Assessing reliability of NDE flaw detection using smaller number of demonstration data points

Ajay M. Koshti, D. Sc. PE
NASA Johnson Space Center, Houston

March 2019

SPIE Smart Structures/NDE 2019
Denver, CO, U. S. A.

Outline

- Acknowledgements
- References
- Abstract/Introduction
- Overview
- Simulation approach
- Approach for assessing technique reliability in x-ray radiography
- Conclusions

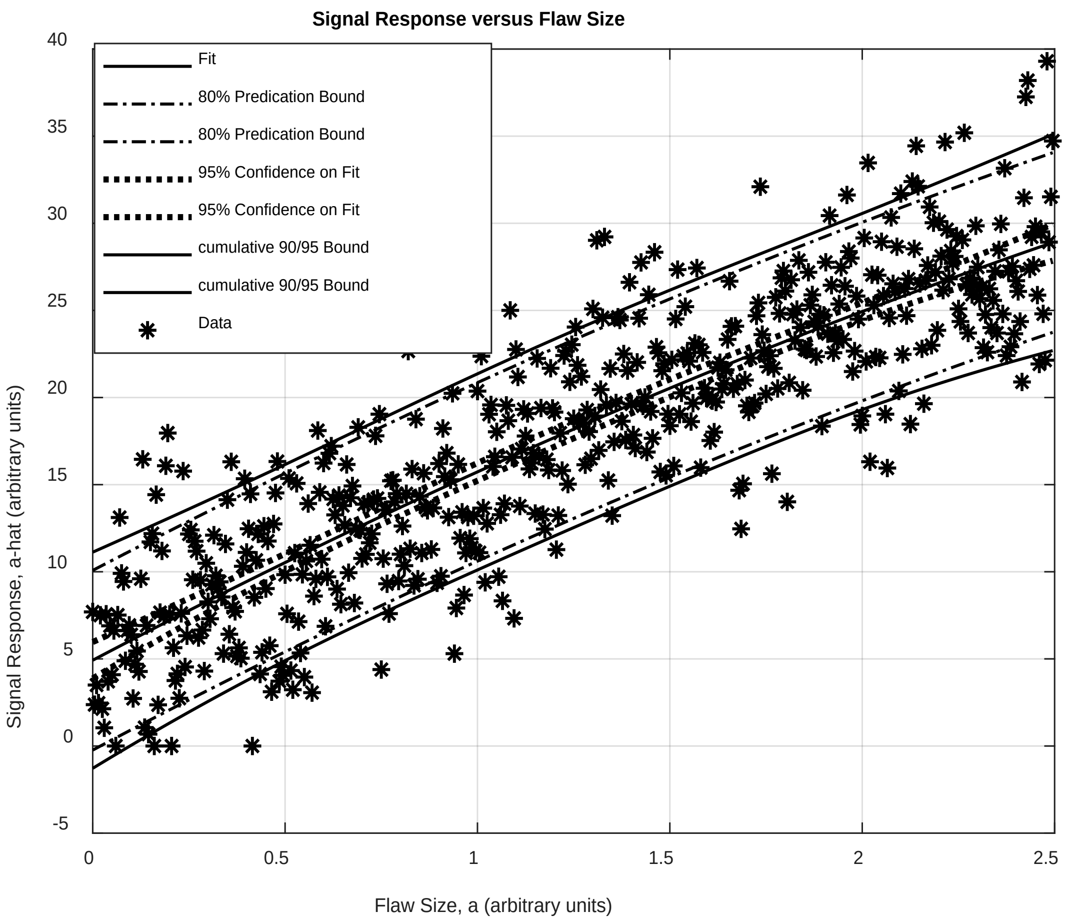


Fig. 1: Signal response versus flaw size.

Acknowledgements

- William Prosser (NASA Engineering Safety Center, NASA Langley Research Center)
- Floyd Spencer (NASA Engineering Safety Center)
- David Stanley (NASA Johnson Space Center),
- Ronald Beshears(NASA Marshall Space Flight Center), and
- James Walker (NASA Marshall Space Flight Center)

References

[1] MIL-HDBK-1823A, "Nondestructive Evaluation System Reliability Assessment," Department of Defense, USA, (2009).

[2] Annis, Charles, "mh1823 POD Software V5.2.1," Statistical Engineering, http://www.statisticalengineering.com/mh1823/, (2016).

[3] ASTM E 2862-12, "Standard Practice for Probability of Detection for Hit/Miss Data," ASTM International, (2012).

[4] Rummel Ward D., "Recommended Practice for Demonstration of Nondestructive Evaluation (NDE) Reliability on Aircraft Production Parts," Materials Evaluation, August Issue 40 pp 922, (1988).

[5] Koshti, A. M., Optimizing probability of detection point estimate demonstration", Nondestructive Characterization and Monitoring of Advanced Materials, Aerospace, and Civil Infrastructure, Proc. of SPIE Vol. 10169, (2017).

[6] Koshti, A. M., "Modeling the X-ray Process and X-ray Flaw Size Parameter for POD Studies," Proc. SPIE 9063, Nondestructive Characterization for Composite Materials, Aerospace Engineering, Civil Infrastructure, and Homeland Security, (2014).

[7] Koshti, A. M., "Simulating the X-ray Image Contrast to Set-up Techniques with Desired Flaw Detectability," Structural Health Monitoring and Inspection of Advanced Materials, Aerospace, and Civil Infrastructure, Proc. of SPIE Vol. 9437, (2015).

