AFRL-RX-WP-TP-2012-0356

STATISTICAL ANALYSIS OF EDDY CURRENT DATA FROM FASTENER SITE INSPECTIONS (PREPRINT)

Jeremy S. Knopp

Materials State Awareness & Supportability Branch
Structural Materials Division

Ramana V. Grandhi and Inseok Park

Wright State University

John C. Aldrin

Computational Tools

July 2012
Interim

AIR FORCE RESEARCH LABORATORY
MATERIALS AND MANUFACTURING DIRECTORATE
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433-7750
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. REPORT DATE ***(DD-MM-YY)***	**2. REPORT TYPE**	**3. DATES COVERED *(From - To)***
July 2012	Technical Paper	1 June 2012 – 1 July 2012

4. TITLE AND SUBTITLE

STATISTICAL ANALYSIS OF EDDY CURRENT DATA FROM FASTENER SITE INSPECTIONS (PREPRINT)

5a. CONTRACT NUMBER In-house

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER 62102F

6. AUTHOR(S)

Jeremy S. Knopp (RXCA)
Ramana V. Grandhi and Inseok Park (Wright State University)
John C. Aldrin (Computational Tools)

5d. PROJECT NUMBER 4349

5e. TASK NUMBER 40

5f. WORK UNIT NUMBER LP110200

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Materials State Awareness & Supportability Branch
Structural Materials Division
Air Force Research Laboratory, Materials and Manufacturing Directorate
Wright-Patterson Air Force Base, OH 45433-7750
Air Force Materiel Command, United States Air Force

8. PERFORMING ORGANIZATION REPORT NUMBER

AFRL-RX-WP-TP-2012-0356

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory
Materials and Manufacturing Directorate
Wright-Patterson Air Force Base, OH 45433-7750
Air Force Materiel Command
United States Air Force

10. SPONSORING/MONITORING AGENCY ACRONYM(S)

AFRL/RXCA

11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)

AFRL-RX-WP-TP-2012-0356

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited. Preprint to be submitted to Journal of Nondestructive Evaluation.

13. SUPPLEMENTARY NOTES

The U.S. Government is joint author of this work and has the right to use, modify, reproduce, release, perform, display, or disclose the work. PA Case Number and clearance date: 88ABW-2012-3396, 13 June 2012. This document contains color.

14. ABSTRACT

Recent work on reliably detecting and characterizing cracks in multi-layer airframe structures has used modeling and simulation to extract features from raw eddy current data, and to assist in the evaluation of probability of detection (POD). This paper focuses on the statistical analysis of the data from these studies. Hit/miss, linear, and physics-inspired methods are employed to evaluate POD. The Box-Cox transformation is used as a remedy for violations of homoscedasticity. In addition, a bootstrapping method is introduced for confidence bound calculation on a 2nd order linear model. The objective of this work is to provide on insight how different models and assumptions impact POD evaluation.

15. SUBJECT TERMS

probability of detection, eddy current, fastener site inspection, bootstrap confidence intervals, box-cox transformation

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT:**	**NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** (Monitor)
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Mark Blodgett
Unclassified	Unclassified	Unclassified	SAR	24	**19b. TELEPHONE NUMBER** *(Include Area Code)* N/A

Statistical Analysis of Eddy Current Data from Fastener Site Inspections

Jeremy S. Knopp[1] Ramana Grandhi[2], John C. Aldrin[3], and Inseok Park[2]

[1]Air Force Research Laboratory (AFRL/RXLP),
Wright-Patterson AFB, OH 45433, USA

[2] Wright State University,
Colonel Glenn., Dayton, OH, USA

[3] Computational Tools,
4275 Chatham Ave., Gurnee, IL 60031, USA

ABSTRACT

Recent work on reliably detecting and characterizing cracks in multi-layer airframe structures has used modeling and simulation to extract features from raw eddy current data, and to assist in the evaluation of probability of detection (POD). This paper focuses on the statistical analysis of the data from these studies. Hit/miss, linear, and physics-inspired methods are employed to evaluate POD. The Box-Cox transformation is used as a remedy for violations of homoscedasticity. In addition, a bootstrapping method is introduced for confidence bound calculation on a 2nd order linear model. The objective of this work is to provide on insight how different models and assumptions impact POD evaluation.

KEYWORDS

Probability of Detection, eddy current, fastener site inspection, bootstrap confidence intervals, Box-Cox transformation

INTRODUCTION

There are two conventional approaches to evaluating POD: hit/miss analysis or â vs a analysis [1]. Hit/miss analysis is still the most widely used to method to determine reliability of Nondestructive inspections, but it is advantageous to use â vs a analysis because the information in the signal response can be used for the parameter estimates and confidence bounds. This provides a more informative reliability assessment and may require less samples than hit/miss analysis. Two major requirements for â vs a analysis are a linear relationship between flaw size and signal response, and constant variance. In many cases, a logarithmic transformation can be applied if the linear requirement is not met, but if this fails, there are limited options other than hit/miss analysis. It is also possible that a logarithmic transformation can address a violation of constant variance, but this isn't always the case. Given that these conditions of linearity and homoscedasticity are often not met with real NDI data, it is useful to explore remedial measures such as transformations so that the full signal response of NDI data can be used more frequently in practice. In addition, if the linear assumption is not met after the data is transformed, there is an additional question of how to properly put confidence bounds on a POD result that is derived from a more complicated measurement model. A case study problem is presented for exploring these issues in POD evaluation. Prior work on detecting subsurface cracks in multi-layer airframe structures used novel methods to extract features useful for POD analysis [2-4]. A preliminary model-assisted POD study was conducted based on those efforts [5-7]. In the previous work [5], hit/miss analysis was chosen because visual inspection of the data indicated that there was a violation of the constant variance assumption and possibly the linear assumption. In this work, a Box-Cox transformation will be used to mitigate, at least in part, concerns about heteroscedasticity. If constant variance can be achieved with this transformation and the linear assumption is met, then â vs a analysis can be performed according to the methods set forth in Berens' classic work on the subject [1], and for the most part, still considered state-of-the-art today [8-10]. If constant variance is achieved, but the linear assumption is not met, then methods need to be developed for more complicated models. It was difficult to determine by visual inspection of the data whether a linear model was most appropriate for the data set, so additional modeling and simulation studies have been conducted to determine the model form of the response that can be expected with this type of inspection. While the model itself is not used in this study, it inspired the use of a 2nd order linear model; thus it is referred to as a "physics-inspired" model rather than a physics-based model. Lastly, it has been found that bootstrapping is a very easy and useful method for providing confidence bounds on POD curves, and its use will be illustrated with some examples.

