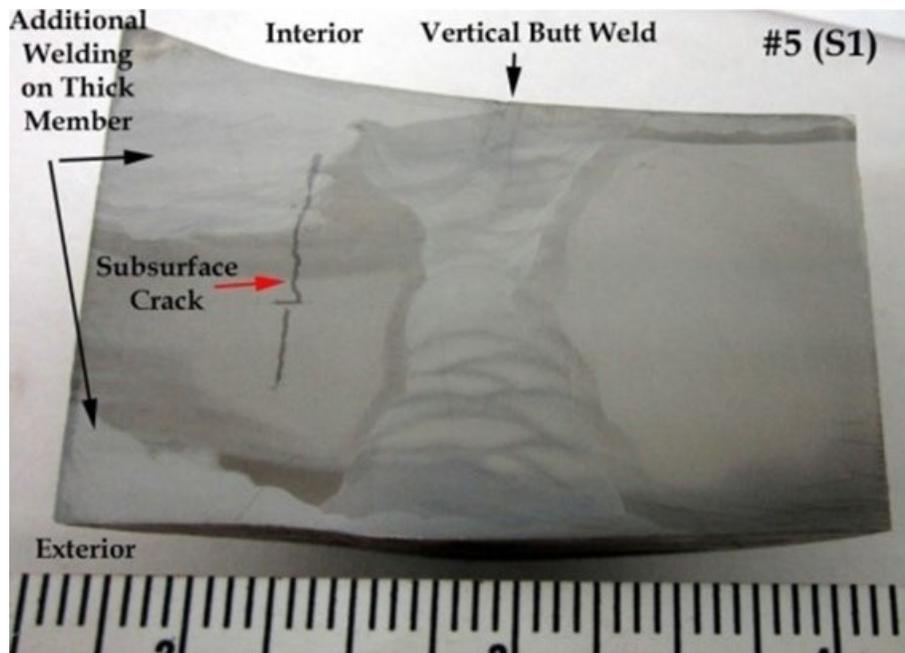
A. Polished and etched crack at magnification of 1x.
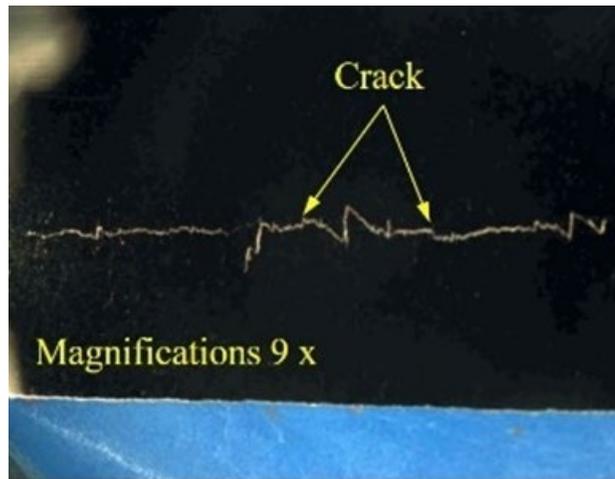
B. Crack extending at magnification of 9x.
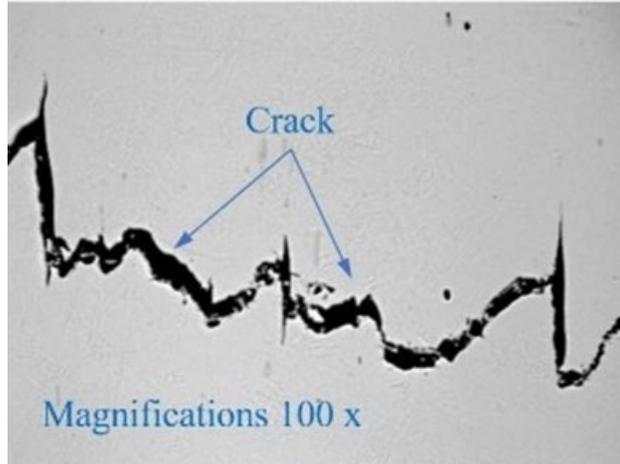
© 2012 ATS.

C. Crack at 100x magnification.

Figure 16. Photos. Core no. 5.

