USING A BAYESIAN MODEL TO JOINTLY ESTIMATE THE FLAW SIZE DISTRIBUTION AND THE POD FUNCTION

M. Li, W. Q. Meeker, and P. Hovey

Citation: AIP Conf. Proc. **1211**, 687 (2010); doi: 10.1063/1.3362461

View online: http://dx.doi.org/10.1063/1.3362461

View Table of Contents: http://proceedings.aip.org/dbt/dbt.jsp?KEY=APCPCS&Volume=1211&Issue=1

Published by the American Institute of Physics.

USING A BAYESIAN MODEL TO JOINTLY ESTIMATE THE FLAW SIZE DISTRIBUTION AND THE POD FUNCTION

M. Li[1], W. Q. Meeker,[1] and P. Hovey[2]

[1]Center for Nondestructive Evaluation and Department of Statistics, Iowa State University, Ames, IA 50011
[2]Department of Mathematics, University of Dayton, Dayton, OH 45469

ABSTRACT. In this paper we extend previous work by the authors to jointly estimate the flaw size distribution and the POD function from simulated field inspection data. Similar to our previous work, we assume that when a crack is above a detection threshold, both the signal amplitude and the flaw size are recorded. For a signal that is above the noise floor, but below the detection threshold, only the amplitude is recorded. At all other locations we know only that the signal is below the noise floor, i.e. left censored. Now our model allows different airplanes to have different crack growth rates, and the distribution of crack growth rates is to be estimated from the data. To estimate the parameters of the model, we use a Bayesian formulation that provides a convenient structure for estimating the plane-to-plane differences. The Bayesian formulation also allows the use of prior information based on knowledge of physics or previous experience with similar inspection situations. For example, there may be useful information about crack growth rates and about the slope in the amplitude and crack size relationship. Use of such information can importantly improve estimation precision.

Keywords: POD, Bayesian, Bivariate Lognormal Distribution, Censored Data, Reliability
PACS: 43.60.UV, 43.60.Cg, 81.70.Cv, 02.50.Sk

INTRODUCTION

Background

Nondestructive evaluation (NDE) methods are widely used in many industries, such as aerospace applications, to detect flaws or cracks in structures by non-intrusive physical measurements. There are two types of research at NDE community: carefully designed laboratory experiments and field studies. Carefully designed laboratory experiments provide flexibility to study the effect of particular experimental factors. Laboratory experiments are usually based on artificial cracks or other flaws in test specimens. The measurement response is usually modeled as a function of crack size and this model is used to estimate the probability of detection (POD). This is described in detail in [1]. The laboratory studies are usually for new detection methods exploration and validation purpose. After a detection method is developed, tested in the laboratory, and put into use, there is often a desire or a need to do a field study to assess performance with real

CP1211, *Review of Quantitative Nondestructive Evaluation Vol.29,*
edited by D. O. Thompson and D. E. Chimenti

applications. There are significant differences between those two types of studies and various statistical methods can be adopted, depending on the situation.

Motivation and Overview

Many of the statistical methods developed in NDE research are based on complete laboratory study dataset, while features presented at field study data sets such as censoring, truncation, and missing information make the analysis more complicated. In a field study the subjects to be inspected are real parts used in applications such as airplane engine fan blades and rivet holes, and the purpose of the inspection is to determine whether there is a crack in the parts and what is the approximate size of the crack. For a particular inspection method, there is a detection threshold, often based on laboratory experience, model-based theory and operator's experience. If the measurement for an inspected part is below the detection threshold, that part will be classified as no flaw or crack or no risk for future operation. When the measurement response is below the detection threshold, sometimes there is no measurement response value, resulting in left censored data. There will be no crack size information available even thought there are possible small cracks inside those parts. For parts with measurements above the detection threshold, a crack existence decision will be made and those parts will be repaired or removed from future services.

Last year, in [2], we introduced the bivariate lognormal maximum likelihood method to jointly estimate the crack size distribution and POD based on a set of simulated airplane lap-splice rivet holes field data with only one fixed crack growth rate. In this paper, we extend the bivariate lognormal joint estimation method with a Bayesian approach to deal with a more complicated situation with different crack growth rates for different airplanes.