[8] Koshti, A. M., "Crack detection flaw size parameter modeling for x-rays at grazing angle to crack faces", Nondestructive Characterization and Monitoring of Advanced Materials, Aerospace, and Civil Infrastructure, Proc. of SPIE Vol. 10169, (2017).

[9] Koshti, A. M., "X-ray ray tracing simulation and flaw parameters for crack detection," SPIE Smart Structures and NDE, Proc. of SPIE Vol. 10600, (2018).

[10] Koshti, A. M., "Methods and Systems for Characterization of an Anomaly Using Infrared Flash Thermography," National Aeronautics and Space Administration, U.S. Patent US8577120 B1, (2013).

[11] Koshti A.M., "Flash Infrared Thermography Contrast Data Analysis Technique," NASA Technical report server, http://ntrs.nasa.gov, Doc. ID. 20140012757, NASA Tech Briefs Webinar; Meeting Sponsor, NASA Johnson Space Center; Houston, TX, (2014).

[12] Koshti, A. M., "Infrared Contrast Analysis Technique for Flash Thermography Nondestructive Evaluation, Document ID: 20140012756, Report/Patent Number: AK3-14, JSC-CN-32013; Patent US 8,577,120, Technical Reporthttps://ntrs.nasa.gov/, (2014).

[13] Koshti, A. M., "Measuring and Estimating Normalized Contrast in Infrared Flash Thermography," NASA Technical report server, http://ntrs.nasa.gov, Doc. ID: 20130009802, NASA Tech Brief, NASA Johnson Space Center; Houston, TX, United States (2013).

[14] Koshti, A. M., "Infrared Contrast Data Analysis Method for Quantitative Measurement and Monitoring in Flash Infrared Thermography," Structural Health Monitoring and Inspection of Advanced Materials, Aerospace, and Civil Infrastructure 2015, edited by Peter J. Shull, Proc. of SPIE Vol. 9437, 94370X-2, www.spie.org, San Diego, (2015).

[15] Koshti, A. M., "Methods and Systems for Measurement and Estimation of Normalized Contrast in Infrared Thermography," National Aeronautics and Space Administration, US Patent US 9,066,028 B1, (June 23, 2015).

[16] Koshti, A. M., "Normalized Temperature Contrast Processing in Infrared Flash Thermography," SAMPE Long Beach, (2016).

[17] Koshti, A. M., "Applicability of a Conservative Margin Approach for Assessing NDE Flaw Detectability," Aging Aircraft 2007 Palm Springs CA, http://ntrs.nasa.gov/, (2007).

[18] Koshti, A. M., "Assessment of dye penetrant crack detectability in external corners using similarity analysis," SPIE Smart Structures and NDE, Proc. of SPIE Vol. 10599, (2018).

[19] Koshti, A.M., "Reliably detectable flaw size for NDE methods that use calibration", Nondestructive Characterization and Monitoring of Advanced Materials, Aerospace, and Civil Infrastructure, Proc. of SPIE Vol. 10169, (2017).

[20] Koshti, A. M., "Eddy current crack detection capability assessment approach using crack specimens with differing electrical conductivity," SPIE Smart Structures and NDE, Proc. of SPIE Vol. 10599, (2018).

[21] Koshti, A. M., "NDE flaw detectability validation using smaller number of signal response data-points", NASA Technical Report Server, http://ntrs.nasa.gov, NASA Johnson Space Center, (2018).

[22] Koshti, A. M., "NDE flaw estimation using smaller number of hit-miss data-points", NASA Technical Report Server, http://ntrs.nasa.gov, NASA Johnson Space Center, (2018).

Abstract/Introduction

- The paper provides an **engineering analysis approach** for assessing reliability of NDE flaw detection using smaller number of demonstration data points.
- Uses the **most basic POD and POF a-hat versus "a" model** for developing the approach empirically.
- It explores relationship of **probability of detection (POD**) and **probability of false positive (POF)** with **contrast-to-noise ratio (CNR)**, and **net decision threshold-to-noise ratio (TNR)** in a simulated data; and draws some generically applicable inferences to devise the approach.
- POD analysis of inspection test data results in an estimated flaw size, denoted by $a_{90/95}$. The flaw size has 90% POD and minimum 95% confidence.
- POD demonstration requires specimen with flaws of known size. In many situations, it is very expensive to produce the large number of flaws required for the POD analysis. In some situations, only real flaws can truly represent the flaws for demonstration. Real flaws of correct size and location within part may be difficult to produce, if not impossible.
- Based on **applicability of simulation model assumptions**, a technique is considered reliable,
  - If it provides flaw detectability size equal to or better than the **theoretical a_{90}** used in simulation and
  - If it provides a POF less than or equal to a chosen value (i.e. 0.1% or 1%)
- Engineering analysis is performed when **NDE procedure is controlled** and it is **assumed that there a_{90} exists**
  - $a_{90/95}$ is not estimated due to lack of adequate data,
  - Instead limited validation flaw size a_{lv} is estimated such that $a_{90} < a_{lv}$ with high confidence.

