AUT QUANTIFICATION METHODOLOGY

METHODOLOGY 2

Methodology for Design, Fabrication and Fingerprinting of Quantification Welds

Issued by:
Edison Welding Institute

Revision: 5

Revision Date: February 18, 2011

Prepared by:

Evgueni Todorov, Ph.D.
Senior Engineer
Nondestructive Evaluation

Approved by:

Roger Spencer
Senior Engineer
Nondestructive Evaluation

February 18, 2011
Date Approved

Methodology for Design, Fabrication, and Fingerprinting of Quantification Welds

1. Scope

1.1 This methodology covers the design, fabrication, and fingerprinting of practical trial specimens used for the automated ultrasonic testing (AUT) quantification process. The designated quantification administrator will have primary responsibility for the design, fabrication, and fingerprinting described in the document. This methodology is part of the guidelines for quantification of AUT of girth welds.[1]

1.2 Specimens used for establishing detection capabilities can also be used for establishing sizing capabilities through destructive measurements or through fingerprinting using other, more accurate nondestructive techniques if destructive testing is not conducted.

2.0 Definitions

a_{50} – flaw size with 50% POD. This means that 50% of the flaws with this size and larger will be detected.

a_{90} – flaw size with 90% POD. This means that 90% of the flaws with this size and larger will be detected.

$a_{90/95}$ – flaw size with 90% POD and 95% confidence. This is the most quoted parameter in the literature. It means that 90% of the flaws with this size and larger will be detected and this is true in 95% of the inspections under similar conditions (equipment, examiners, environment, etc.).

3.0 General Design Requirements

3.1 Specimens must be representative of the actual structure and account for input parameters defined in the technical justification (TJ) such as:

- Girth weld material, geometry, and other conditions
- Flaw location and parameters
- Inspection objectives related to inspection performance and expressed through required detection and sizing capabilities.

3.2 Simplification of test specimen design and number of specimens is possible through TJ. Validated modeling tools are particularly useful in justifying the simplified design and reduced number of specimens.

3.3 Specimens shall be representative of those conditions and parameters that cannot be justified through previous experience, physical considerations, previous trials or modeling.

3.4 Quantification specimens shall be fabricated from a pipe representative of the material, heat treatment, surface finish, and acoustic properties.

3.5 Flaw types shall be representative of those expected for the welding process to be used.

3.6 Typically flaws should be evenly distributed throughout the weld, both in the through-wall direction and lateral direction.

3.7 The circumference of the girth weld shall be divided into sectors (grading units) normally in the range from 40 to 75 mm (1.6 to 3 in.) in length. The sectors shall be uniformly distributed around the girth weld circumference.

3.8 The locations of sectors with and without flaws shall be randomized around the girth weld circumference to avoid statistical dependence.

3.9 If a weld sector does not contain flaws, it shall have at least 1-in. length on either side of the sector without flaws. If the sector contains a flaw longer than the sector length (1.6 to 3 in.), it shall have at least 1-in. length without flaws on either side of the flaw.

4.0 Number of Flaws Needed

4.1 The number of flaws used in the quantification will be affected by the following:

- Essential parameters requiring justification through practical trials
- Balance between TJ and practical trials
- Statistical method to be used for quantification

4.1.1 A sufficient number of weld sectors with and without flaws shall be built to effectively estimate sizing error distributions, and determine the detection capabilities such as probability of detection (POD) and probability of false positive (PFP) for specific flaw types and sizes.

4.1.2 Two statistical options are recommended (refer to Appendix 1 for advantages and disadvantages of each):

- Option 1. Binomial Method: Specimens shall contain weld sectors with flaws of identical size, or a size that varies in very small range. Typically, the flaw size is specified to be the critical or very close to the critical size for the girth weld to be inspected.

- Option 2. Binary Regression Method: Specimens shall contain weld sectors with flaws in a range of sizes. The size range shall cover flaws that are expected to be detected, that will be missed and that will easily be detected. The flaws that are expected to be detected are usually with sizes smaller than the critical size.

4.2 Number of Flaws for Option 1 – Binomial Method.[(2)] The number of weld sectors with flaws is determined using a standard binomial detection test. The minimum number of flaws will depend on the required POD and required POD confidence as shown in Table 1. The number of flaws will strongly depend on the number of misses as well.