EXPERIMENTAL DATA

The experimental problem of interest is the detection of cracks under installed countersunk fasteners in airframe structures. The description of the data and how it was processed is provided in detail in prior papers [4, 5]. The sample set contained over 300 fastener sites with cracks in the 1st layer and 2nd layer at the faying surface. In this paper, only the 1st layer cracks are considered, and there are a total of 171 observations. The dimensions for the thickness of the top and bottom layers measured 3.96 mm and 2.54 mm respectively. Conductivities of 1.87 E7 S/m for the aluminum layers and 1.79 E6 S/m for the titanium fasteners were considered. The radius of the fastener hole was 4.04 mm. The probe was operated at 600 Hz and had coil dimensions of 6.0 mm in height, 3.0 mm in inner radius and 6.0 mm in outer radius. A corner crack model for the first layer was considered with the assumed aspect ratio length, a, to width, b, of 1:1. Crack lengths in the experimental samples were available between 0.0 to 4.3 mm.

Model-based image processing methods were used to extract features in the scans that correlate to flaw size [3]. This model-based approach essentially fits models based on first-principles to image data in order to enhance crack indications in the presence of coherent noise from the fastener site, adjacent fasteners and panel edges. The final step is to extract a quantitative metric associated with the crack condition off-axis from each fastener site center. This same analysis process was applied to all experimental and simulated data to facilitate proper comparison. The raw data is displayed in Figure 1. A previous analysis of data from these samples used binary logistic regression because visual inspection of the data revealed that the homoscedastisticy assumption was violated. It is clearly observed that the variance increases as a function of flaw size. There is also another current study investigating a similar set of inspection data, with an alternative approach to the statistical analysis [11].

LINEAR MODEL ANALYSIS WITH BOX-COX TRANSFORMATION

In this analysis, 'â' is the magnitude of the eddy current signal response, and 'a' refers to crack length. For cases where there is a relationship between the mean response and variance, the Box-Cox transformation is used to stabilize the variance. This method assumes that the relationship between the error variance σ_i^2 and and mean response μ_i can be described with a power transformation on â in the form of equation 1. The new regression model in equation 2 will include the additional λ parameter which will also need to be estimated.

$$\hat{a}' = \hat{a}^{\lambda} \tag{1}$$

$$\hat{a}_i^{\lambda} = \beta_0 + \beta_1 a_i + \varepsilon_i \tag{2}$$

Following a method outlined in Kutner et al [12], a numerical search procedure is set up to estimate λ. The â observations are first standardized so that the order of magnitude error sum of squares isn't dependent on the value of λ.

The standardized observations are:

$$g_i = \frac{1}{\lambda c^{\lambda-1}}\left(\hat{a}_i^{\lambda} - 1\right),\ \lambda \neq 0, \tag{3}$$

$$g_i = c(\ln(\hat{a}_i)),\ \lambda = 0, \tag{4}$$

where $c = (\prod \hat{a}_i)^{1/n}$, and n is the total number of observations, which happens to be the geometric mean of the observations. Once these standardized observations are obtained, they are then regressed on 'a', which in this case is crack length, and then the sum of squares error (SSE) is obtained. The optimization problem is formulated such that the objective is to minimize SSE with λ as a single parameter to be adjusted. Microsoft Excel's Solver add-in was used to determine the value of λ which minimizes SSE.

Before this procedure is illustrated, a 0.02 offset is added to the raw data. This facilitates the analysis using these transformations. For this data, λ = 0.45 is the transformation that minimizes the SSE. Note that if λ = 0.5, it is simply a square root transformation. This procedure only provides a general estimate of a preferred transformation, and is not quantitative in a rigorous sense, so for the sake of using a familiar transformation, further analysis will use the square root transform. Both values of λ will be used to provide an idea of the sensitivity of POD results to the choice of transformation. The transformed data, â vs a analysis, and the POD curve is shown in Figure 2, 3, and 4 respectively for λ = 0.45. The left censor value is chosen to be 0.13, the right censor is not used, and the detection threshold is set to 0.23. The following parameter estimates are obtained for the linear regression model: $\hat{\beta}_0$ = 0.166, $\hat{\beta}_1$ = 0.045 and τ = 0.026, where τ is the regression standard deviation. The a_{90} value is 2.176 mm and the $a_{90/95}$ value is 2.327 mm.

The same analysis is conducted for the λ = 0.5 transformation, since that will be used from now on. The detection threshold for this transformation is 0.195, and the left censor value is 0.14. The transformed data, â vs a analysis, and the POD curve is shown in Figure's 5, 6, and 7 respectively. The following parameter estimates are obtained for the linear regression model: $\hat{\beta}_0$ = 0.135, $\hat{\beta}_1$ = 0.043 and τ = 0.024. The a_{90} value is 2.102 mm and the $a_{90/95}$ value is 2.257 mm.

ANALYSIS WITH PHYSICS-INSPIRED MODEL

There is some precedent for using a physical model of an inspection to improve the evaluation of POD in ultrasonic inspections [13]. In the work of Thompson and Meeker, a "kinked" regression model was developed to describe the impact of hard-alpha inclusions on POD. In particular, the physics model provided a better understanding of the small flaw regime. If the flaw is significantly smaller than the ultrasonic wavelength, it is in the Raleigh scattering regime which has a cube relationship with the flaw dimensions. Thus, 2 different linear models were needed depending on the flaw size range, and this enabled a more accurate POD analysis.