Abstract/Introduction - Continued

- **Linear correlation** is used between the signal response data and flaw size.
  - POD software mh1823 uses generalized linear model (GLM) in POD analysis after transforming the flaw size and signal response, if needed, using logarithm. Therefore, this approach is in agreement with the linear signal correlation used in mh1823.
- Using the **simulated POD analysis of data**, generic conditions on **contrast-to-noise ratio (CNR)** and **net decision threshold-to-noise ratio (TNR)** are derived for reliable flaw detection empirically.
  - The conditions **may be obtained theoretically*** but that is not part of this paper.
- In order to assess **technique reliability** using the engineering approach,
  - **1. signal response-to-flaw size correlation** about the flaw size of concern is needed.
  - 2. In addition, **measurement of noise** is also needed.
  - If the technique meets the above requirements, **assumption of linear signal-to-flaw size correlation** and **conditions on noise**, then the technique can be assessed using this analysis as it fits the underlying POD model used here.
- The approach is conservative and is designed to provide a larger flaw size compared to the POD approach.
  - Such NDE technique assessment approach, although, not as rigorous as POD, can be cost effective if the larger flaw size can be tolerated.
  - Typically, this is a situation for all quality control NDE inspections. Here, an NDE technique needs to be reliable and the true a_{90} is not known, but the assessed flaw size is assumed to be larger than the unknown a_{90} due to conservative factors or margins.
- Applicability of the approach for assessing reliability of flaw detection **in x-ray radiography and 2D i**maging in general is also explored.

** Review comment by Floyd Spencer, NASA NESC*

Empirical Model for Assessment of Flaw Detectability

This approach is based on hypothesis that simulated data used in $\hat{a}$ versus "a" curve-fit POD or $\hat{a}$ versus "a" mh1823 POD analysis can be used to devise necessary conditions for engineering analysis for assessment of technique reliability. Therefore, if POD methods are used to determinate POD curves, perform noise analysis, choose decision threshold, and perform POF analysis, then this information can be used to devise the necessary conditions for the engineering analysis. The following linear signal response versus flaw size model is used. Signal response $\hat{a}$ relates to flaw size "a" as follows.

$$\hat{a} = \beta_1 a + \beta_0 + \delta, \quad (1)$$

where, β_0 and $\beta_1$are constants. Noise δ is assumed to have **Normal distribution** with constant standard deviation σ. First, a symmetrical POD function curve based on **error function (erf)** is chosen. This is given by cumulative density distribution of a probability density function, which is chosen to be a Normal distribution. This meets the key assumption that POD increases with flaw size. **Probability density function (**PDF), in the form of Normal distribution, is given by,

$$f(a) = \frac{1}{\sigma^* \sqrt{2\pi}} e^{-\frac{(a-\mu)^2}{2{\sigma^*}^2}} \quad (2)$$

POD function is given by cumulative density distribution function (CDF) of the Normal distribution function PDF. It is given by,

$$g(a, \mu, \sigma^*) = \frac{1}{2}\left[1 + erf\left(\frac{a-\mu}{\sigma^* \sqrt{2}}\right)\right], \quad (3)$$

where, μ is mean of the PDF and CDF functions at a given decision threshold. σ and σ^* are standard deviations of noise δ and for PDF (or CDF) function at a given decision threshold. 90% POD is given by following expression, $0.9 = g(1.2815, 0, 1)$.

