

CIVA Modelling Module for Zonal Discrimination Method Part 2-System Qualification Aspects

Ed Ginzel¹

¹ University of Waterloo, Waterloo, Ontario, Canada
e-mail: eginzel@mri.on.ca

Dave Stewart²

²Simply AUT, Edmonton, Alberta, Canada
e-mail: david.stewart@simplyaut.ca

2024.02.10

Abstract

Automatic Ultrasonic Testing (AUT) systems for pipeline girth weld inspections are usually subjected to a qualification process. This process consists of assessment of performance capabilities including detection, sizing, temperature range and repeatability. One of the most demanding qualification programmes is described in DNV ST-F101 Appendix E [1]. With the introduction of the AUT module in the 2023 edition of CIVA simulation software, the incorporation of metamodeling allows users to predict detection reliability via probability of detection (POD) and sizing accuracy of the zonal discrimination method (ZDM) using amplitude apportioning. Part 1 of this project demonstrated the ability of CIVA to simulate the calibration block and strip chart output. In Part 2 we illustrate the techniques that can be used to predict POD and sizing accuracy as might be done in a system qualification. Results compare well to field data.

Keywords: CIVA, phased-array, ultrasonic, Zonal Discrimination Method

1. Introduction

Some years ago, the authors demonstrated [2] the ability of CIVA to match strip chart responses of the ZDM for both calibration targets and flaws. This was not particularly surprising considering the many validations, especially the QNDE Benchmark programme, that CIVA has participated in [3]. Some have suggested that CIVA is limited in its ability to simulate 3D shaped defects and small sized defects. Clearly this was disproved in the 2011 paper for the defect types typically involved in such projects [2]. Validating CIVA results for the ZDM against field results was demonstrated in that paper and again more recently in Part 1 [4] of this project where the calibration block results used were actually part of a recent field project.

In this paper we compare the CIVA AUT module results of sizing and POD to those obtained by a qualified AUT system using the same calibration setup as used for a field validation. It is demonstrated how the use of metamodeling provides a large-scale sample on which to draw data for POD and Sizing accuracy studies. These results can then be compared to system qualification results to confirm if the system performance remains intact. Because of the ability of the metamodel to examine multiple variables at the same time, predictions can be made about the range of parameter variability that a validation could be extended to. This could be a significant benefit when multiple sizes of pipe are being considered for systems' validation beyond the parameters that they were initially qualified for.



2. Qualification Parameters

An AUT system qualification in accordance with DNV ST-F101 uses performance assessments based on:

1. Detection capabilities
2. Sizing capabilities
3. Temperature capabilities
4. Repeatability tests

Of these assessments, temperature and repeatability tests would be mechanically reliant performance tests so are not candidates for computer simulations. However, detection and sizing capabilities can be assessed using simulations as these results will be determined by the configuration of the ultrasonic components and software processing of data.

Upon successful qualification a document is issued (by DNV) that tabulates the range of parameters for which the system is qualified. These parameters include;

- Weld bevel geometry
- Bevel angle
- System setup (pulse-echo and pitch-catch)
- Welding method (e.g. SMAW, GMAW, etc.)
- Material (carbon steel, corrosion resistant alloy, clad, etc.)
- Thickness
- Pipe diameter
- Calibration reflectors (e.g. notches for surface detections and flat bottom holes for subsurface)

Aspects of a qualification that can make use of simulations include;

1. Preparation of the calibration block
2. Optimisation of delay laws to ensure adequate zone spacing and over-trace
3. Positioning of targets for volumetric channels
4. Observe the effect of added gain for scanning
5. Predict a suitable evaluation threshold commensurate with the proposed fracture-mechanics based acceptance criteria
6. Predict a likely POD that the configuration can achieve
7. Assess the algorithm used to size vertical extent of flaws
8. Predict the flaw size that the system makes to ensure that under-sizing error tolerances give less than or equal to 5% probability of under sizing when used in relation to any ECA specified defect sizes
9. Predict the range of parameter variables that would influence the qualification of the system and require validation tests to extend these parameter ranges

When preparing validation welds, companies must try to fabricate flaws near the critical flaw size as determined for the acceptance criteria. Ideally, they would use flaws around the qualified POD sizes for the various thresholds; however, in reality the flaws are seldom the desired size. With modelling, the validation flaws could be adjusted for the sizes required.

Part 1 of this programme investigated the first 4 aspects of the above list. The other aspects of simulation will be examined in this report.

3. CIVA AUT Sensitivity Algorithm

After more than 30 years of experience using AUT ZDM, the generally accepted targets on which to set sensitivity for AUT systems are 1mm deep surface notches and 3mm diameter flat bottom holes aligned with the theoretical weld bevel. These targets have been used by many service companies and thousands of macro images have been generated to evidence the detections and sizing accuracies achieved. Flaw sizing and POD are affected by the threshold used to identify which indications are evaluated. Typically, analysis of indications is carried out when the signal response exceeds 20% full screen height (FSH) when the reference amplitude from the calibration target is set to 80% FSH (i.e. 25% of reference). In some instances, as per service provider procedures or client specifications, extra gain is added to some channels to increase detection capabilities. With the exception of the volumetric channels, this added gain for scanning is not applicable to the POD assessment at the time of qualification, since the POD can easily be adjusted by varying the evaluation threshold relative to the reference amplitude. Varying the evaluation threshold is equivalent to adjusting the gain when assessing POD.