Table 1. Minimum Number of Weld Sectors with Flaws for a Given POD, Confidence and Number of Misses

Confidence	Number of Misses	Number of Weld Sectors with Flaws		
		POD 80%	POD 90%	POD 95%
50%	0	3	7	14
	1	7	17	32
	2	11	27	51
	3	15	37	70
	4	19	47	89
	5	23	57	108
	10	43	107	203
	20	83	207	394
90%	0	11	22	45
	1	17	38	77
	2	23	52	105
	3	29	65	132
	4	34	78	158
	5	39	91	184
	10	64	152	306
	20	112	267	538
95%	0	13	29	59
	1	21	46	93
	2	27	61	124
	3	33	76	153
	4	39	89	181
	5	45	103	208
	10	72	167	336
	20	121	286	577

4.2.1 As an example, the minimum required number of flaws for 90% POD with confidence 95% is 29 (Table 1). If the AUT operator misses 1 flaw from the set, then, a second trial is conducted where the AUT operator must detect 45 out of 46 flaws for the selected flaw size. It can be done by asking the operator to detect all of the additional 17 flaws to the set of 29 or the initial set of 29 flaws can be mixed and randomized with the additional 17 flaws to form a set of 46 flaws for the flaw size of interest to demonstrate 90% POD score at 95% confidence where 1 of the flaws can be missed.

4.2.2 If several flaw sizes or flaw size intervals are needed to be quantified, the minimum number of flaws for the POD, confidence and anticipated number of misses must be assembled in each of those intervals. For example, a minimum number or 116 flaws shall be assembled (29 in each size interval) if the quantification is for 4 flaw height intervals – 1 to 1.5 mm, 1.5 to 2 mm, 2 to 2.5 mm and 2.5 to 3 mm is required with 90% POD and 95% confidence. The number of flaws in each interval shall be increased in accordance with Table 1 if any misses are expected.

4.3 Number of Flaws for Option 2 – Binary Regression Method.[3] There are two approaches for the binary regression method: The first approach is referred to as **â vs a** (a hat versus a) and should only be used when the AUT system response (amplitude) **â** is proportional to the flaw size **a**. The second approach is referred to as **hit/miss** when only two conditions of the AUT system are considered: **hit** (pass) – where a flaw with size **a** was detected; and **miss** (fail) – where a flaw with size **a** was missed.[3] The number of flaws needed will depend on the statistical approach used.

4.3.1 **â vs a** Approach: At least 40 flaws are required to build a reliable POD(a) curve when **â vs a** approach is implemented because additional information is available. In addition, other conditions (linearity, error normality, uncorrelated measurements, and uniform variance) shall be verified to ensure that the POD(a) estimate is valid.

4.3.2 **hit/miss** Approach: At least 60 flaws (ideally 120) are required to build a reliable POD(a) curve when the **hit/miss** approach is implemented.

4.3.3 Generally, flaw sizes will depend on the inspection objectives and limits of the AUT system. In instances where $\mathbf{a_{90}}$ is unknown, preliminary open trials with the AUT system may be required to approximately determine the range of flaw sizes necessary for the POD quantification.

4.3.4 If $\mathbf{a_{90}}$ is approximately known from other quantification studies, 80% of flaw sizes should be distributed uniformly on a Cartesian scale (as opposed to log scale) around this value. Of the remaining 20%, 10% of flaw sizes should be undetectable ($\mathbf{a} < \mathbf{a_{10}}$) and 10% of flaw sizes should be readily detectable ($\mathbf{a} > \mathbf{a_{90}}$).

4.3.5 In general, the sizing accuracy of the AUT system decreases and the POD increases with an increase in flaw height. When detection and sizing capabilities are determined with the same set of specimens, additional flaws, with a size larger than $\mathbf{a_{90}}$, may be necessary to determine the sizing accuracy for larger flaws.

EXAMPLE: If $\mathbf{a_{90}}$ = 1 mm, $\mathbf{a_{10}}$ = 0.5 mm, $\mathbf{a_{90/95}}$ = 2 mm, the following distribution of flaw sizes might be applicable for both POD and sizing accuracy determination (Table 2):

Table 2. Example Flaw Size Distribution

Flaw Size Range, mm	0-0.5	0.5-1.0	1.0-1.5	1.5-2.0	2.0-2.5	2.5-3.0	3.0-5.0
Number of Flaws	7	16	16	8	8	8	7

4.3.6 It is important to note that the POD(a) may not be accurately determined, or may not be determined at all, if only small number of flaws are missed and or flaw sizes of missed flaws do not overlap with the sizes of the detected flaws. Practical trials where all flaws are detected or missed are of little value in determining AUT detection capabilities expressed as POD(a).

4.4 All weld sectors with and without flaws shall be randomized, if possible, during the quantification to reduce the chances of statistical dependence.