Measured noise δ and Standard deviation σ^* of POD fit

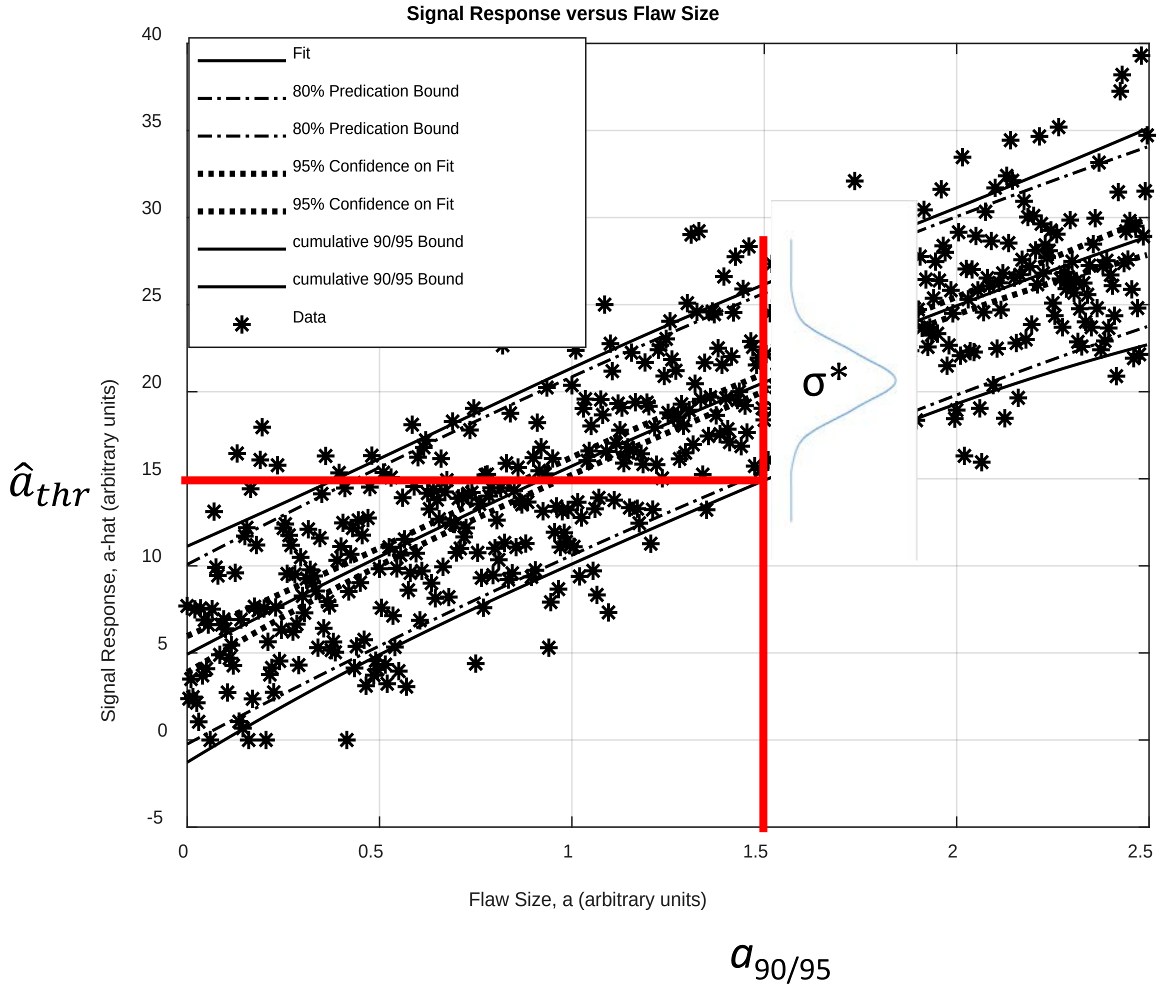


Fig. 1: Signal response versus flaw size.

Noise δ is assumed to have Normal distribution with constant standard deviation σ.

$$f(a) = \frac{1}{\sigma^*\sqrt{2\pi}} e^{-\frac{(a-\mu)^2}{2\sigma^{*2}}}. \qquad (4)$$

Ratio of standard deviation of 90/95 fit of POD Model to standard deviation of noise is also called ratio of standard deviation of noise here. It is denoted by $R_\sigma$and is given by,

$$R_\sigma = \sigma^*/\sigma. \qquad (5)$$

Noise δ is measured in flaw free area.

Standard Deviation of Noise Ratio, R_σ

The noise ratio R_σ is plotted below. It is between 1.06 to 1.2.

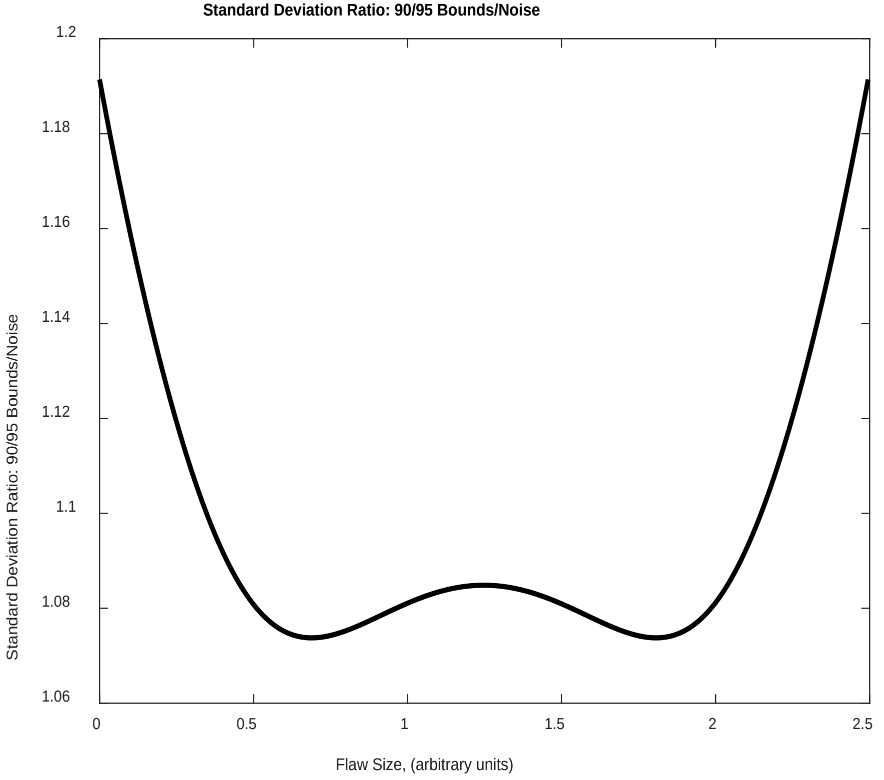


Fig. 2: Standard deviation ratio versus flaw size.