In the process of setting up the calibration on the standard targets, the CIVA AUT module sets each zonal target to a reference (0dB) independent of the other zonal channels. These reference values are then used by the CIVA POD module computations in assessing the responses from flaws to calculate POD and to estimate flaw size.

Perhaps the greatest limitation of the present practice for qualifications in the DNV process is the limited number of samples required to assess POD and Sizing. The number of deliberate imperfections required for an AUT qualification is between 91 to 120, depending on the weld bevel shape and material (carbon steel or corrosion resistant alloys).

With over 20 years of qualifications from numerous AUT systems and service companies, there are now thousands of flaws documented that provide assurance that the reference targets used can provide the sensitivity required to assure the required reliability of inspections. However, the process described for reliability (i.e. POD calculation) of these qualifications is perhaps a departure from normal practice as it applied to NDT reliability assessments by POD. A POD is typically applied to a particular inspection and flaw type. For example, this would imply that pulse-echo testing is treated differently than tandem testing in UT, or surface-breaking flaws that can be better detected by way of corner effects are treated differently than embedded flaws.

Yet in the DNV qualification process for AUT systems, the approach lumps all of the inspection approaches and flaw types together to obtain an overall POD. As an example, for a J-bevel weld in carbon steel, the numbers of flaws are as follows;

Imperfection	Number of imperfections
Outside Surface	29
Hot Pass	29
Buried	29
Inside Surface	29
Volumetric	2
Inter-run lack of fusion	2

It is noted that the volumetric flaws are not to be included in the POD calculations so there are 118 imperfections identified for the overall assessment of POD.

POD assessment of AUT system reliability calculates the minimum flaw size that can be detected at 90% probability with a 95% confidence. If the normal requirements of 90|95 were to be used with only the 29 flaws in any one imperfection grouping being used, no more than 1 flaw could be missed. The assembly of a POD experiment makes the process even more difficult. Flaws that are used to assess an inspection system's reliability are supposed to be scattered in the range of interest as determined by the critical flaw size determined for fracture mechanics-based acceptance criteria. This would then require some of the deliberate flaws to be smaller and some flaws to be greater than this critical size. Missing more than 1 flaw in 29 would make it impossible to achieve the 95% confidence level. Yet having some of the flaws missed is a requirement for a realistic POD curve. If all of the flaws are "detected" then 100% POD exists and no "minimum flaw size" can be calculated from the data.

CIVA AUT module does not permit the determination of an "overall POD" but rather conforms to the MIL HDBK-1823A [5] recommendations. CIVA calculates separate PODs for each flaw group. Cap, Root, Hot Pass and Fill zones are calculated separately with option to group all the Fill zones in one POD computation.

Guidance on the number of flawed locations is provided in MIL-HDBK 1823A. They suggest that in order to provide reasonable precision in a POD, the specimen test-set should contain at least 60 sites if the system provides only *hit/miss* analysis and at least 40 sites if the system is to provide a POD based on \hat{a} vs a . They further recommend that these numbers are minimums. MIL-HDBK 1823A further notes that for *hit/miss* responses, 120 sites will result in a significantly more precise estimate of a_{50} , and thus a smaller value for $a_{90/95}$.

CIVA AUT module makes use of the advanced parametric study capability of Meta-modelling. Meta-modelling allows analysis of the effects of multiple parametric variations. Having established the parameters that will be used for a girth weld inspection project (probe, calibration block material and targets), the next step in the CIVA simulation is to run a "Sensitivity" study. The "Sensitivity" module carries out a parametric study for each channel. Using the CIVA canonical rectangular flaw, variations are run on the specimen and flaw variables in a range of user-defined uncertainties.

Flaw variables include:

- Height
- Length
- Tilt
- Depth from the outside surface

Test variables include;

- Material velocity
- Pipe thickness
- Probe offset.

For the purposes of AUT applications, the flaw height would be considered the characteristic value and a linear set of variations assigned over a range. Typically, 500 or more samples are then made using a randomising of the other parameters in the range selected. This results in a very large database from which to extract samples. For example, if 500 samples were run for each Fill Zone in a 4 Fill calibration, this would provide 2000 samples from which to build the metamodel. The POD construction can then be made by interpolating values from the resulting metafile curves. This makes the potential variations for the POD infinite and also provides the potential to test different scenarios with varying sampling without the need for more simulations.