6.0 Considerations for Determination of False Positives

6.1 The false positive rate (FPF), or probability of false positive (PFP), is the ratio of false positive indications to the number of weld sectors without flaws. Consequently, a percentage of weld sectors will need to contain no weld flaws in order to determine the false positive rate.

6.2 To estimate false positives with good accuracy and reliability, the ratio of "no-flaw" sites or weld sectors to the flaw weld sectors should ideally be at least 3 to 1. Since this can add considerable expense, it is acceptable in most cases to reduce this ratio; however the ratio should not be less than 1 to 1.[(3)]

7.0 Specimen Design Documentation

7.1 During design, each specimen shall be assigned a unique identification and have a zero scan location point from which all flaw locations will be referenced. In addition, one side of the weld shall be designated as the upstream (US) side and the other side as the downstream (DS) side.

7.2 Flaw locations and sizes should be documented in table format as shown in the example in Appendix 2.

7.3 It is recommended that fabrication drawings be produced for use during sample fabrication. The drawing should show weld bevel dimensions, pipe diameter, wall thickness, as well as, location and size of each intentional flaw.

8.0 Specimen Fabrication

8.1 Raw material used in preparation of specimens shall be tested before welding for flaws, attenuation and sound velocity. Quality certificates of compliance may be requested from the raw material manufacturers.

8.2 Each specimen shall be impression stamped (or other approved permanent marking) showing the specimen identification, zero location, weld sector markings, US identification, and DS identification. All permanent markings shall be in a location where they do not interfere with scanning or AUT performance.

8.3 The welding process and method used for specimen preparation shall be similar to the welding of pipes to be inspected. The administrator may review or audit the process at specimen fabrication site.

8.4 Natural or implanted (seeded) flaws simulating the ultrasonic response from actual flaws are recommended for the practical trial specimens.

8.5 It is recommended that intentional flaws be at least 12 mm in length unless it is a requirement to detect shorter flaws. Likewise, it is recommended that adjacent flaws be separated by at least 12 mm and that the use of stacked flaws is avoided.

8.6 Implantation flaws may introduce additional flaws due to implantation process. The area surrounding the implant shall be inspected to ensure no other satellite flaws are present between the implant and the original material.

8.7 A possible difference in macrostructure between the implant and original material shall not cause significant difference between the signals from the implanted flaw and the natural flaw that is simulated.

9.0 Fingerprinting and Quality Assurance

9.1 During specimen manufacture, the administrator shall:

- Provide detailed specimen specifications to the specimen fabricator
- Review the work progress regularly
- Review the compliance certificates from the specimen manufacturer when the specimens are delivered.

9.2 It is recommended that nondestructive fingerprinting of the flawed specimens be performed to document the as fabricated condition and to assure that the intended flaws were produced prior to proceeding with the quantification process. Fingerprinting may also be used where validation of flaw location and size through destructive measurements is not possible. Following are possible nondestructive methods for fingerprinting:

- X-ray radiography (film, digital, or computed tomography) primarily for volumetric flaws
- Ultrasonic method primarily for planar flaws
- Eddy current, magnetic particle, or liquid penetrant methods for surface-breaking flaws. (The eddy current method is recommended because it is possible to do the inspection with minimal surface preparation.)

9.3 If alternative nondestructive methods are to be used in lieu of destructive methods for estimation of AUT detection and sizing capabilities, the capabilities of the alternative methods shall be assessed through comprehensive destructive testing measurements. The detection capabilities shall be better and the errors in location and flaw size measurements shall be significantly smaller than the errors of the AUT technique being quantified.

9.4 Limited destructive measurements of flaw location, size, and flaw type may be required in the initial stages of specimen fabrication to validate and optimize the flaw implantation and fabrication process.

9.5 The quantification administrator must keep specifications, drawings and any other information confidential.

9.6 Specimens used for open trials shall not be used for blind trials.

9.7 Flaw locations and sizes determined after fingerprinting should be arranged in table format as shown in the example of Appendix 3.

10.0 Documentation

10.1 The quantification administrator shall maintain documents related to specimen design, fabrication, and fingerprinting results.

10.2 Flaw locations and sizes should be summarized in table format as shown in Appendices 2 and 3 for specimen fabrication and fingerprinting, respectively. Information related to specimens designed for blind trials shall be kept confidential and shall not be accessible to unauthorized personnel and parties involved in the quantification process.