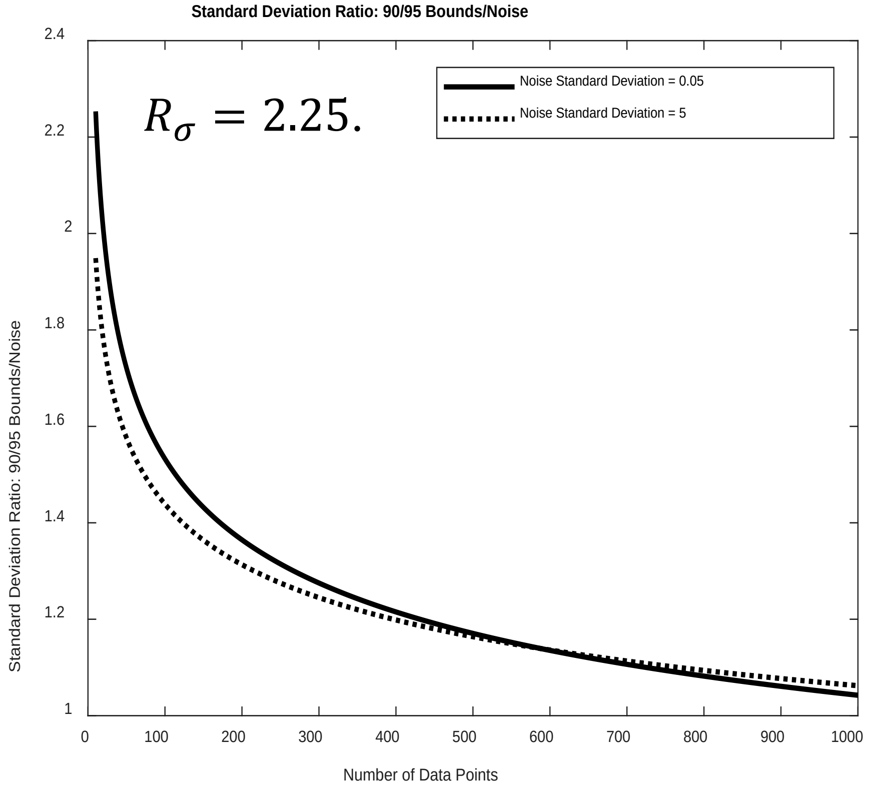


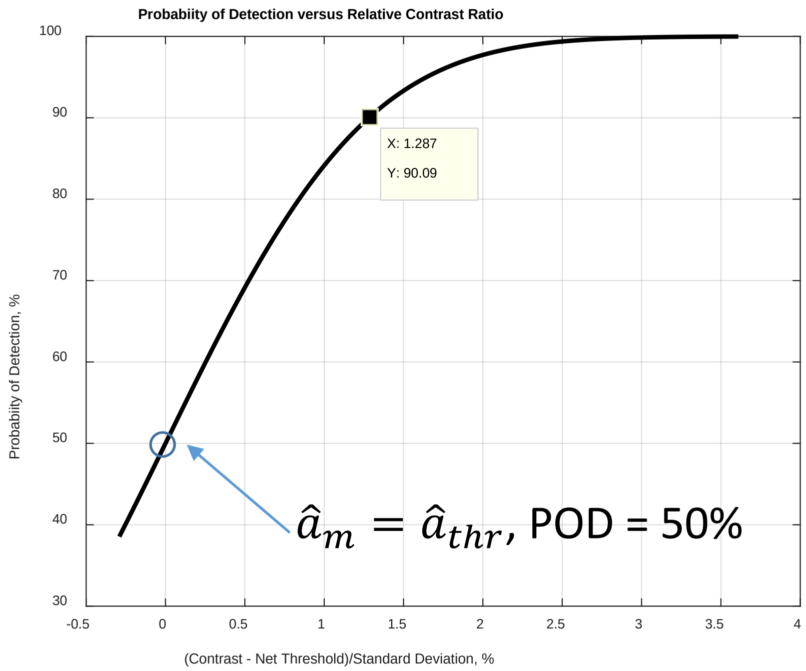
Fig. 3: Standard deviation of noise ratio versus data points.

The upper curve is more conservative. Fit equation for the upper curve is given below.

$$R_\sigma = 3.313 \times \mathrm{n}^{-0.1674}. \qquad (6)$$

Notice that the noise ratio R_σ range is from 1 for over 1000 data points to 2.25 for 10 equally distributed data points around target flaw size. Conservatively, we can take 2.25 as the worst case value in this paper.

POD Model



Contrast is given by,

$$c = \hat{a}_m - \beta_0. \quad (9)$$

Decision threshold $\hat{a}_{thr} = \hat{a}_{90/95}$ (for this simulation).

Net decision threshold is given by,

$$\hat{a}_{thr\_net} = \hat{a}_{thr} - \beta_0. \quad (10)$$

Using the assumed POD model (Eq. (3)), a condition based on relative contrast ratio $c_{rel}NR$, is given below.