The sensitivity module allows the user to rank the relative effect of each variable by computing Sobol indices for all essential variables, and ranking the impact of essential variables to the amplitude of each channel. Figure 2 illustrates the relative impact of the selected variables on the Fill 2 zone of the 4-Fill calibration used in this project. It is obvious from this display that the main influencing variables are Flaw Height, Flaw Depth and Flaw Tilt. Thickness variation was assumed to lie in a narrow range, typical of rolled plate (+/-0.5mm) and has little effect on the sensitivity. Length of interest was assumed to be flaws greater than 6mm in length. Since the probe used is focussed in the passive plane, flaw length is seen by the Sobol indices to have no impact on sensitivity. Probe position relates to the accuracy of the band-placement to maintain a fixed distance to the weld centreline. Within the allowed +/-0.5mm, this parameter has no significant impact on sensitivity. The last parameter in the display is acoustic velocity. It was given a tolerance of +/-30m/s and the effect on sensitivity is similar to, but less than, the tilt of the flaw. Since variation in acoustic velocity affects the refracted angle, it could be considered similar to what might occur as a result of temperature variation.

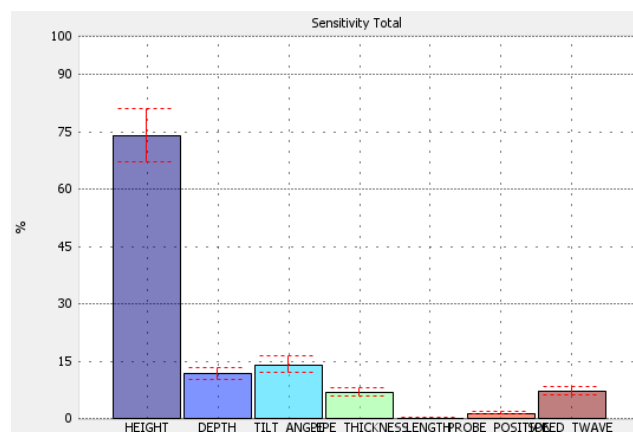


Figure 1 Sobol indices for Fill 2 parameters of the 15.88mm J-Bevel inspection

4. CIVIA AUT POD computation

In part 1 [4] of this project the calibration block and delay laws provided nearly identical results to the validation calibration scan obtained in the field. Compared to the field settings, the simulation overtrace values were slightly less and some of the delay laws had the start elements different by just a single element in order to centre the beams on the targets. The calibration used in part 1 provided the foundation for the sensitivity metamodel in part 2.

Results of the meta-model are then used to derive the probability of detection (POD). As noted, separate PODs are calculated for Cap, Root, Hot Pass and Fill regions. As well, in the Fill region it is possible to calculate PODs for just a single Fill or one can select multiple Fill zones.

The images in Figure 2 are the POD plots and associated size-versus-amplitude response plots for the 4 regions of the J-Bevel weld. Sensitivity values were made using 800 LHS (Latin Hypercube Sampling) shots for each Zone. The plots were constructed using 10 height values with 50 samples for each height. This sampling could be further refined at no extra computation costs because it relies on a metamodel built from the simulations. Data was converted to a percentage applied to the A-scans whereby 100% Reference was the response from the calibration reference target. This would be equivalent to 80% FSH. For the purpose of the POD calculations the “detection” threshold is set to 25% of the reference which is equivalent to 20% FSH when the reference target is set to 80% FSH.