10.3 The documentation will be part of the quantification file.

11.0 References

(1) Guidance for Quantification of Automated Ultrasonic Testing Systems for Examination of Pipeline Girth Welds, Edison Welding Institute (EWI).

(2) Nordtest Report NT TECHN REPORT 394, Approved 1998-04.

(3) MIL-HDBK-1823A (2009), Nondestructive Evaluation System Reliability Assessment, Department of Defense Handbook.

Appendix 1 – Advantages and Disadvantages of Statistical Options

Option 1 - Binomial Method

Advantages

- Smaller number of weld sectors with flaws are required for low POD with low confidence where misses are not expected with tested AUT process.
- Relatively fast estimates are possible for a single flaw size of interest.
- Data processing and evaluation is relatively simple.

Disadvantages

- Number of required flaws will increase significantly for higher POD with higher confidence and possible misses.
- It is difficult and expensive to produce flaws at multiple locations with identical or very narrow range of sizes. This is particularly true for surface and subsurface flaw height where destructive validation is not conducted and flaw characterization or fingerprinting of quantification flaws is conducted with alternative NDT techniques.
- Only one flaw size – the largest is quantified for a given flaw size range. For example, if 29 flaws with height from 1.6 to 2.45 mm were detected during quantification in the height range 1.5 to 2.5 mm for 90% POD with 95% confidence, the AUT process is quantified for the flaw with the largest height of 2.45 mm.
- It is difficult and expensive to investigate the AUT performance for a larger range of flaw sizes so that improved AUT performance, if demonstrated, in detecting smaller flaws leading to larger inspection intervals is taken into account when service life of critical joints is considered.
- The sizing accuracy will only be available for the narrow range of flaw sizes tested. There will be no data on how the sizing accuracy varies as a function of the flaw size.

Option 2 - Binary Regression Method

Advantages

- A smaller number of flaws might be required for higher POD with higher confidence compared to Option 1.
- The AUT performance is quantified in a range of flaw sizes by definition.
- The number of misses will automatically be reflected in the estimation of detection capabilities (e.g., $\mathbf{a_{90/95}}$) and no retesting will be necessary with larger number of flaws.
- It is easier to compare the performance of AUT vendors and operators with a minimum number of trials on the same set of specimen weld flaws.
- The AUT performance is easier to optimize as well. For example, the analysis of the AUT process for vendors that consistently and reliably detect smaller flaws provides objective criteria and direction for improvement to vendors that do not perform to the same level.

Disadvantages

- Prior knowledge of detection capabilities is necessary to produce flaws with adequate size distribution.
- The data processing will require specialized computing tools developed by outside parties for estimation of POD. Careful analysis of underlying principles and assumptions imbedded in these computational tools may be required.

Appendix 2 – Sample Table for Flaw Location and Size during Fabrication

Date:
Project No.:
Weld No.:
Specimen Fabricator:
Specimen Design Drawing or Specification:

Location and Size of Flaws for Fabrication

Sector No.	Sector Start (mm)	Flaw Type	Axial Position (US/DS)	Circumferential Position (mm)		Length (mm)	Depth (mm)	Height (mm)	Comments
				Start	Stop				
S1	0	NF							
S2	40	A3	US/DS	60	60	0	10	2	Transverse crack
S3	80	NF							
S4	120	NF							
S5	160	A5	US	170	190	20	15	1	Crack
S6	200	NF							
…	…	…	…	…	…	…	…	…	…

NF – No flaw
US/DS – Upstream (US) or Downstream (DS)
A3 and A5 – Codes of flaws (other codes might be used)

Appendix 3 – Sample Table for Specimen NDE Fingerprinting after Fabrication

Date:
Project No.:
Weld No.:
Specimen Fabricator:
Specimen Design Drawing or Specification:
NDE Method Used for Fingerprinting (each method requires separate table)

___ Visual, ___ X-Ray Radiography, _X_ AUT, ___ Eddy Current, ___ Magnetic Particle, ___ Liquid Penetrant.

Description of NDE Equipment:
NDE Procedure:
NDE Operator:
Location of Fingerprinting:

Location and Size of Flaws after NDE Fingerprinting

Sector No.	Flaw Type	Axial Position (US/DS)	Circumferential Position (mm)		Length (mm)	Depth (mm)	Height (mm)	Comments
			Start	Stop				
S1	NF							
S2	Transverse Crack	US/DS	63	63	N/A	13	3.5	
S3	NF							
S4	NF							
S5	Crack	US	174	202	28	10.2	1.6	
…	…	…	…	…	…	…	…	…

NF – No flaw
US/DS – Upstream (US) or Downstream (DS)