STATISTICAL MODEL SETUP

Bivariate Normal Distribution

The multivariate normal distribution is widely used to model the joint distribution of more than two random variables. The multivariate normal distribution has nice mathematical properties, described, for example, in [3]. The bivariate normal distribution is a special case of multivariate normal with dimension two. The density function of bivariate normal is:

$$f(x_1, x_2; \mu_1, \mu_2, \sigma_{11}, \sigma_{22}, \rho) = \frac{1}{2\pi\sqrt{\sigma_{11}\sigma_{22}}\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2}Q\right) \text{ with}$$

$$Q = \frac{1}{1-\rho^2}\left[\frac{(x_1-\mu_1)^2}{\sigma_{11}} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sqrt{\sigma_{11}\sigma_{22}}} + \frac{(x_2-\mu_2)^2}{\sigma_{22}}\right]. \tag{1}$$

An important property of the bivariate normal distribution used in this paper is that the conditional distribution of one of the random variables, conditional on a fixed value of the other random variable, is a univariate normal distribution. For example, conditional on a fixed value of $X_1 = x_1$, the distribution of X_2 is normal with $\mu_{X_2|X_1=x_1} = \beta_0 + \beta_1 x_1$ and $\sigma^2_{X_2|X_1=x_1} = \sigma_{22}(1-\rho^2)$ where $\beta_1 = \sigma_{12}/\sigma_{11}$ and $\beta_0 = \mu_2 - \beta_1\mu_1$.

Introduction to Bayesian Method

Likelihood based methods are well developed and are familiar to many researchers outside the statistics community. There are many likelihood based software packages including one described in [1] available for POD analysis. Recently, due to the development of faster computer hardware and more efficient statistical software algorithms, Bayesian methods are becoming increasingly popular. The free Bayesian software WinBUGs [4] has been an important factor in this increase of use.

In this paper we will not provide theoretical details of Bayesian simulation methods such as the Markov Chain Monte Carlo (MCMC) algorithm, but instead focus on the application and advantages of Bayesian methods. One important advantage of Bayesian methods is it allows integration of prior knowledge about the inspection procedures and the information about the specimens through the use of prior probability distributions. Then the prior information and the inspection data are combined with Bayes' theorem to provide estimates of the relationship between signal response and crack size and POD in the form of a joint posterior distribution. One other advantage of Bayesian methods is the flexibility of model setup and the kinds of data that can be used. For example, it is easy (relative to trying to do the same with maximum likelihood methods) to use WinBUGs with random effects models and censored data. The credible intervals in Bayesian methods are straight forward and easy to obtain from the posterior distributions while the confidence intervals in likelihood or least square based methods are difficult to obtain especially for a high dimension parameter space. When we use diffuse (flat or approximately non-informative) prior distributions, the results from Bayesian estimation methods will be very close to what one would obtain using likelihood-based approaches.

Bayesian Model for Crack Growth

We simulate field inspection data describing cracks growing out of airplane lap-splice rivet holes. We assume that all holes have an initial crack size that is too small to detect. The cracks grow deterministically as a function of the airplane's service time: $a(t) = a_0 \exp(\lambda t)$ and λ is the unknown crack growth rate. For a fleet of airplanes, different airplanes may have different λ and we assume that λ follows a normal distribution with mean μ_λ and variance σ_λ^2; and that the initial crack sizes follow a log-normal distribution $\log_{10}(a_0) \sim N(\mu_{a0}, \sigma_{a0}^2)$ where μ_{a0} and σ_{a0}^2 are the mean and variance, respectively of the normal distribution. Periodically, such as every 1000 hours of service, an inspection will be made at each rivet hole, and the measurement responses of these measurements will follow a simple linear relationship: $\log_{10}(\hat{a}(t)) = \beta_0 + \beta_1 \log_{10}(a(t)) + \varepsilon$ where $\varepsilon \sim N(0, \sigma_{\hat{a}}^2)$ is the measurement error. Similar to the simpler case of equivalence between the bivariate normal distribution and simple linear regression without crack growth (i.e. a single inspection time), we can model the above crack growth model through a bivariate joint distribution of crack size and measurement response through:

$$\begin{pmatrix} \log_{10}(\hat{a}_{ij}(t_k)) \\ \log_{10}(a_{ij}(t_k)) \end{pmatrix} \sim BVN\left(\begin{pmatrix} \mu_{\hat{a}_j}(t_k) \\ \mu_{a_j}(t_k) \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} \right) = BVN\left(\begin{pmatrix} \beta_0 + \beta_1(\mu_{a0} + \lambda_j t_k) \\ \mu_{a0} + \lambda_j t_k \end{pmatrix}, \begin{pmatrix} \sigma_{\hat{a}}^2 + \beta_1^2 \sigma_{a0}^2 & \beta_1 \sigma_{a0}^2 \\ \beta_1 \sigma_{a0}^2 & \sigma_{a0}^2 \end{pmatrix} \right) \quad (2)$$

where n_1 is the number of rivet holes in each airplane, n_2 is the number of airplanes, n_3 is the number of inspection times with $i = 1,\ldots,n_1$, $j = 1,\ldots,n_2$, $k = 1,\ldots,n_3$, and $\lambda_j \sim N\left(\mu_\lambda, \sigma_\lambda^2\right)$. The relationship between the bivariate normal distribution parameters and the crack growth model parameters can be expressed as $\beta_1 = \sigma_{12} / \sigma_{22}$, $\sigma_{a0}^2 = \sigma_{22}$ and $\sigma_{\hat{a}}^2 = \sigma_{11} - \sigma_{12}^2 / \sigma_{22}$.

Prior Distributions for Crack Growth Model

For computational reasons, the most efficient prior distributions for MCMC simulation are conjugate. Conjugate priors, however, usually impose certain constraints, such as parameter independence, that may not be realistic. For example, in our bivariate normal crack growth model, the mean vector and the variance-covariance matrix elements are not independent. The parameter β_1 appears in both the mean vector and variance-covariance matrix and it is no longer possible to use conjugate priors in MCMC simulations. We have to put an individual prior distribution on each of the model parameters including μ_{a0}, σ_{a0}^2, β_0, β_1, $\sigma_{\hat{a}}^2$, μ_λ, and σ_λ^2. To make our results comparable with what we would obtain using likelihood-based methods, we use diffuse prior distributions for all seven parameters. In particular, μ_{a0}, β_0, β_1, and μ_λ have normal prior distributions $N(0,1000)$ while σ_{a0}^2, $\sigma_{\hat{a}}^2$, and σ_λ^2 have Gamma prior distributions $G(0.001, 0.0001)$.

SIMULATED AIRPLANE FIELD DATASET

At present we do not have access to the real data corresponding to our model and setup. Therefore we use simulated rivet hole data to illustrate the use of the model and

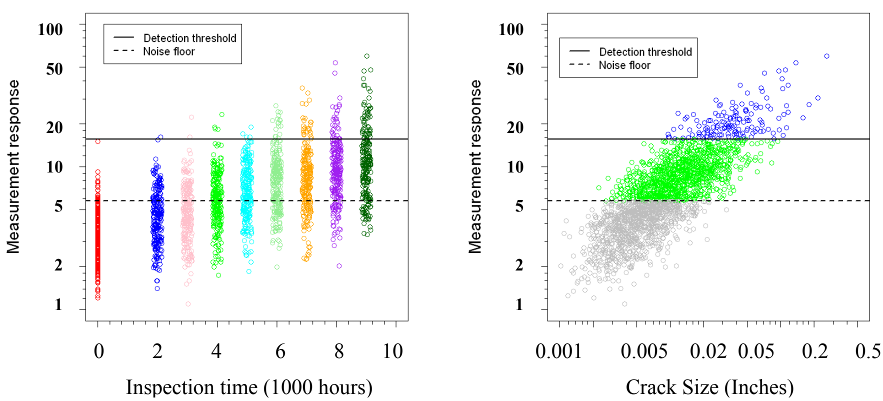

FIGURE 1. Simulated airplane engine fan blades field data set: the measurement responses as function of inspection times (left) and measurement responses as function of crack sizes (right). The horizontal solid line is the pre-set detection threshold and the horizontal dashed line is the noise floor.