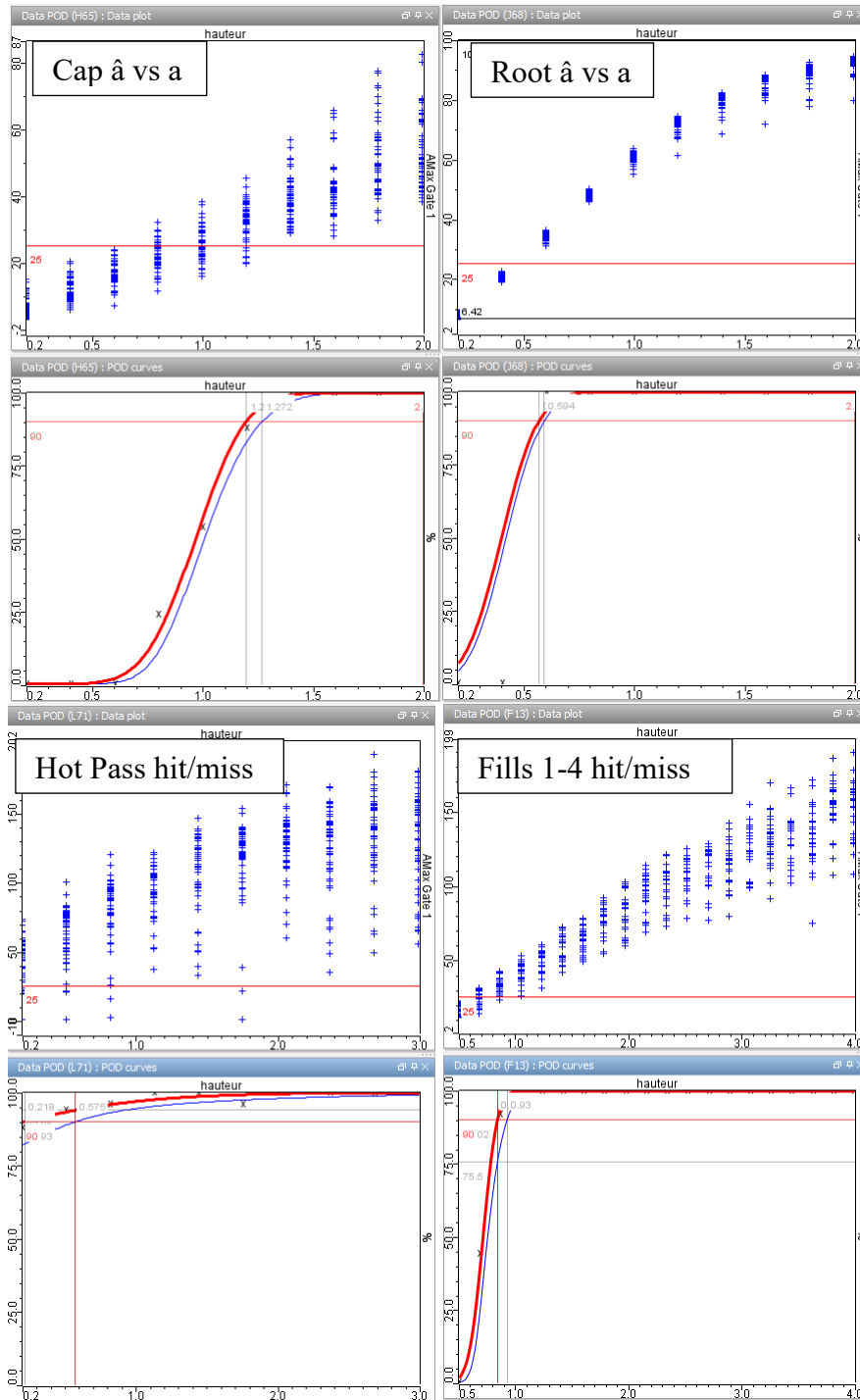


Figure 2 POD curves and amplitude vs size plots for 4 regions of the J-Bevel

The Root and Cap regions show a distinct trend to a proportionality between flaw size and amplitude response so it was elected to compute the POD using the “ \hat{a} vs a ” method.

Variance of the data for the Hot Pass and Fills exhibited more scedasticity so it was considered more appropriate to use the hit/miss approach on the data in these zones.

From the curves in Figure 2 the values for the a50 and a95 were extracted.

Zone(s)	a50 (mm)	a95 (mm)
Cap	1.2	1.27
Root	0.57	0.59
Hot Pass	0.22	0.58
Fills	0.84	0.93

The rather sensitive detection capability in the Hot Pass zone can be attributed to the use of 1.5mm diameter FBHs instead of the standard 3mm diameter FBHs. This change was as a result of a client requirement.

It is interesting to note how the uncertainty of the POD increases (moves to the right) when the sample size is reduced. In Figure 3 the 500-samples used to derive the POD for the 4 Fill zones is seen on the right and to the left in Figure 3 the same 4 Fill zones using only 48 samples produces a higher value for both the a50 and a95. When configuring the parameters for the POD on the left side of Figure 3, 8 bins were distributed equally over the depth range of the four Fill zones. This would be equivalent to taking 12 flaw samples from each Fill zone. The critical a95 value moves from 1.16mm to 2.75mm when a lower sampling is used. If the 29 samples required for the Fill zones by the qualification process were used separately instead of as part of an overall POD, the uncertainty that would result could be significantly increased.