$$POF = 1 - cdf(\hat{a}_{thr}) \quad (7)$$

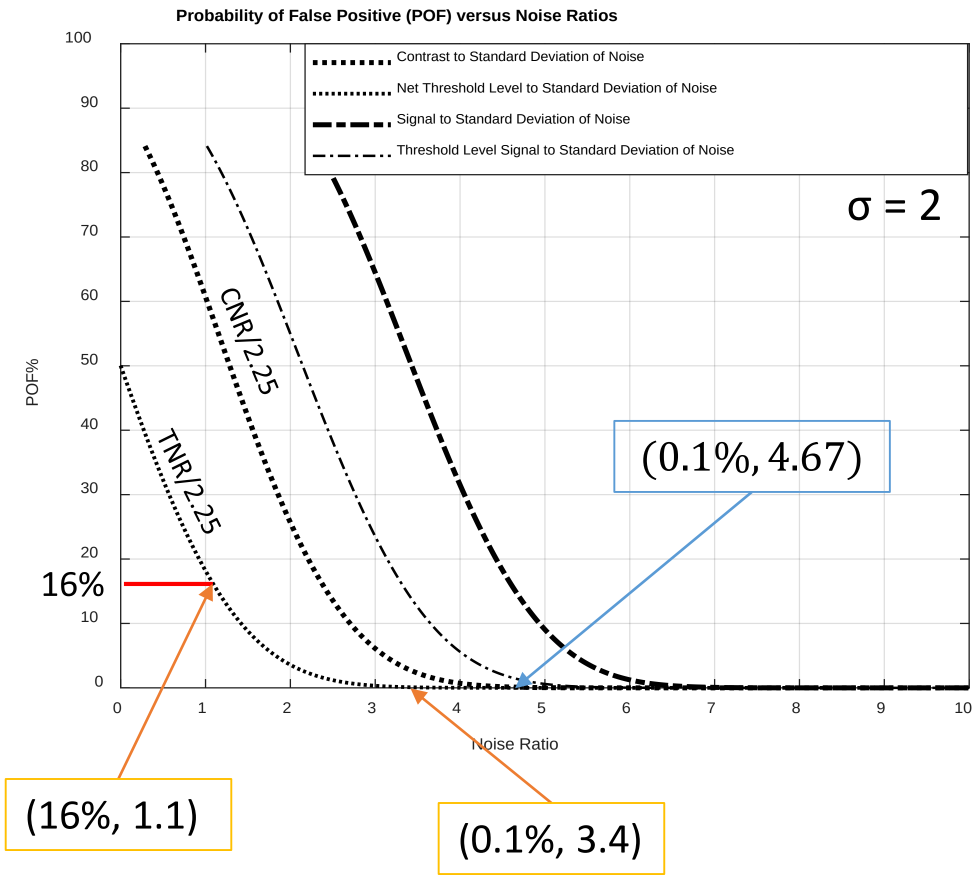
$$c_{rel}NR = \left((\hat{a}_m - \beta_0) - (\hat{a}_{thr} - \beta_0) \right) / \sigma^* \geq 1.285. \quad (11)$$

Based on modeling noise as Standard Distribution, 90% percentile or as cumulative noise,

$$n_{90} = 1.285\, \sigma^* \quad (8)$$

CNR and TNR for 0.1% POF

Change Decision Threshold and compute CNR and TNR a for different σ = 2 and 4.



CNR and TNR versus POF are invariants with respect to noise.

Contrast to Standard Deviation of Noise (CNR) is given by,

$$CNR = (\hat{a}_m - \beta_0)/\sigma = R_\sigma \times 4.4 = 2.25 \times 4.67 = 10.5. \quad (12)$$

Net threshold to Standard Deviation of Noise (TNR)is given by

$$TNR = (\hat{a}_{thr} - \beta_0)/\sigma = R_\sigma \times 3.4 = 2.25 \times 3.4 = 7.65. \quad (13)$$

Contrast to net threshold ratio is given by,

$$CTR = CNR/TNR \quad (14)$$

Conditions for Reliable Flaw Detection

Table 1: Conditions for reliable flaw detection, noise ratio = 2.25

Condition	Description	Abbreviation	POF = 0.1%	POF = 1%	Change
1	Difference in contrast and net threshold normalized to standard deviation of 90/95 bounds, Eq. (18)	$C_{rel}NR$	≥1.285	≥1.285	0
2A	Contrast-to-standard deviation of noise ratio, Eq. (21)	CNR	≥ 10.5	≥ 8.66	1.845
2B	Contrast-to-net noise ratio, Eq. (24)	CNR	≥ 4.45	≥ 3.67	0.78
3A	Net threshold-to-standard deviation of noise, Eq. (22)	TNR	≥ 7.65	≥ 5.76	1.89
3B	Net threshold-to-net noise, Eq. (25)	TNR	≥ 3.24	≥ 2.44	0.8
4	Ratio of the contrast-to-net threshold ratio, Eq. (26)	CTR	~1.37	~1.5	-0.13

Assumed true

Calculated using condition 1

Merit Conditions CNR, TNR and CTR Can be calculated for NDE application with signal response that follows condition 1.

Multi-hit POD and POF in X-ray Radiography
CNR = 2.5, Threshold = Average Signal, Conservative Flaw Detection

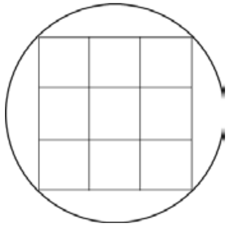


Fig. 10: Boundary of image of round void or hole of IQI and cluster of resolution size pixels.

If probability of detecting a single pixel size flaw is *P*, then probability of flaw detection P_m in cluster, is given by,

$$POD_m = 1 - (1 - POD_1)(1 - POD_2) \ldots\ldots (1 - POD_i), \quad (15)$$

If each pixel has same POD,

$$POD_m = 1 - (1 - POD_i)^N \quad (16)$$

where, *N* = number of pixels in a cluster, i.e. N = 9 in this case. POF for pixel cluster is given by,

$$POF_m = POF_1 POF_2 \ldots . POF_i. \quad (17)$$

If each pixel has same POF,

$$POF_m = POF_i{}^N \quad (18)$$

Assume that each pixel samples non-overlapping area of discontinuity.

This can be assumed to be true if pixel size ≥ resolution size.