Figure 17 shows subsurface flaws in core no. 16 (figure 17-A) and core no. 19 (figure 17-B). Testing showed a flaw in core no. 16 as 0.1 inch in length from the metallurgical report.[18] No UT indication associated with the flaw in core no. 16 was reported, whereas the RT report indicated an incomplete fusion flaw at this location. In contrast, testing and a metallurgical analysis showed that core no. 19 had a flaw length as 0.08 inch. UT reported the length as 1.25 inches, while the RT reported no flaw at this location. For both cores no. 16 and no. 19, the metallurgical report assessed that these were "minor welding discontinuities confined to individual weld passes and not indicative of active cracking mechanisms."[18]

© 2012 ATS.

A. Core no. 6.

31
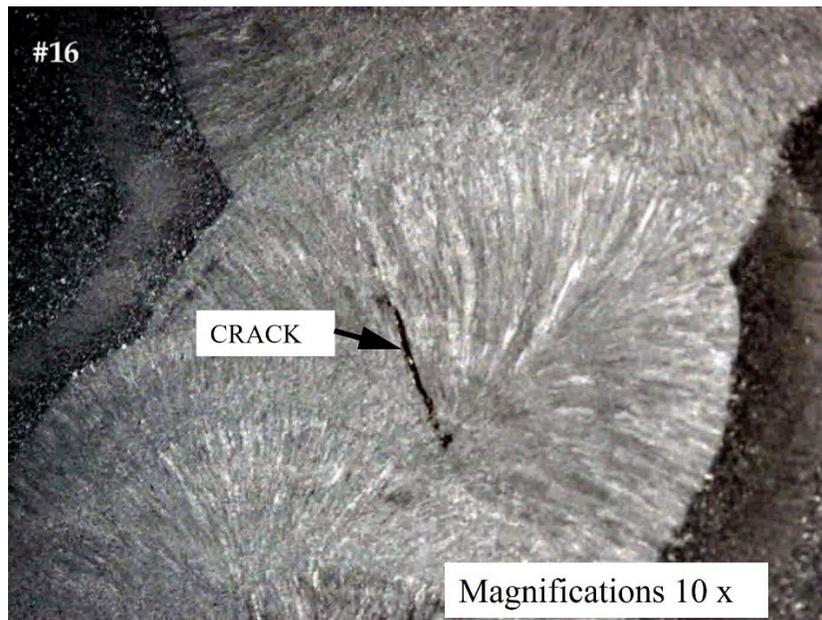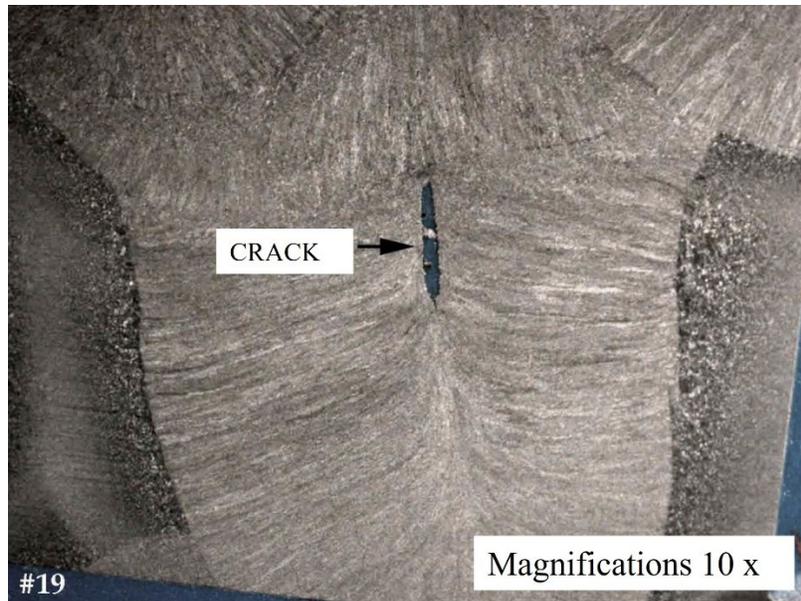
© 2012 ATS.

B. Core no. 19.

Figure 17. Photos. Flaws in cores.