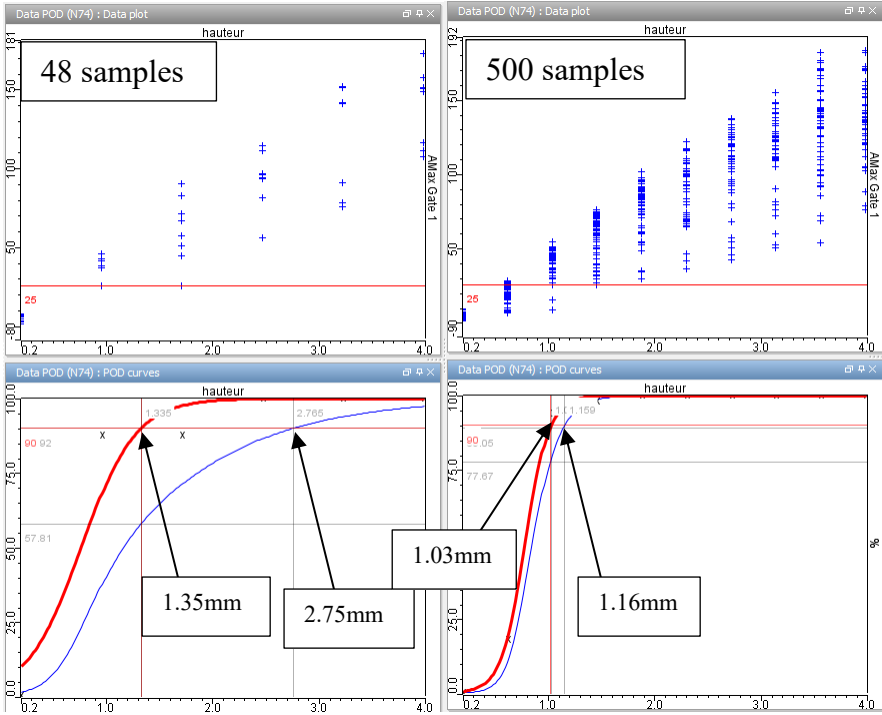


Figure 3 Comparing the effect of sample size on POD curves

5. CIVIAUT Sizing Algorithm

From its earliest applications, ZDM was used in conjunction with fracture-mechanics acceptance criteria. Service companies using ZDM were given tables or graphs showing allowed lengths for estimated flaw heights. Sizing estimates were based on apportioning amplitudes in each zone from the flaw responses seen on the strip charts.

Figure 4 illustrates a common approach whereby the flaw amplitudes from the weld on the strip chart are compared with the amplitude of the responses seen from the calibration. Calibration overtrace, i.e. the presence of a response on an adjacent channel from the FBH target, is required by the inspection standards to verify that there is no significant loss of sensitivity between zones. Overtrace is simply the off-axis pressure from the beam detecting a reflector. If a flaw is seen on two or more zones, the zone with the maximum amplitude is considered the main zone and at least some of the amplitude in the adjacent zone(s) will be attributable to the off-axis beam pressure. Amplitude apportioning reduces the amplitude of the response in the zone adjacent to the main response by the overtrace seen on the calibration. The remaining amplitude is considered to be from a reflecting area greater than the FBH in the calibration block.

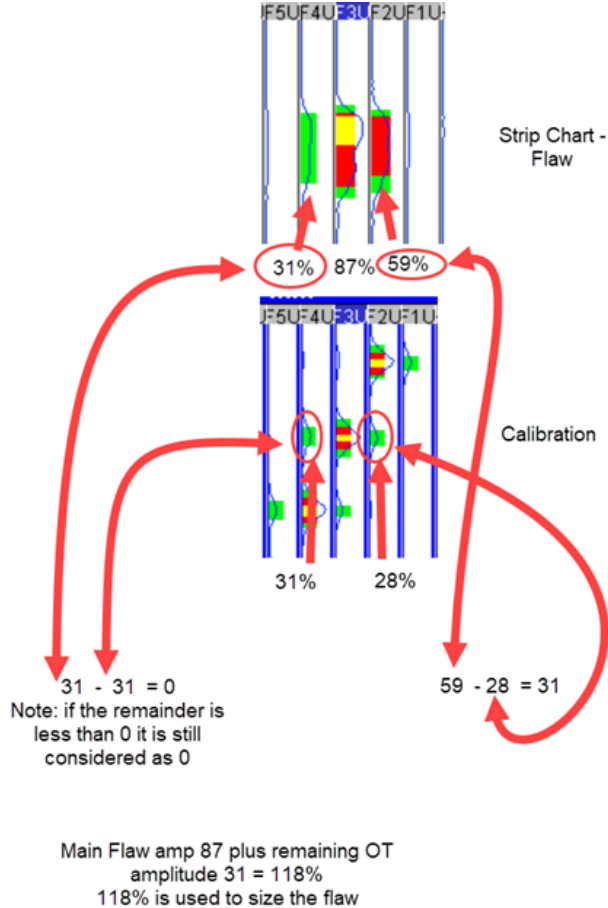


Figure 4 Principles of sizing by amplitude apportioning

Meta-model data is also used as part of the sizing component of the CIVA AUT module. The known vertical size parameter is corrected for the projected height using the Cosine function and becomes the source of known flaw sizes. Amplitude responses computed for each condition of parametric variation are extracted from the meta-model. The default sizing algorithm in CIVA applies a logic sequence similar to that described in Figure 4. The algorithm identifies the channel with maximum amplitude and examines the amplitude responses in adjacent channels. It uses the calibration responses from the main and adjacent zones to apportion the amplitude contributions to calculate flaw height. Results using the default options in the CIVA AUT Sizing module were found to be reasonably close to those obtained in the field. There is an option to use the CIVA Script module to modify the sizing algorithm if so desired.

As seen in the POD amplitude versus size plots (Figure 2), the ZDM process is surprisingly linear. This provides an opportunity to indicate the sizing trend and derive the safety factor against under sizing.

Amplitude versus size plots for the root and cap zones rely on a corner effect. Since flaws of interest are usually much longer than the beam is wide, the reflecting area is proportional to the flaw height. And since the signal amplitude is proportional to flaw area, it follows that signal amplitude will be proportional to flaw height. At only 1.25mm, the root and cap zones are approximately the same size as the beam. When the flaw height starts to exceed 1.5mm, the amplitude versus height in the Cap zone has a bit more scatter and in the Root zone the linear proportionality tends to drop off. However, in both zones, the concept of a linear relationship between amplitude and flaw height is easily made.