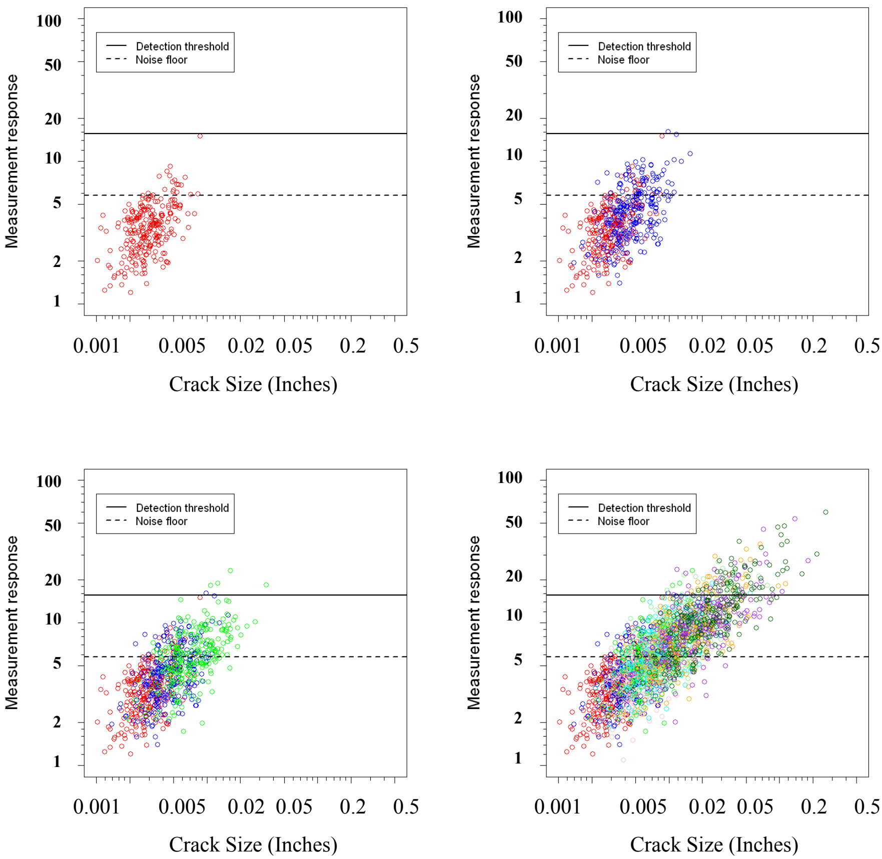

FIGURE 2. Simulated rivet-hole field data with the trend of crack growth for inspections at 0 hours (top left), 2000 hours (top right), 4000 hours (bottom left), and all inspection times (bottom right).

Bayesian estimation method. Measurements are taken on each hole at each inspection time. In actual applications, once a crack signal is above the threshold, a repair is effected, and the size of the crack is determined during the repair process, but then no further data is obtained from that location.

The simulated data are illustrated at Figure 1. The left figure shows measurement responses for $n_3 = 9$ inspection times (0, 2, 3, … , 9 thousands hours) in service; the right figure shows the relationship between the measurement responses and the crack sizes. The detection threshold and noise floor are also indicated in Figure 1 by horizontal solid and dashed lines, respectively. There are $n_1 = 5$ fan blades sampled from each airplane and $n_2 = 50$ airplanes in the fleet. Each airplane has a unique crack growth rate λ_j which is sampled from a normal distribution $N(0.216, 0.0050)$. In our $\hat{a}$ versus a model, the log measurement responses were simulated as a linear function of log crack sizes with intercept $\beta_0 = 4.5$ and slope $\beta_1 = 0.56$ and with a random measurement error from a normal distribution $N(0, 0.13)$. The logarithms of the initial crack sizes were sampled from $N(-5.81, 0.147)$. The crack growth patterns are shown in Figure 2 with different colors representing measurement data from different inspection times.

We apply the following procedure to the original simulated data set. First for any specimens with measurement response below the detection threshold, we will assume the crack is small enough to put the specimen back into service without risk. So there will be no crack size information for those specimens with measurement below detection threshold. Second, we have a noise floor to indicate that the minimum meaningful readings are limited by the inspection instruments and any response below the noise floor is considered to be left censored. That is, our inspection instrument can only provide precise readings above the noise floor and any reading below the noise floor are only known to be below the noise floor (left censored). Third, at any inspection time, if a specimen has a measurement response above the detection threshold, the specimen will be removed from future service and the crack size information will be obtained during the repair. Because the specimen will be removed from services and there will be no further future inspections for that location. Thus there is at most one crack size reading for each specimen. The simulation procedure is illustrated in Figure 3. The plot on the left shows simulated measurements at all inspection time, even after cracks would be repaired. The plot on the right shows that actual structure of data that would be observed in actual applications.