Cores Without Flaws

Table 9 lists five core samples removed from the S-M Bridge that revealed no flaws, according to the metallurgical report. Four cores were at sites where RT had detected indications. UT did not detect any indications at the location of these four cores. An assessment of core no. 7 had an indication as reported by UT, but RT did not detect any indication at the core no. 7 location. The metallurgical assessment of this core was no indication found (NIF). Of note is that for the indications reported by UT, two different inspectors from two different inspection firms both reported indications. One inspector reported a class A indication at a depth of 0.8 inch at that location and did not provide its length, possibly because inspection procedures do not require reporting a length when reporting the indication as rejectable if the amplitude is sufficient to classify the indication as class A. The second inspection of the same location by a different inspector reported a rejectable indication with a depth of 0.885 inch and a length of 1.625 inches. Of note is that for the location of this core, inspectors working for different firms conducted two different UT tests but found a similar flaw at that location. However, the RT did not report a flaw at that location, and when the laboratory researchers examined the core, they noted that the metallurgical analysis found no indication.

Table 9. Data for cores that did not contain indications through metallurgical analysis.

Core No.	Method	NDT-Reported Indication	Metallurgical Assessment
6	RT	Incomplete fusion: L = 1.40 inches	NIF
7	UT	Subsurface flaw: Class A, L = 1.625 inches	NIF
17	RT	Porosity: L = 0.2 inch	NIF
18	RT	Porosity: L = 0.8 inch	NIF
20	RT	Porosity, including fusion: L = 1.1 inches	NIF

L = length.

Comparison of NDT and Core Results

The project team analyzed the core data to consider the accuracy and detection capabilities of the RT and UT technologies. They included 18 of the 21 cores that were removed from the bridge in this analysis. The team did not include cores no. 1, no. 2, and no. 3, as discussed in the In Situ Testing section of this report. The metallurgical analysis of the cores removed from the bridge included five cores that contained no indication. For the remaining 13 cores shown in table 7, a total of 16 flaws were reported in the metallurgical analysis of the cores.

The metallurgical analysis indicated that core no. 8 included three flaws. MT reported two of these flaws, but neither RT nor UT reported the flaws. The inspectors used both methods at this location, but both tests reported no indications in the area of the core. However, the analysis of detection rates considered the three flaws identified by metallurgic analysis.

Core no. 14 included two cracks. The UT results indicated one long indication to include the core location. Therefore, the team assumed that UT detected the larger flaw and did not detect the second flaw. The RT results for core no. 14 called out four separate indications. Correlation occurred for the largest indication reported by RT with the larger crack in the core sample and for the second largest RT indication with the second largest crack in the core sample. Excluded from correlation were the remaining two indications reported by RT because the RT indications may have been included in the cracking reported by metallurgical analysis of the core.

No indication was found during the metallurgical analysis for five cores. Of these five cores, four had indications reported by RT, and one had an indication reported by UT. The analysis of the core data reported here focused on the detection rates of UT and RT; consequently, the analysis excluded the cores without flaws. Therefore, the analysis included 16 flaws confirmed through metallurgical analysis.

Detection Rates

The project team analyzed the detection rates for RT and UT to assess the performance of NDT technologies in detecting flaws in situ in the S-M Bridge. This analysis considered only those cores with flaws, excluding five cores without flaws. To determine the true positive rate (TPR), they used the number of indications reported by the NDT technology divided by the total number flaws considered in the analysis. The calculation used to identify the false negative rate (FNR) is the complement of the TPR, or the number of flaws missed divided by the total number of flaws under consideration.

To assess the impact of the shelf plates on the detection rates, the project team included the following assumptions. The analysis included the five cores that contained subsurface flaws because none of the cores' locations were in the vicinity of shelf plates. For the remaining 11 cores, three were neglected (cores no. 1, no. 2, and no. 3) as mentioned previously in the In Situ Testing section of this report. The project team performed a subjective assessment based on the NDT results and photographs of core removal sites. Based on this assessment, the team did not include the results for cores no. 10, no. 12, and no. 13 in the analysis of detection rates because of the potential for interference from the shelf plate that may have affected the effectiveness of the NDT technology. The remaining 10 cores (5 with subsurface flaws and 5 with surface-breaking flaws) contained 13 flaws confirmed through metallurgical analysis.

The resulting analysis shown in table 10 indicates the TPR and the FPR for each technology. The TPR for UT was 46 percent (6/12), and the RT detection rate was 54 percent (7/12). The project team also found that 77 percent of the flaws were detected by at least one of the two technologies, but 23 percent (3/12) of flaws were not detected by either RT or UT, and 23 precent (3/12) of the flaws were detected by both RT and UT, as shown in table 10.

Table 10. Detection rates for core samples, including all applicable flaws (top) and considering the interference of shelf plates (bottom).