In the Fill zones, flaws tend to move along the bevel and may cross from one Fill zone to the next without the convenient separation that is provided by the abrupt angle change that exists at the Hot Pass and Root zones. The effectiveness of the amplitude apportioning for sizing in this region can be seen in a plot of the sizing for just Fill zone flaws seen in Figure 5.

In Figure 5 two sizing plots extracted from the same meta-model are compared. To the left 100 samples are used and to the right 200 samples are used. The range of flaw heights considered is from 0.1mm to 4.5mm. The CIVA display indicates the perfect sizing (red diagonal line) and computes the quantile of under and oversizing. CIVA does not compute a mean and standard deviation around the mean that would be based on a normal distribution. A blue line is drawn parallel to the red line where 5% of the samples are less than the intercept to the X-axis. Where 100 samples are used, on the left 5 blue dots are seen below that blue line. This is a measure of the 5-percentile under sizing. For the 100 sample plot the intercept is at -1.82mm. For the same variables, using 400 samples improves the under sizing so it is reduced to -1.58mm.

1.58mm under sizing is actually not very good for sizing accuracy of an AUT system. A critical factor in all statistical assessments is the range of variables used. In the plots used in Figure 5 the range of variables was much greater than might be used in the field. A machined bevel can be controlled to within 1° so instead of using $\pm 3^\circ$ tolerance for the angle we can reduce that to $\pm 1.5^\circ$. And it is unlikely that we would detect flaws less than 0.5mm for evaluation so the lower limit for the tolerance can be increased from 0.1mm to 0.5mm. Similarly, we can elect to reduce the upper flaw height limit to 4mm.

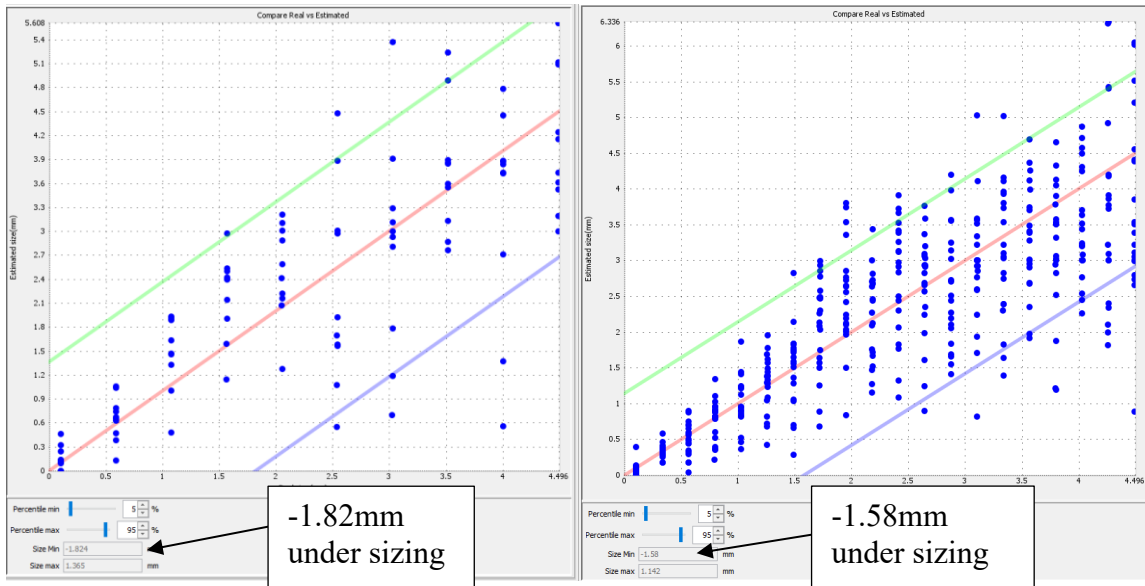


Figure 5 Comparing Fill Zone sizing using 100 flaws and 400 flaws

When the range of analysis is more likened to the field conditions, the plot of sizing accuracy more closely represents results that have been achieved in AUT system qualifications.

Figure 6 illustrates 400 samples with tilt tolerance reduced to +/- 1.5° and the range of flaw height from 0.5mm to 4.0mm. The resulting 5 percentile under sizing is indicated as 0.78mm, which is the same overall value achieved by the AUT system used in this project.