BAYESIAN POD ANALYSIS

Analysis of the simulated field inspection data using the Bayesian method requires complicated statistical modeling because we have left censored data, missing crack size information, a random effect for crack growth rates, and incomplete longitudinal inspections. Also, we would like to have quantification of the statistical uncertainty in the relationship between measurement responses and crack sizes and POD as a function of crack size. The traditional likelihood based methods are difficult to formulate especially for the confidence bounds and there is no commercially available software that can handle

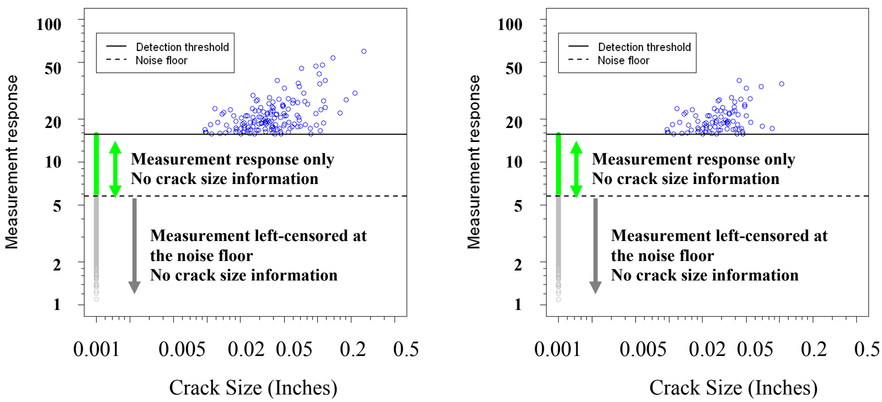

FIGURE 3. In both plots we see that there is no crack size information below the detection threshold and only left-censored inspection measurement for observations below the noise floor. The left plot shows the data before any specimens' replacement, while the right plot shows the final data with specimens removed after measurement response above the detection threshold.

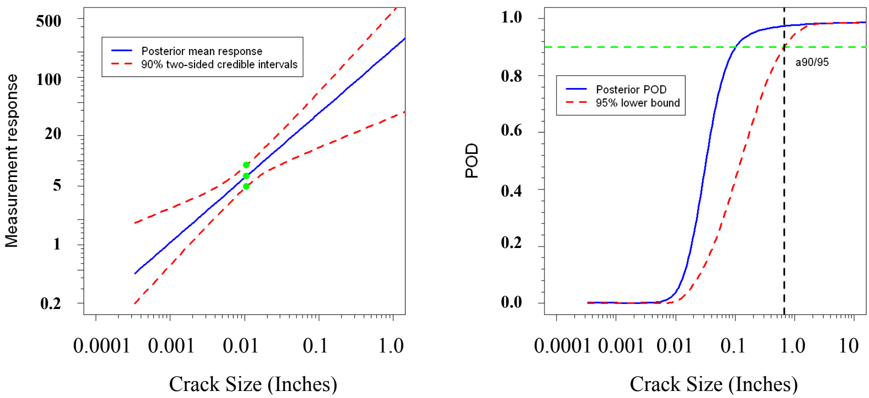

FIGURE 4. The relationship between measurement responses and crack sizes with 90% credible bounds (left) and the POD with 95% lower credible bounds (right) from posterior distributions.

the complicated data and random effects. We can, however, formulate this problem relatively easily in the Bayesian framework and let WinBUGs do the computations. WinBUGs also provides tools for computing the credible bounds from the output of the MCMC simulation. Because we use non-conjugate priors, the computation time for each MCMC step is longer than when using conjugate prior distributions and more steps are needed for MCMC to converge. The total simulation time is around 5 hours for our dataset using an Intel® Dual Core 3.2GHz computer. The relationship between measurement responses and crack sizes with 90% credible bounds and the POD with 95% lower credible bound from posterior distributions are shown at Figure 4. Finally the WinBUGs simulation results are compared between using the full simulated data set (i.e. all the measurement responses and crack sizes are exact numbers) and using the processed final simulated field inspection data sets. The results are shown at Table 1 with the true parameters used in the data simulation. The parameter estimates from the full data set are very close to the true values when all data responses and crack sizes are available.