Figure 8 shows the expected signal response based on simulations in VIC-3D® [2]. The same image processing methods were applied to the simulated data. This type of response is not quite linear, so the next analysis will be with a 2nd order regression model of the form:

$$\hat{a} = \hat{\beta}_0 + \hat{\beta}_1 a + \hat{\beta}_2 a^2 + \varepsilon. \tag{5}$$

The statistical significance of a^2 in the standard regression model is 0.001, and the adjusted R-square value for the model including a^2 is 0.7754 which is slightly above 0.7619 which is for the model that includes only 'a', so there is good reason to include it in the model. Given the square root transform or $\lambda = 0.5$, the estimates for $\hat{\beta}_0$, $\hat{\beta}_1$, and $\hat{\beta}_2$ are 0.137, 0.027, and 0.005 respectively. τ is 0.0229. The censored regression has the same left censor and threshold as the 1st order model. The a_{90} value for this 2nd order model is 2.277 mm. There are no published procedures to find the $a_{90/95}$ value for this type of model. This paper introduces a very useful bootstrapping method to address this issue.

BOOTSTRAP METHODS FOR CONFIDENCE BOUND CALCLUATION

The algorithm to generate confidence bounds on more complicated models is quite simple. The main idea is to use "sampling with replacement", which interestingly wasn't used much in the statistics community until relatively recently [14], and has been used with good success in engineering [15, 16].

To illustrate how bootstrap confidence bounds are calculated, and to verify against standard methods, we need to go back to a previous â vs a analysis where confidence bound calculation method are well established. To verify, the case of the transformation parameter $\lambda = 0.5$ with the threshold set to 0.195 and 0.14 is used. This time a new data set generated by the sampling with replacement of the original data, and this new set is used to calculate a_{90}, and this process is repeated 1,000 times. The a_{90} results are then sorted in ascending order. For the case of 1,000 samples, the 950th a_{90} value is considered the value for $a_{90/95}$. Table 1 summarizes the results of this process.

	a_{90}	$a_{90/95}$
Wald Method	2.102 mm	2.257 mm
Bootstrap 1,000	2.096 mm	2.281 mm
Bootstrap 10,000	2.099 mm	2.299 mm
Bootstrap 100,000	2.099 mm	2.297 mm

Table 1: Comparison of Wald and Bootstrapping with 1,000, 10,000, and 100,000 samples.

No significant difference in a_{90} exists, and although there is a slight difference in $a_{90/95}$, the bootstrap results are on the conservative side. Based on these results, it doesn't seem necessary to sample more than 1,000 times.

This bootstrap approach was applied to the 2nd order model. Figure 9 shows the fitted 2nd order model with the transformed $\lambda = 0.5$ data. The $a_{90/95}$ using the bootstrap method with 1,000 samples is 2.472 mm.

One of the advantages of adding a^2 to the model is that there is less dependence on subjective decisions regarding censoring values and threshold values. The small flaw region is better represented with this model. Future work will involve a sensitivity study of the left censor value and threshold and the impact they have on the a_{90} and $a_{90/95}$ results.

HIT/MISS ANALYSIS

Since, the data have been examined with â vs a analysis and also with a 2nd order linear model, it is interesting to compare it with hit/miss bernoulli analysis since that is still overwhelming used to this day. The analysis will be conducted 2 different ways. At most 1 false call was recorded in the previous analysis, so one analysis shown in Figure 10 forces the false call rate to 1 by setting the threshold to 0.187. At this threshold, a_{90} = 1.72 mm and $a_{90/95}$ = 2.04 mm which are considerably smaller than the corresponding POD parameters for the other types of analysis. Secondly, the threshold is lowered substantially to 0.167 so that 2 additional flaws are detected, and this results in 11 false calls. Even smaller POD parameters are determined with a_{90} = 1.498 mm and $a_{90/95}$ = 1.907 mm. Note that this was performed with the transformed data with $\lambda = 0.5$, but it was also performed with the original data, and the exact same POD parameters were obtained corresponding with the false call rates of 1 and 11.

CONCLUSIONS AND RECOMMENDATIONS

analysis method	λ	left censor	detection threshold	false calls	a_{90} (mm)	$a_{90/95}$ (mm)	a_{90} - $a_{90/95}$ % difference
1st order linear	0.45	0.13	0.23	0	2.176	2.327	6.9%
1st order linear	0.5	0.14	0.195	1	2.102	2.257	7.3%
1st order linear	0.5	0.195	0.195	1	2.269	2.53	11.5%
2nd order linear	0.5	.14	0.195	1	2.277	2.472	8.5%

2nd order linear	0.5	0.195	0.195	1	2.197	2.428	10.5%
hit/miss	1		0.187	1	1.72	2.04	18.6%
hit/miss	1		0.162	11	1.498	1.907	27.3%

Table 2: Summary of results for different models, thresholds, and left censoring values.

Multiple statistical analysis methods were used to examine data from an eddy current inspection of fastener sites in multi-layer structures. There were notable differences in a_{90} and $a_{90/95}$ estimates for the different models. The bernoulli model contains the least information, but produces the most attractive POD. No hard conclusions can be made about this trend, but it does at least show that in at least one real case, the hit/miss results may be optimistic when compared to analysis that contains more information.

It is also interesting to note that the physics-inspired model produced similar results for the POD parameters of interest regardless of the chosen value of the left censor. Further investigations will systematically study the effect of censoring on linear and higher order models. Preliminary evidence suggests that the $a_{90/95}$ value may be invariant to the choice of the left censor value.

As more sophisticated models begin to be used in analysis of inspection data, bootstrapping is an easy and accurate way to produce confidence bounds on POD results. This was demonstrated for the the usual â vs a analysis which provided confidence (no pun intended) in the bootstrap approach. It was practical to use this method for putting confidence bounds on the 2^{nd} order model.

Future work will include Bayesian analysis using model calibration methods proposed by Kennedy and O'Hagan [17]. This will allow the physics-based model to be used directly as opposed to using physics-inspired models.

ACKNOWLEDGEMENTS

The R software environment for statistical computing and graphics was used for all statistical computations and most plots. R is open-source (free) software and is available to download here:

http://www.r-project.org/. In addition, all of the statistical analysis performed in this work made use of the MH1823 R package developed by Charles Annis [18].