NDT Method	TPR (percent)	FNR (percent)
UT	46	54
RT	54	46
Either UT or RT	77	23
Both UT and RT	23	23

Accuracy

The project team analyzed the NDT data to determine the accuracy of UT and RT for measuring the length of an in situ flaw. The analysis considered only true positive outcomes, resulting in seven measurements for RT and six measurements for UT. Table 11, which presents the data from the analysis, shows the core number for each flaw considered, the measured flaw length from the metallurgical analysis, and the results from NDT. The error shown is the difference between the actual flaw length and the length reported by the NDT technology, either RT or UT. If the NDT indication was reported as shorter than the actual crack length, the error value is presented as a negative value. The absolute error is the absolute value of the error. Notable data from the analysis include that UT consistently overestimated the length of flaws, such that the average error and the absolute error were the same value. RT flaw lengths were considerably more accurate than UT lengths, and, in some cases, overestimated very small flaws lengths and underestimated large flaw lengths. The standard deviation of the error for RT was 0.42 inch, while the standard deviation of the UT measurements was 2.41 inches; therefore, UT length scatter was significantly greater than the RT length measurements.

Table 11. Actual flaw lengths, flaw lengths from RT data, flaw lengths from UT data, and error analysis.

Core No.	Core Data (inch)	RT Data (inch)	Error (inch)	Abs. Error (inch)	Core No. (inch)	Core Data (inch)	UT Data (inch)	Error (inch)	Abs. Error (inch)
Core 15[1]	0.09	0.35	0.26	0.26	Core 21[2]	0.01	1.00	0.99	0.99
Core 16[1]	0.10	0.20	0.10	0.10	Core 19[2]	0.08	1.25	1.17	1.17
Core 9[1]	0.63	1.00	0.37	0.37	Core 15[2]	0.09	1.10	1.01	1.01
Core 14[1]	1.00	0.30	−0.70	0.70	Core 17[2]	1.20	2.50	1.30	1.30
Core 4[1]	1.13	0.50	−0.63	0.63	Core 11[2]	1.44	2.50	1.06	1.06
Core 5[1]	1.20	1.00	−0.20	0.20	Core 14[2]	1.50	8.50	7.00	7.00
Core 14[1]	1.50	1.17	−0.33	0.33	—	—	—	—	—
Avg.[1]	**0.81**	**0.65**	**−0.16**	**0.37**	**Avg.[2]**	**0.72**	**2.81**	**2.09**	**2.09**
Std. Dev.[1]	**0.55**	**0.40**	**0.42**	**0.22**	**Std. Dev.[2]**	**0.73**	**2.87**	**2.41**	**2.41**

—No data.
[1]Data from RT.
[2]Data from UT.
Abs = absolute; Avg. = average; Std. Dev. = standard deviation.

Figure 18 shows the results from the analysis. This figure plots the results of indications reported by UT and RT. If UT or RT did not report an indication near a flaw, the symbol appears on the horizontal axis. The dashed line in the figure indicates the actual crack lengths according to the metallurgical analysis. This figure graphically illustrates that UT results overestimate the length of the flaw in all cases, whereas the RT results are mixed. Generally, RT reported indication sizes larger than the actual size when flaws were small and less than the actual size when flaws were large.
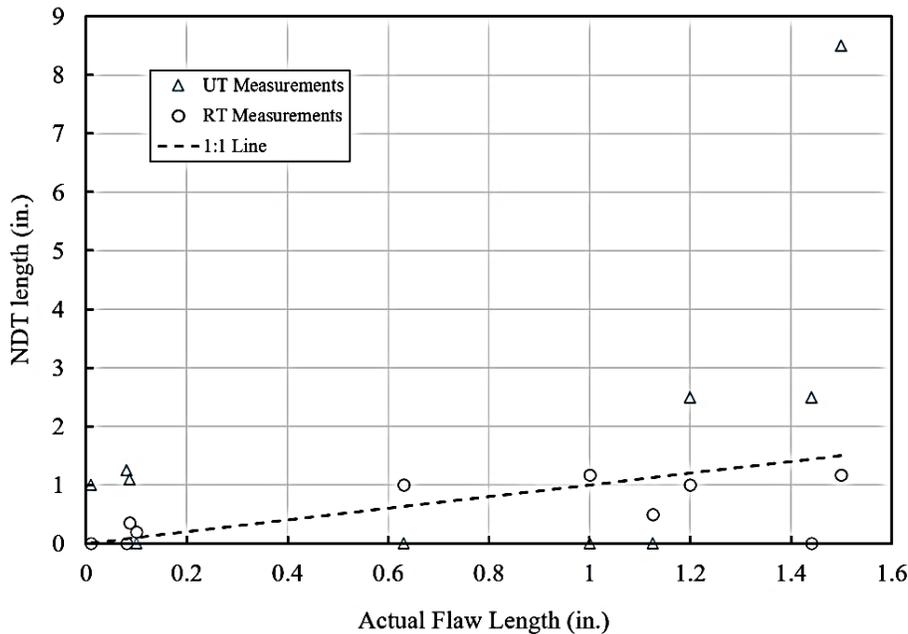
Figure 18. Graph. Length measurements reported by RT and UT as compared with actual flaw lengths.