TABLE 1. Comparison of model parameters estimate between WinBUGs simulation using full unprocessed simulated data set and processed field inspection data set with the true parameter value used in the data generation.

Model	Parameter	True value	Full data	Field data
$\log\left(a_{0,ij}\right) \sim \mathrm{N}\left(\mu_{a0}, \sigma_{a0}^2\right)$	μ_{a0}	-5.81	-5.812	-5.359
	$\sigma_{a_0}^2$	0.147	0.144	0.194
$a_{ij}[t_{ijk}] = a_{0,ij} + \lambda_j t_{ijk}$ $\lambda_j \sim \mathrm{N}\left(\mu_{\lambda 0}, \sigma_{\lambda 0}^2\right)$	μ_{λ_0}	0.216	0.216	0.099
	$\sigma_{\lambda 0}^2$	0.0050	0.0059	0.0032
$\log(\hat{a}_{ij}[t_{ijk}]) =$ $\beta_0 + \beta_1 \log(a_{ij}[t_{ijk}]) + \varepsilon_{ijk}$ $\varepsilon_{ijk} \sim \mathrm{N}\left(0, \sigma_{\hat{a}}^2\right)$	β_0	4.500	4.513	4.499
	β_1	0.560	0.546	0.576
	$\sigma_{\hat{a}}^2$	0.130	0.131	0.155

With only fraction of the total information, the parameter estimates from the simulated field data are still close to the true values except for the crack growth rate distribution parameters $\mu_{\lambda 0}$ and $\sigma^2_{\lambda 0}$. The bias in the estimates of $\mu_{\lambda 0}$ and $\sigma^2_{\lambda 0}$ is caused by a sampling issue similar to the well known length-biased sampling problem (e.g., [5]) because we are likely to get more crack-growth rate information on cracks that are growing slower. Further investigations are needed to address this problem with a better model.

CONCLUSIONS

In this paper we showed how to extend the bivariate joint estimation idea to a more complex real world situation through a simulated nondestructive evaluation field inspection data set. The Bayesian's approach was used to handle the more complicated data structure through the WinBUGs software package. This general approach provides an efficient and versatile alternative method to likelihood-based methods. The Bayesian method provides credible intervals for specified quantities of interest. These intervals can be obtained by using the marginal posterior distributions of these quantities. Because of its versatility, we believe Bayesian methods will have more impact in future nondestructive evaluation research. Future work will include solving the length-bias issue and integrating the noise interference model [6] into the model, and applying the results to real data.

ACKNOWLEDGEMENTS

This material is based upon work supported by the Federal Aviation Administration under Contract #DTFA03-98-D-00008, Delivery Order #0058 and performed at Iowa State University's Center for NDE as part of the Center for Aviation Systems Reliability program.

REFERENCES

1. MIL-HDBK-1823 (2007), *Nondestructive Evaluation System Reliability Assessment*, Standardization Order Desk, 700 Roberts Avenue, Philadelphia, PA 19111.
2. P. Hovey, W. Q. Meeker, and M. Li, "Joint estimation of the flaw size distribution and POD function" in *Review of Progress in QNDE,* 28B, edited by D. O. Thompson and D. E. Chimenti, AIP Conference Proceedings Vol. **1096**, American Institute of Physics, Melville, NY (2009), pp. 1832-1839.
3. A. R. Johnson, and D. W. Wichern, *Applied Multivariate Statistical Analysis 4th Edition*, Prentice Hall, New Jersey (1998).
4. D.J. Lunn, A. Thomas, N. Best, and D. Spiegelhalter, "WinBUGS - a Bayesian modeling framework: concepts, structure, and extensibility", *Statistics and Computing*, **10**:325-337 (2000)
5. R. Scheaffer, "Sized-biased Sampling", *Technometrics*, Vol. **14**, pp. 635-& (1972)
6. M. Li and W. Q. Meeker, "A noise interference model for estimating probability of detection for nondestructive evaluations", in *Review of Progress in QNDE,* 28B, edited by D. O. Thompson and D. E. Chimenti, AIP Conference Proceedings Vol. **1096**, American Institute of Physics, Melville, NY (2009), pp. 1769-1776.