A point of 15.68% POF (net threshold-to-standard deviation of noise ratio TNR = 1.1)

This point was chosen so that contrast-to-standard deviation of noise ratio (CNR) is 2.5 when decision threshold is same as average contrast. Each pixel is assumed to have CNR ≥ 2.5.

$$POF_m = POF_i{}^N = 0.1568^9 = \sim 0\%. \quad (19)$$

$$(\hat{a}_{thr} - \beta_0)/\sigma = 2.25 \times 1.1 = 2.5. \quad (20)$$

$$\hat{a}_{thr} = \hat{a}_m \quad (21)$$

$$CNR = (\hat{a}_m - \beta_0)/\sigma = 2.25 \times 1.1 = 2.5. \quad (22)$$

$$POD_m = 1 - (1 - 0.5)^9 = 99.8\%. \quad (23)$$

Multi-hit POD and POF in X-ray Radiography
CNR = 2.5, Threshold = Crossing Point, Extreme Visual Detection

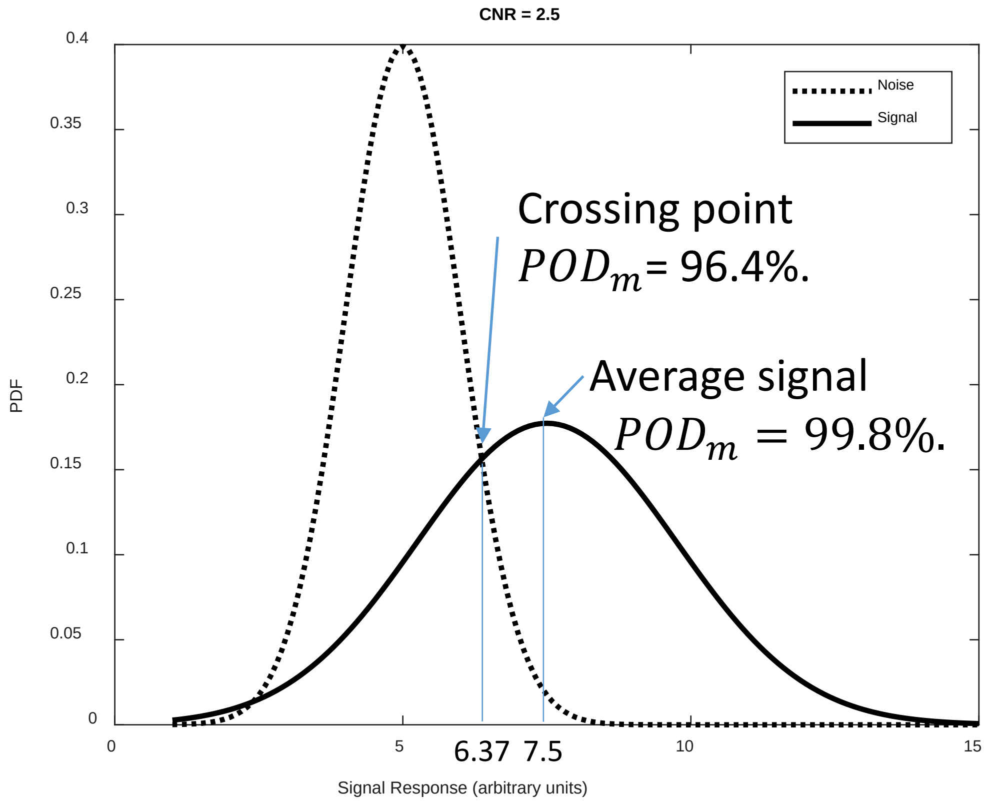


Fig. shows Normal probability density distributions for signal and noise for CNR = 2.5.

For visual detection, the decision threshold is likely to be between the crossing of the two distributions i.e. at signal response of 6.37 units in Fig. and the average signal response is 7.5 units.

Here, we consider the extreme POD case of visual flaw detection with decision threshold at the crossing point i.e., a_{thr} = 6.37 units.

$C_{rel}NR$ ~ 0.5 => POD_i = 69.2% for a single pixel.

Therefore, using for 9 pixels, the POD_m= 96.4%.

POF_i = 8.5%. $POF_m = POF_i{}^N = 0.08^9 = \sim 0\%$.

This POD is smaller than 99.8% which was previously calculated for decision threshold at average contrast level. But visually detected flaw size is smaller than that detected by using decision threshold as average signal response of a target flaw. Thus, for visual detection also, the detection for CNR = 2.5 is reliable.