REFERENCES

[1] Berens, A.P.: NDE Reliability Data Analysis, American Society for Metals Handbook Nondestructive Evaluation and Quality Control, **17**, 689-701, ASM International, (1989)

[2] Knopp, J.S. and Aldrin, J.C.: Fundamental feature extraction methods for the analysis of eddy current data, *Electromagnetic Nondestructive Evaluation Proceedings*, **31**, (2008)

[3] Aldrin, J.C. and Knopp, J.S.: "Crack Characterization Method with Invariance to Noise Features for Eddy Current Inspection of Fastener Sites", *Journal of Nondestructive Evaluation*, **25**(4), 165-181, (2006)

[4] Aldrin, J.C., Knopp, J. S., and Blodgett, M.P.: Model-based Image Processing in Eddy Current Nondestructive Evaluation, Electromagnetic Nondestructive Evaluation Proceedings, **33**, 254-261, (2010)

[5] Knopp, J.S., J.C. Aldrin, Lindgren, E. and Annis, C.: Investigation of a Model-Assisted Approach to Probability of Detection Evaluation, *Rev. Prog. Quan.t Nondestr Eval.*, **26**, 1775-1782, (2007)

[6] Aldrin, J. C., Knopp, J. S., Lindgren, E. A., Jata, K. V.: Model-assisted Probability of Detection (MAPOD) Evaluation for Eddy Current Inspection of Fastener Sites, *Rev. Prog. Quan.t Nondestr. Eval.*, **28**, 1784-1791, (2009)

[7] Aldrin, J.C. and Knopp, J.S.: Modeling and Simulation for Nondestructive Testing with Applications to Aerospace Structures, *Mat. Eval.*, **66,** 53-59, (2008)

[8] U.S. Department of Defense, MIL-HDBK-1823A Nondestructive Evaluation System Reliability Assessment, (August 2010).

[9] Gandossi, L. and Simola, K.: Derivation and use of probability of detection curves in the nuclear industry, *Insight*, **52**(12), (2010)

[10] Drury, C.G., Ghylin, K.M., and Holness, K.: Error Analysis and Threat Magnitude for Carry-on Bag Inspection, Proceedings of the Human Factors and Ergonomic Society 50th Annual Meeting, **50**, 1189-1193, (2006)

[11] Li, M., Nakagawa, N., Larson, B.F., and Meeker, W.Q.: Statistical Assessment of Probability of Detection for Automated Eddy Current Nondestructive Evaluation Inspection," Preprint available at: http://www.stat.iastate.edy/preprint/articles/2011-05.pdf (last accessed May, 2012).

[12] Kutner, M.H., Nachtsheim, C.J., Neter, J., Li, W.: Applied linear statistical models 5th edition, McGraw-Hill/Irwin, (2004)

[13] Thompson, R.B. and Meeker, W.Q.: Assessing the POD of HardAlpha Inclusions from Field Data, *Review of Progress in QNDE*, **26**, 1759-1766, (2007)

[14] Efron, B., and Tibshirani, R. J.: An Introduction to the Bootstrap, Chapman & Hall, New York, NY, (1993)

[15] McCulloch, C.C. and Murphy, J.: Local Regression Modeling for Accurate Analysis of Probability of Detection Data, *Mat. Eval*., **60**(12), 1438-1143, (2002)

[16] Amarchinta, H.K. Tarpey, T., and Grandhi, R.V.: Probabilistic Confidence Bound Framework for Residual Stress Field Predictions", *12th AIAA Non-Deterministic Approaches Conference*, AIAA-2010-2519, Orlando, FL, (2010)

[17] Kennedy, M.C. and O'Hagan A., Bayesian Calibration of Computer Models, *J. R. Statist. Soc. B.*, **63**, (2001)

[18] Annis, C.: Statistical Best-Practices for Building Probability of Detection (POD) models, R package mh1823, version 2.5.4.1, http://statisticalengineering.com/mh1823

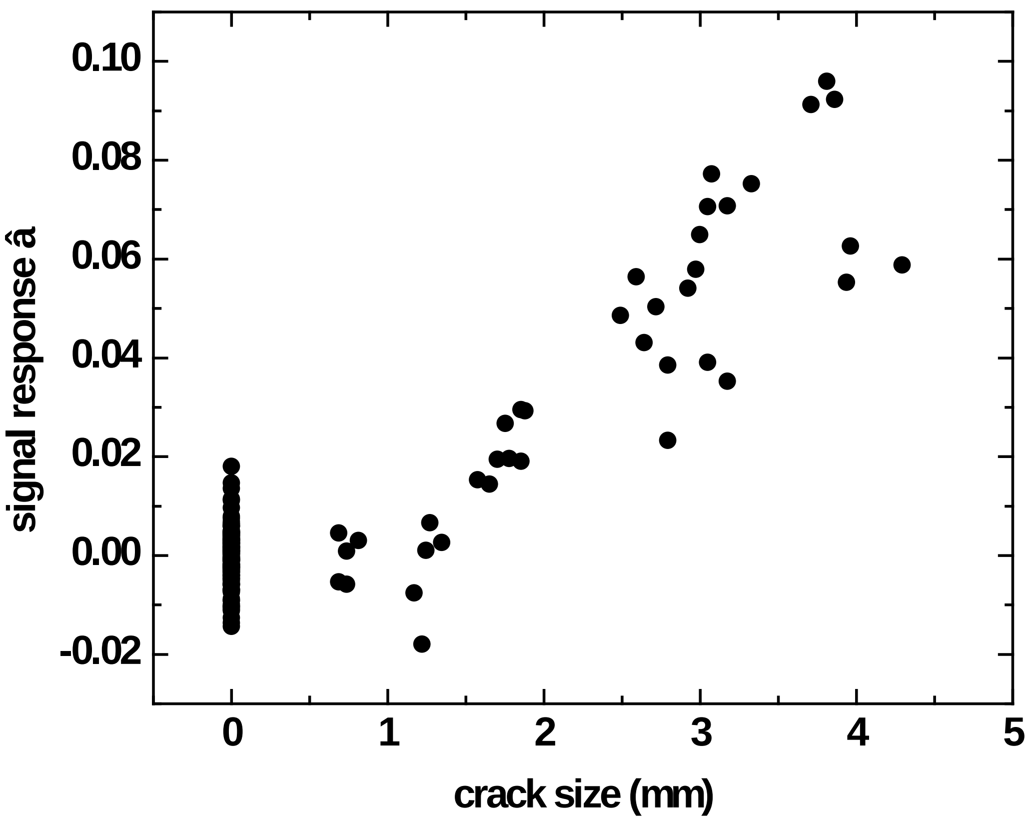

Figure 1. Original data from model-based processing of fastener site inspection data.