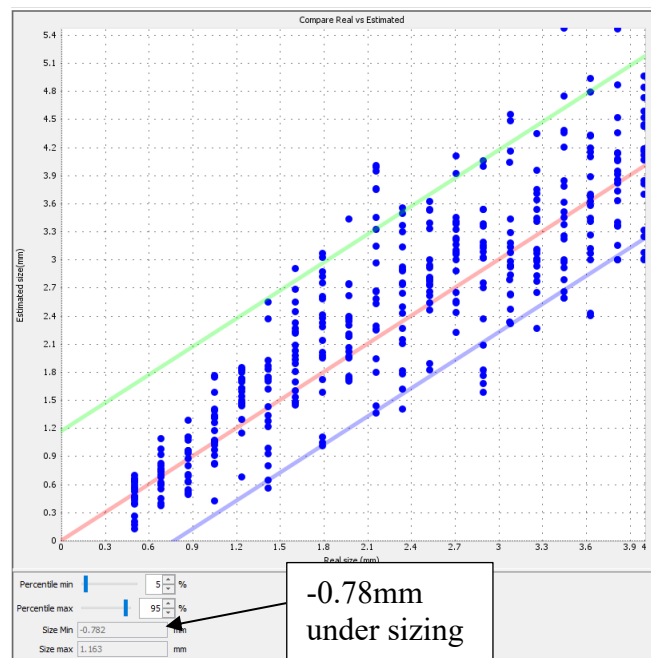


Figure 6 Fill Zone sizing using 400 flaws at field tolerances

Ranking the under sizing of an AUT system using the 5-percentile undersized provided by CIVA works well if the AUT system has a mean error of zero. However, it is possible that there is a systematic over sizing or under sizing. Systematic error is identified by the mean error being more or less than zero.

To address the systematic error that might be present in a system, DNV derived an equation to provide a Safety Factor against under sizing.

$$SF = \chi - (1.64 * \sigma)$$

Where: χ is the mean sizing error

σ is the standard deviation of error

1.64 is the one-sided z -score for 95% of the population being to the left in a normal distribution

CIVA can be used to calculate this safety factor because the data behind the plot in Figure 6 is provided in tabular form. Copying the columns of values for the Real and Estimated flaw sizes and putting them in a spreadsheet provides:

Average			0.17
StdDev			0.57
Safety Factor against under sizing			-0.76

In this case, the safety factor is little changed from the 5-percentile indicated in the CIVA plot because there is a small systematic over sizing of 0.17mm.

6. Probability of Rejection (POR)

In 1997 Forli [6] wrote that the idea that a flaw had a probability of being rejected could be equated to its probability of detection. This was rationalised by the fact that any flaw that was to be rejected first required that it be of sufficient amplitude to be evaluated (detected). In the 2007 Edition of the DNV OS F101, the concept of Probability of Rejection (POR) added.

POR is, in principle, similar to the reliability evaluation of POD. Unlike POD, POR uses the 85% POR at 95% confidence level. The 85% POR accounts for both a 90% POD and a 95% probability of avoiding under-sizing. The DNV code requirement states that the flaw height at 85% POR at 95% confidence level, shall be equal to or below the smallest “allowable” flaw height in the acceptance criteria. POR will not necessarily say anything about the smallest flaw that is possible to detect with the system at a certain set-up. The difference between POD and POR is the threshold applied for hit and miss. For POR the threshold is set for the AUT reported flaw height rather than amplitude.

Using 400 data points from Figure 6 and the ability to import data into the CIVA POD module, CIVA can be used to establish the AUT system POR using the sizing accuracy data. Note, the POR can be just as easily determined using field data imported to the CIVA POD module.

Using the \hat{a} vs a method, the data derived from the table used to generate Figure 6 calculates the 85% POR at 1mm threshold to be 1.4mm with the 95% confidence at 1.5mm.

Selecting the method of analysis (i.e. hit/miss or \hat{a} vs a) is based on the type of data available.

When the data indicates that there is some proportionality with the size of the flaw, the preferred process is to apply \hat{a} vs a in the derivation of the curve. The curve in Figure 6 was made using the \hat{a} vs a option in the CIVA POD module.