Example of CNR = 2.5 and CNR = 10 Indications

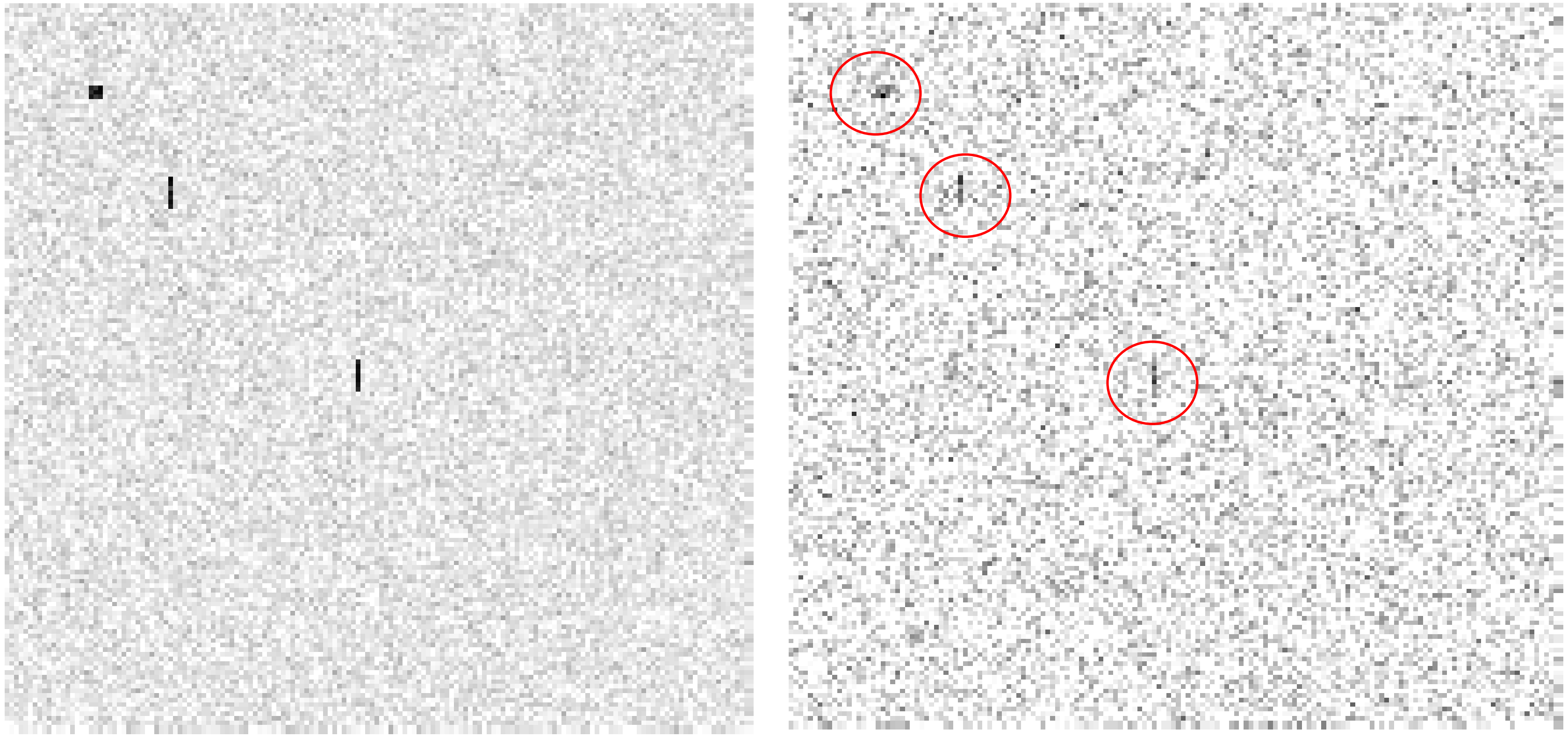


Fig. 12: Simulated CNR = 10 in left image and CNR = 2.5 in right image. One 3 x 3 cluster indication in top left and two 7 x 1 linear indications.

The images confirm that the indications are visually detectable at CNR = 2.5.

Theoretical Resolution in X-ray radiography

Estimated resolution or total unsharpness U_{lm} per ASTM E2698 is given by,

$$U_{lm} = \frac{1}{\nu} \sqrt[3]{\left(U_g\right)^3 + (1.6\,SRb)^3} \text{ and,} \quad (24)$$

$$U_g = (\nu - 1)ø, \quad (25)$$

where U_g is geometric unsharpness. v is the largest geometric magnification present in the image which happens at maximum distance of point on object from detector. ø is the x-ray source focal spot size per ASTM E1165 and the detector basic resolution SRb is calculated using method specified in ASTM E2597.

$$SRb \cong 1.3\, x\, d \quad (26)$$

where, d = detector pixel size.

Therefore, minimum contrast sensitivity at a void needed to meet contrast-to-noise ratio of 2.5 is given by,

$$CS_{void} = \frac{CNR \times CS}{(2.36 MTF_{3x3})} = \frac{2.5 \times 2\%}{2.36 \times 0.8} = 2.65\%. \quad (27)$$

Conditions for Reliable Flaw Detection in X-ray Radiography

Table 2: Conditions for reliably detecting 4.2 x U_{lm} diameter void with minimum 2.65% thickness.

Condition	Description	
1A	Contrast-to-standard deviation of noise ratio on void	≥ 2.5
1B	Signal-to-standard deviation of noise in acreage	≥ 130
2	Contrast sensitivity	≤ 2%

Conclusions

- Engineering analysis rule of thumb or cook-book conditions are given based on analysis of simulated data to assess reliability of an NDE technique.
  - If merit ratio conditions are met and POD model assumption is valid, then the NDE technique is considered to be reliable for engineering estimate. Although $a_{90/95}$ flaw size is not estimated due to lack of sufficient number of flaw detection datapoints.
  - The NDE technique is considered reliable for those applications where $a_{90/95}$ is not needed.
  - The approach assumes linear correlation between signal and flaw size.
  - Noise is assumed to have constant standard deviation.
  - Minimum 10 data points are recommended in signal correlation and noise measurements. The approach is conservative and is designed to provide a larger flaw size compared to the POD approach.
  - Three merit ratio conditions (CNR, TNR and CTR) are provided. All three should be checked.
  - The assessed flaw size in this analysis has high confidence that it is larger than the unknown true a_{90} due to conservative factors used in the analysis.
- Assessment of reliability of x-ray radiography NDE, including film, DR, CR and CT, is also considered.
  - For reliable detection of 4.2 x U_{lm} diameter void with minimum 2.65% thickness, two merit ratio conditions are provided
    - CNR or SNR and Contrast Sensitivity
  - This approach is also applicable to assessment of reliability of flaw detection in other 2D imaging techniques.
- The analysis indicates that, multi-hit detection in 2D pixel cluster to image flaw is inherently more reliable than using just single-hit detection similar to that using only real time A-scan signal display.