Overall NDT Results

The S-M Bridge included 352 vertical butt weld locations that required NDT inspection. Several different firms conducted NDT on the bridge with MT, RT, and UT inspection technologies. The prime contractor produced a summary of the results from the field inspections on a spreadsheet. These spreadsheet files contained the results from the RT and UT testing on the bridge. Researchers at the University of Missouri analyzed data from these records and reduced them to a single spreadsheet consisting of data formulated in hit or miss format to analyze the probability of detection for RT and UT.[17]

The data sheets indicated that NDT technologies recorded a total of 802 indications over the course of testing the S-M Bridge. For most of these indications, no verification was available for the NDT result. Consequently, analyzing the reliability of these data in terms of the detection of flaws in the bridge was not possible. However, these data are useful in illustrating the difference in the outcome of the inspections using RT as compared with UT. The researchers' analysis determined the percentage of 802 indications detected by each technology as follows:

- UT—79 percent of indications (634/802) in the overall field inspection.
- RT—33 percent of indications (265/802) in the overall field inspection.
- UT and RT—12 percent of indications (96/802) in the overall field inspection.

These data illustrate the differences between the RT and UT technology in terms of the outcomes of the inspection, with UT returning a significantly larger number of indications in the field inspection as compared to RT. These results suggest that, generally, UT was more sensitive than

RT for reporting indications in welds in the S-M Bridge. Of note is that most of the 802 indications in the S-M Bridge were not verified in the field, except for the results associated with the 21 cores previously described in the section titled In Situ Testing.

DISCUSSION

The results reported herein describe some of the limitations of applying NDT in the field for the purpose of detecting flaws in welds. The performance testing conducted on the S-M Bridge indicated that UT commonly reported flaws to be significantly longer than the actual flaw length, especially flaws less than 1 inch in length. The analysis of cores removed from the bridge also showed that the length of flaws determined by UT was greater than the actual flaw length.

The amplitude of indications reported by inspectors during performance testing varied significantly between different inspectors assessing the same flaws. When compared with the acceptance criteria, the variation in the reported amplitudes was sufficient to possibly result in one inspector reporting a class A rejectable indication while another inspector may not report the indication at all because it is below the threshold amplitude for reporting.

The results comparing the performance of RT with UT with metallurgical analysis indicated that detection rate for either technology was only about 50 percent (46 percent for UT and 54 percent for RT). When combined, the detection rate was 77 percent. These data indicate that using the two different technologies would improve the overall detection rate for flaws. These data also suggest that RT and MT are not detecting the same flaws in many cases.

Considering the overall results of NDT on the S-M Bridge, the review of the data suggested UT was much more likely to report a flaw as compared with RT. The project team also found little agreement between the two technologies, with only about 12 percent of 802 indications being consistent between the two technologies. The analysis of the core results was similar, with only 23 percent of the flaws being detected by both RT and UT.

CONCLUSIONS

The analysis of NDT data from the S-M Bridge indicated significant variation in the NDT results. The data suggest that inspectors and engineers should take care to ensure the quality of NDT results in the field and consider using performance testing similar to that described in this report to assure that inspectors meet at least the minimum qualification standards. Using only fully qualified inspectors can significantly reduce the variability of NDT results. Other QC procedures may be beneficial. For example, using different NDT technicians to reinspect some portions of the welds examined in the field as a QC procedure may help identify errors and reduce the number of missed flaws, thereby improving the quality of inspection results. Overall, based on the experiences described in this report, the team concluded that careful QC procedures are required when using NDT for the inspection of welds in the field.