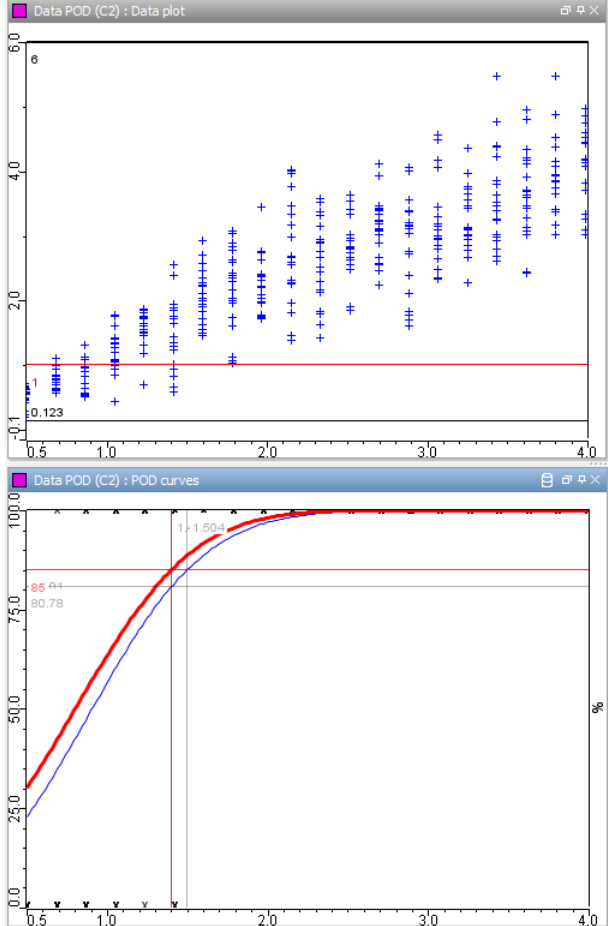


Figure 7 Fill Zone POR sizing using 400 flaws at 1mm threshold

No sizing data is provided in the CIVA AUT module for the cap and root regions; however, when the linear plots of amplitude versus size in Figure 2 are considered, it is evident that the sizing in these small regions will be as good or better than the sizing by apportioned amplitudes in the multi-Fill zone region.

6. Conclusions

The purpose of this paper was to examine how simulation of AUT processes for POD and Sizing might compare with field results. This involved designing a calibration block and delay laws that matched those used in a recent field validation.

The purpose of an AUT “validation” is to simply assure the client that the reliability of the system, as established in the “qualification”, is maintained. The number of samples collected in a validation is inadequate to carry out the statistical analyses involved in POD and Sizing Accuracy tests. This makes the validation process a qualitative assessment since judgement must be used to conclude that the performance reliability established in the qualification can be repeated on the project.

Simulation affords the ability to demonstrate, in a quantitative way, if the AUT system setup for the project is similar to the qualified system.

Since the goal of the validation is to indicate “similarity” of reliability to the “qualified” system, we can compare the main parameters from the qualified system to those found using simulations. The main items of concern for comparison are POD and Sizing Accuracy (height). Comparing values of the qualified system to the values derived using the simulation data, based on the validation calibration block and delay laws, indicates a close similarity.

Parameter	Qualified ¹	Simulated ²
³ POD at 20% threshold	0.7	0.93
POD at 30% threshold	1.0	1.19
POD at 40% threshold	2.2	1.28
Sizing (under sizing)	0.7	0.76

¹ average of all zones
² fill zones only
³ DNV calculations use binning of hit/miss data CIVA does not bin data

With the exception of the 90|95 flaw size for POD at 40% threshold, the values for reliability as determined by POD are similar. Caveats are noted. The qualified system used only 120 observations covering the full range of zones whereas the simulated results use only data derived from the Fill Zones using 500 samples. Although both the qualified and simulated data was processed using a hit/miss algorithm, the qualified data used a process whereby the 120 samples were put into 5 bins (i.e. 24 samples were grouped into each bin based on flaw size with bins ranging from 0.4mm to 2.2mm). CIVA does not bin data when calculating POD.

Although POR was not an item qualified on this system, the CIVA software demonstrated it was easily calculated if required. Using the Fill Zone sizing data and a 1mm height threshold, the POR at 85|95 was found to be to be 1.5mm.

Simulation has provided evidence that the validation data is consistent with historical results. Being able to use several hundred samples significantly improves the confidence curve values derived for these computations.

Generally, it is considered that upon a simulated design of a suitable calibration setup, as demonstrated by the responses on the strip chart, the CIVA AUT simulation module can then be used to predict sizing accuracy and reliability of the procedure setup based on probability of detection.

Strip chart output is reserved for the calibration scan and not available for embedded flaws. However, the standard “Inspection” capabilities are available on a per-zone basis that can be assembled using the echo-dynamic curves to assess a similar display.

TOFD is often included in AUT systems to aid in characterisation of flaws, discriminate between flaws and false signals in ZDM zonal channels and as an aid to sizing when tip diffraction signals are available. TOFD is not yet integrated in the calibration aspect of the AUT module but it is an available feature for the inspection component in the AUT module. It should be noted that TOFD is not used in the POD analysis of ZDM qualifications.

Acknowledgements

We would like to thank Stéphane Leberre, Fabrice Foucher and Mariana Burrowes for their support in this work.

References

1. DNV ST-F101, “Submarine Pipeline Systems, Appendix E, Automated Ultrasonic Girth weld testing, 2021
2. Ginzel, E., Stewart, D., CIVA Modelling for Pipeline Zonal Discrimination, https://www.ndt.net/article/ndtnet/2011/1_Ginzel.pdf, 2011
3. CIVA validation papers <https://www.extende.com/publications>
4. Ginzel, E., Stewart, D., CIVA Modelling Module for Zonal Discrimination Method Part 1-Calibration Block, https://www.ndt.net/article/ndtnet/papers/CIVA_ZDM_Calibration_Block.pdf, 2023
5. MIL-HDBK 1823A, Department of Defense Handbook Nondestructive Evaluation System Reliability Assessment, 2009
6. Forli, O., <http://www.ndt.net/article/0498/forli/forli.htm>, How to develop acceptance criteria for pipeline girth weld defects, Det Norske Veritas, Oslo (N) European American Workshop Determination of Reliability and Validation Methods of NDE, Berlin - June 18-20, 1997