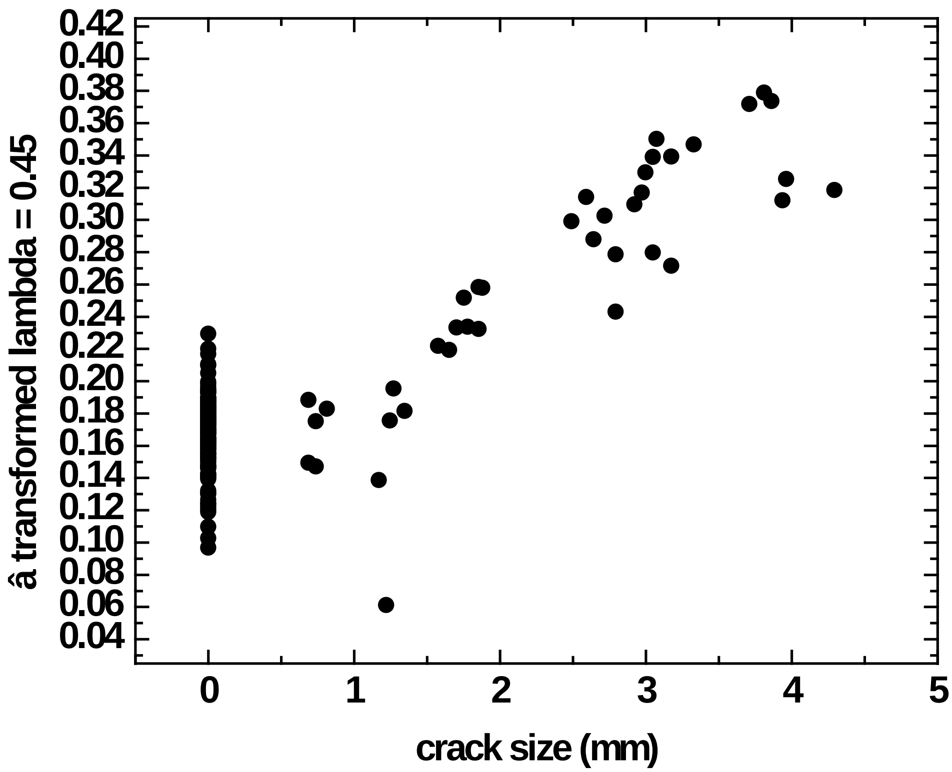

Figure 2. Box Cox transformation of original data with λ = 0.45.

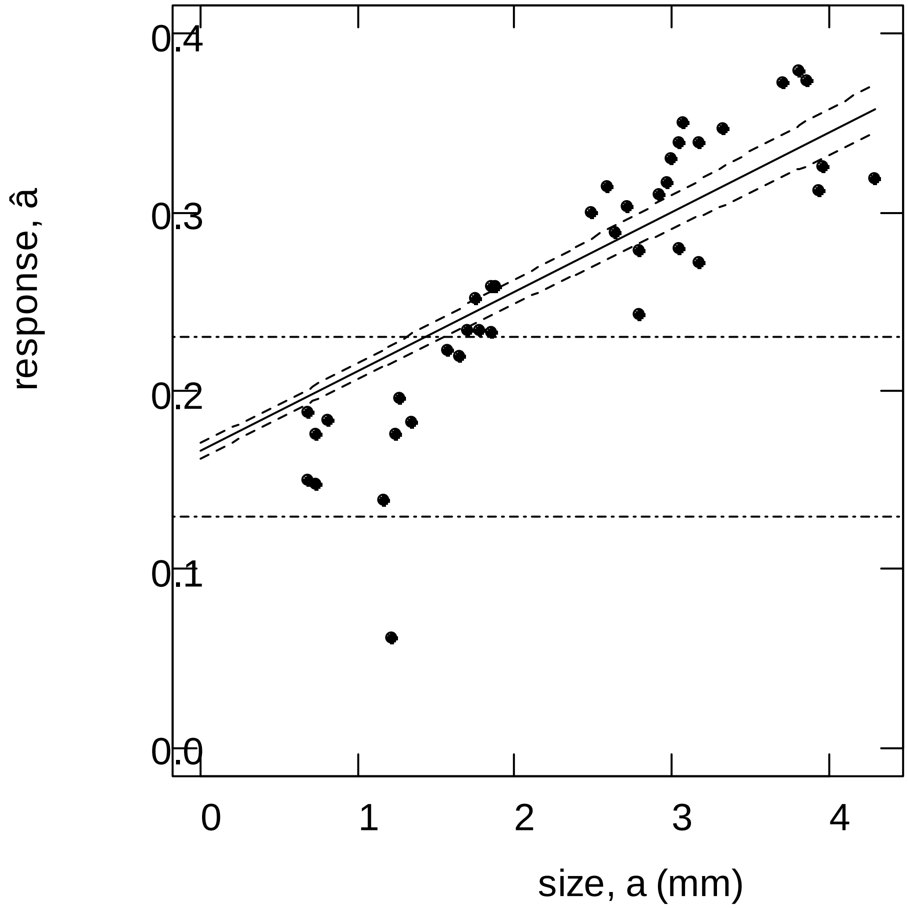

Figure 3. Linear model for transformed data with λ = 0.45.