REFERENCES

1. AASHTO. 1977. "Standard Specification for Structural Steel for Bridges, AASHTO M-270." *Standard Specifications for Highway Materials and Methods of Sampling and Testing, Part I Specifications*. Washington, DC: American Association of State Highway and Transportation Officials.

2. AASHTO. 1974. "Standard Specification for High-Yield-Strength, Quenched and Tempered Alloy Steel Plate, Suitable for Welding, AASHTO M 244." *Standard Specifications for Highway Materials and Methods of Sampling and Testing, Part I Specifications*. Washington, DC: American Association of State Highway and Transportation Officials.

3. ASTM. 1964. "Standard Specification for High-Yield-Strength, Quenched and Tempered Alloy Steel Plate, Suitable for Welding, ASTM A 514." *Annual Book of ASTM Standards, Part 4*. Philadelphia, PA: ASTM.

4. ASTM. 1964. "Standard Specification for Pressure Vessel Plates, Alloy Steel, High-Strength, Quenched and Tempered, ASTM A 517." *Annual Book of ASTM Standards, Part 4*. Philadelphia, PA: ASTM.

5. FHWA. 2011. *Inspection of Fracture Critical Bridges Fabricated from AASHTO M270 Grade 100 (ASTM A514/517) Steel*. Technical Advisory 5140.32. Washington, DC: Federal Highway Administration.

6. Kalla, H. 2021. Hari Kalla to Division Administrators, Directors of Field Services, memorandum, "Non-Destructive Testing of Fracture Critical Members Fabricated from AASHTO M244 Grade 100 (ASTM A514/A517) Steel," December 13, 2021. Washington, DC: Federal Highway Administration. https://www.fhwa.dot.gov/bridge/NDT_FCM_Bridges_Fabricated_T1_Steel.pdf, last accessed December 19, 2023.

7. FHWA. 2023. "National Bridge Inventory (NBI)" (web page). https://www.fhwa.dot.gov/bridge/nbi.cfm, last accessed July 21, 2023.

8. ASTM. 2021. *Standard Specification for Structural Steel for Bridges*. ASTM A 709. Philadelphia, PA: ASTM. https://kupdf.net/download/astm-a-709pdf_5b33d0aae2b6f5553af86392_pdf, last accessed December 14, 2023.

9. AASHTO and AWS. 2020. *Bridge Welding Code*. AASHTO/AWS D1.5M/D1.5:2020. Miami, FL: American Welding Society.

10. ASNT. 2020. *Personnel Qualification and Certification in Nondestructive Evaluation*. Columbus, OH: ASNT.

11. Shaw Jr., R. E. 2002. "Ultrasonic Testing Procedures, Technician Skills, and Qualifications. *Journal of Materials in Civil Engineering* 14, no. 1: 62.

12. Jessup, T. J., P. J. Mudge, and J. D. Harrison. 1981. *Ultrasonic Measurement of Weld Flaw Size*. NCHRP Report 242. Washington, DC: Transportation Research Board of the National Academies of Sciences, Engineering, and Medicine.

13. Connor, R. J., C. J. Schroeder, B. M. Crowley, G. A. Washer, and P. W. Fish. 2019. *Acceptance Criteria of Complete Joint Penetration Steel Bridge Welds Evaluated Using Enhanced Ultrasonic Methods*. NCHRP Research Report 908. Washington, DC: National Academies of Sciences, Engineering, and Medicine.

14. Gruber, G. J., and G. M. Light. 2002. "Supplemental Ultrasonic Code Inspection of Structural Weldments." *Journal of Materials in Civil Engineering* 14, no. 1: 57.

15. AWS. 2020. "Annex F (Normative)—Supplemental Ultrasonic Technician Qualification," in *Structural Welding Code—Seismic*, 4th ed. Miami, FL: AWS.

16. API. 2015. *Recommended Practice for Ultrasonic and Magnetic Examination of Offshore Structural Fabrication and Guidelines for Qualification of Technicians*, API-RP-2X, 4th ed. R15. Washington, DC: API Publishing.

17. Hotaki, H. "Comparison of Nondestructive Testing Technologies Used for Inspection of Welded Joints." Master of Science thesis. Columbia, MO: University of Missouri.

18. ATS. 2012. *Metallurgical Evaluation of Flaw Indications in Tie Chod Transition Butt Weld Cores From the I-64 Sherman Minton Bridge*. Louisville, KY: Applied Technical Services, Inc., p. 207.