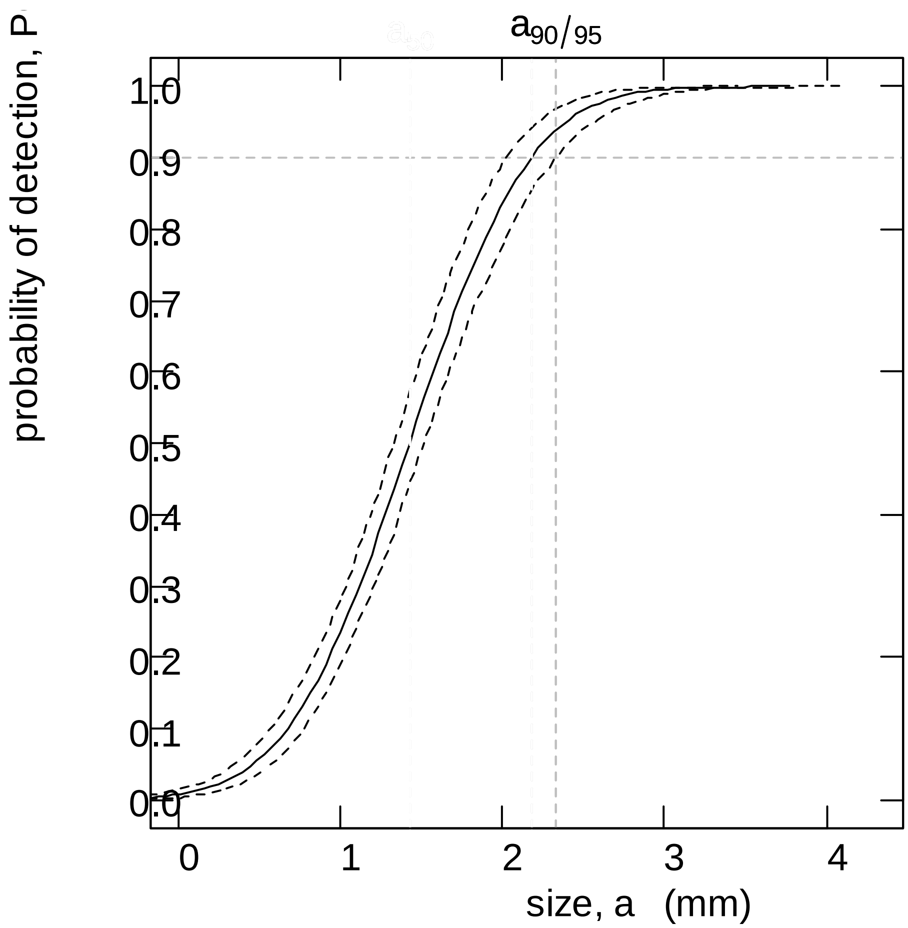

Figure 4. Probability of detection curve for transformed data with λ = 0.45

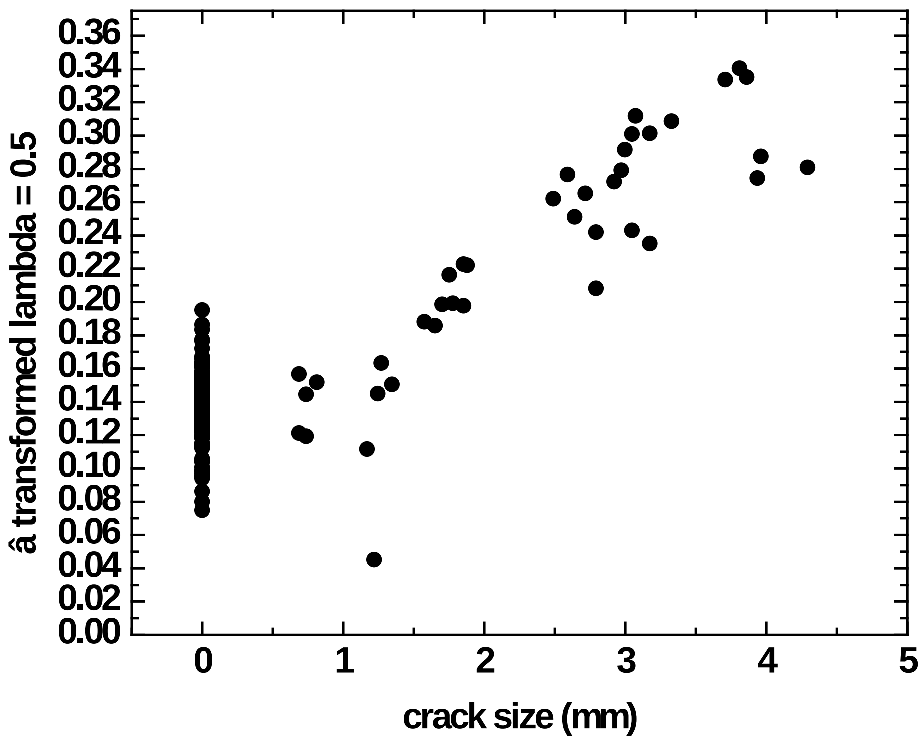

Figure 5. Square root transform of data.

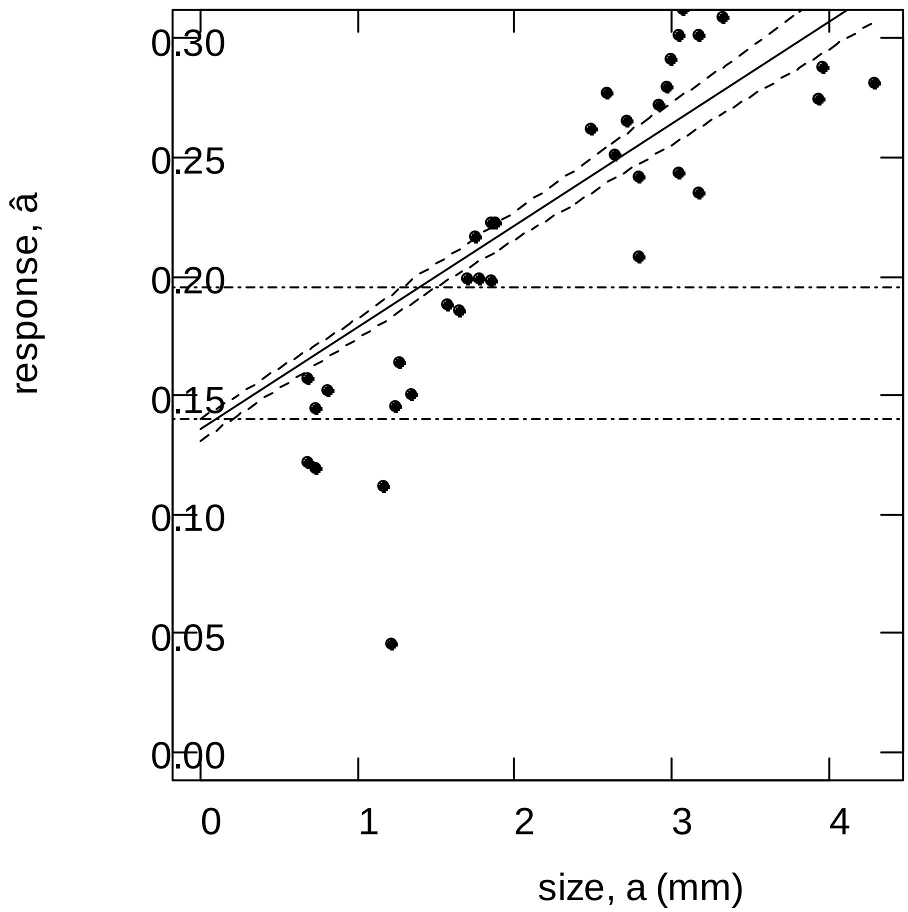

Figure 6. Linear model fit for data with square-root transformation.

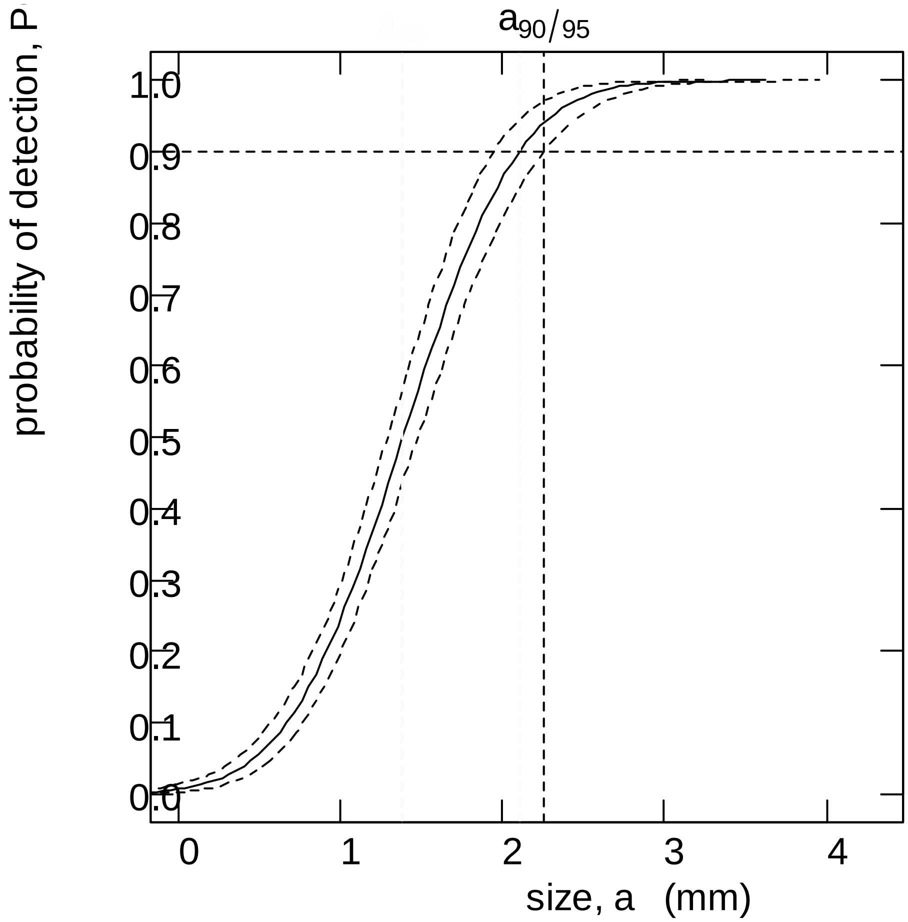

Figure 7. Probability of detection curve for data with square-root transform

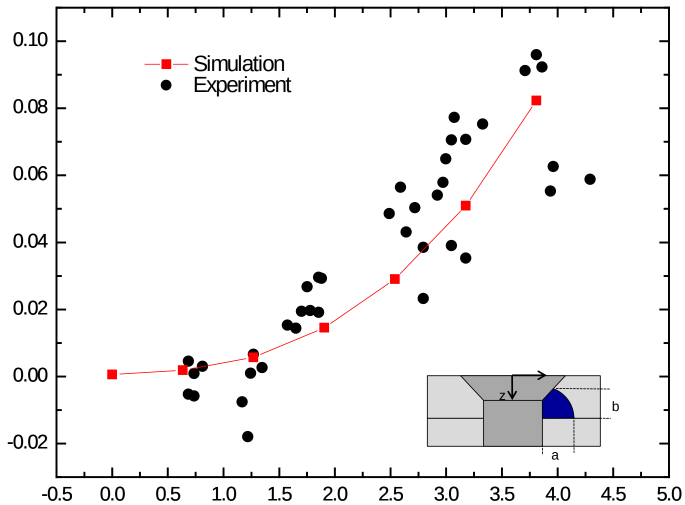

Figure 8: Comparison of experimental and simulated data for varying length of first layer corner crack with aspect ratio, a/b = 1.

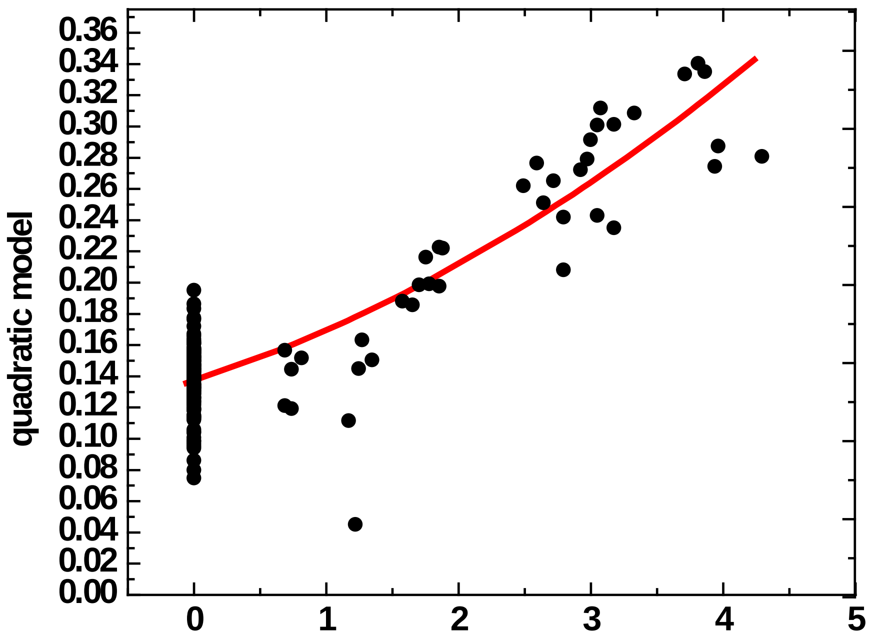

Figure 9. Experimental data with quadratic model fit.

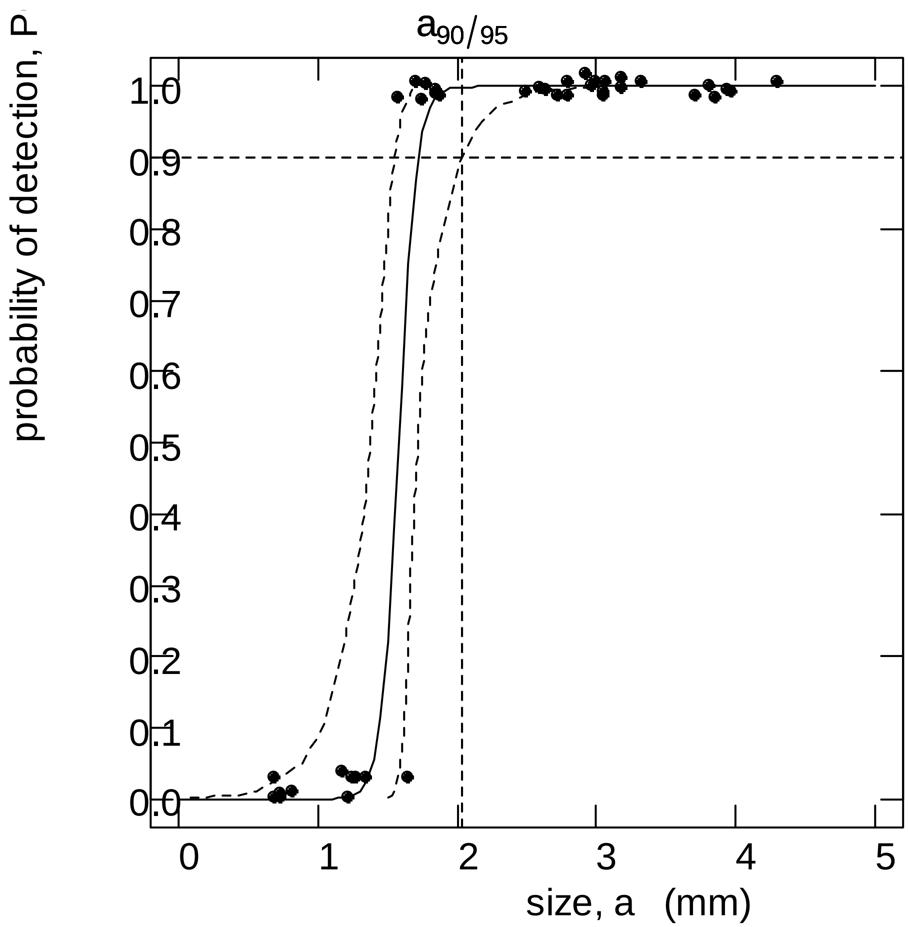

Figure 10: Hit/miss plot with 1 false call.

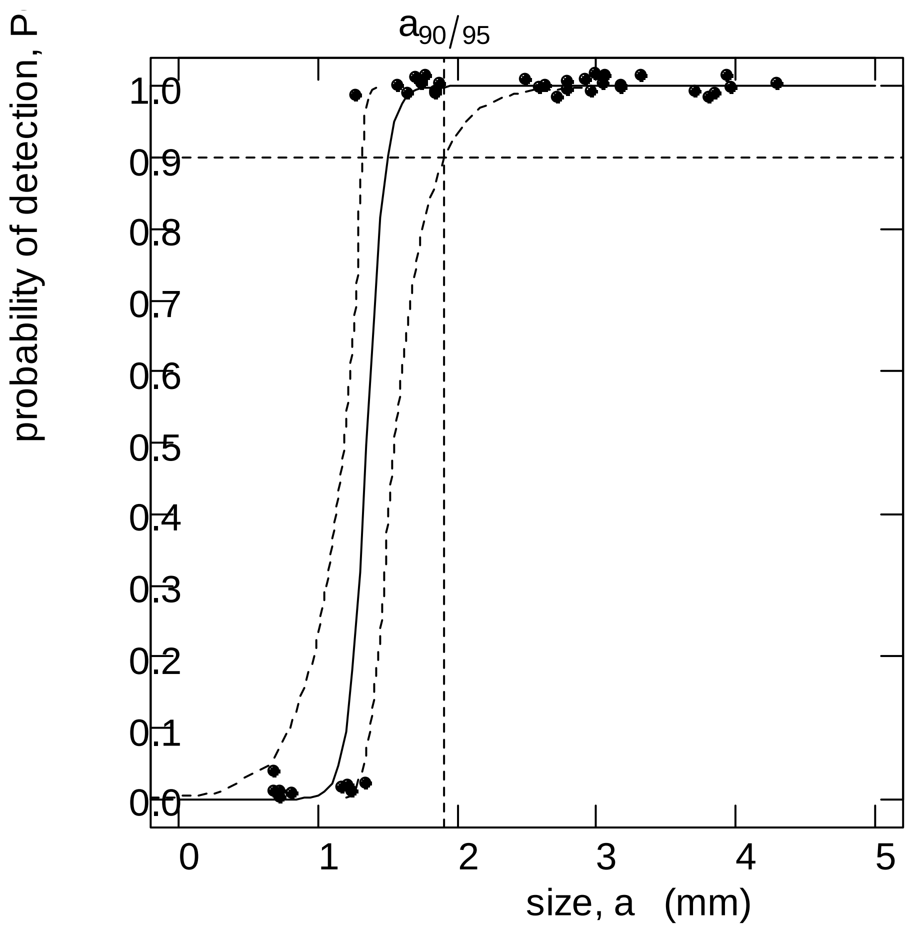

Figure 11. Hit/miss analysis with 11 false calls