INTRODUCTION

In comparing different eddy-current systems (or probes), it is convenient to base the comparison on the magnitude of the signal produced by a certain flaw size and type. However, as in any detection problem, the real effectiveness of such a system in detecting a given flaw is determined by the statistical variations in the flaw signal and the characteristics of the noise added by the system. Hence, the objective of this task has been to develop a statistical model for determining the probability of detecting a given flaw using an eddy-current system. Such a model would be useful not only for comparing different systems, but also for optimizing the detection process. This optimization would consist of using the model to select the probe and operating frequency(ies) that maximize the signal-to-noise ratio (SNR).

In work conducted last year, the basic elements of such a model were developed for an air-core coil probe undergoing vertical lift-off with an output containing Gaussian electronic noise. In addition, the coil radius was assumed to be several times larger than the flaw (e.g., a surface crack). Several improvements to the model were made during the past year. These improvements included adding the capability to utilize measured noise data, and using Monte Carlo simulation, to compute probability density functions. This paper presents a detailed analysis of some measured noise data, and the results of using those data with a probe-flaw interaction model to compute the surface-crack detection characteristics of two different air-core coil probes.
The basis for developing a statistical detection model for an eddy-current system is a measurement model that relates the output voltage of the system, \( V(q, f) \), to its various signal and noise components. We considered only an absolute probe and used an additive measurement model as defined by

\[
V(q, f) = V(f) + V(f, lo, m, r) + V(lo, m, r) + V(n) .
\]  

Here, \( V(f) \) is the voltage produced by the flaw in the absence of probe motion (lift-off, denoted by \( lo \)), material variations (\( m \)), and surface roughness (\( r \)), while \( V(f, lo, m, r) \) is the variation in flaw voltage produced by the combined action of these effects. The sum of these two voltages is the actual voltage produced by a flaw. The quantity \( V(lo, m, r) \) is the voltage produced by probe motion, material variations, and surface roughness in the absence of a flaw, and \( V(n) \) is an electronic noise voltage. When there is no flaw present, the output voltage, \( V(q) \), is composed of only the last two (statistically independent) terms.

To compute the probability of detection (POD) and the probability of false rejection (PFR), one needs to know the probability distributions for each term in Eq. (1) so that the probability densities for \( V(q) \) and \( V(q, f) \) can be determined by Monte Carlo simulation. These distributions can be found using theoretical models, measured data, or a combination of both.

The theoretical probe-flaw interaction models currently available or under development are reviewed by Auld, et al. The validity of these models for surface cracks depends on the ratio of crack depth to skin depth, and on the crack shape. Furthermore, they assume a smooth, flat surface with no material variations. Models that permit arbitrary probe motion are not yet fully developed, so at present one is limited to one-dimensional lift-off motion perpendicular to the surface. Subject to these limitations, one can use these models to compute the probability distributions for \( V(f) \), \( V(f, lo) \), and \( V(lo) \), given the distributions for \( f \) and \( lo \).

Although we can compute \( V(lo) \) for some simple cases, the background "noise," \( V(q) \), is not easy to model because it is very system-dependent. Among other things, it depends on the electronics, scanning mechanism, and type of probe. Thus, the probability distribution of this quantity should be determined by measurement. This straightforward measurement consists of building up a histogram by sampling the voltage at many points along a scan or scans in an unflawed region of a sample having the same geometry as the piece to be inspected.
The distribution of \( V(q,f) \) can also be measured by sampling the peak signals produced by a suitably large number of flaws distributed along a scan. However, it may be impractical to obtain a sample containing such a flaw distribution. Usually, one has a sample containing only a few flaws, and thus one is limited to measuring the peak signal produced by a particular flaw averaged over several scans. Given the difference between this mean flaw signal—or \( V(f) \) as computed using a theoretical model—and the mean of the background signal one can approximate the probability density of \( V(q,f) \) by simply shifting the measured density for \( V(q) \) by this difference of means. This procedure neglects \( V(f,lo,m,r) \) altogether, and also any effect the distribution of \( V(f) \) might have on the shape of the distribution for \( V(q,f) \). However, this is the procedure that was used in the computations presented in this paper.

### STATISTICS OF THE MEASURED BACKGROUND VOLTAGE

The analysis described here is based on 11 sets of background-voltage measurements obtained under various operating conditions. The measurements were made at Martin Marietta using an automated X-Y scanner. The data were taken by using an eddy-current probe to scan an unflawed section of a metal sample (e.g., Ti-6-4) and recording the voltage reading at equally spaced positions along the scan. Each data set consisted of 150 samples.

In all cases, the distribution of the measured background voltage was skewed and had a heavy upper tail; the histogram shown in Fig. 1 for one of the cases is a typical example. The figure also shows that the experimental distribution is fitted by a Gamma probability density function (pdf), as suggested by Kincaid.\(^3\)

Although it is possible to analyze the system's performance using the Gamma pdf, the analysis was simplified considerably by transforming the data so that a Gaussian pdf could be used. We found that a square root transformation achieves the desired result, as is shown in Fig. 2, which displays a histogram of the transformed background voltage and the fitted Gaussian pdf for the same data shown in Fig. 1. Both the Gamma and Gaussian fits were statistically significant at better than the 0.05 level for all data sets.

The Gaussian pdf is completely defined by two parameters: the mean (\( M \)) and standard deviation (\( S \)). These two parameters provide all the information needed to analyze the system's detection performance. Specifically, the POD and the PFR are functions of the SNR, which is itself a function of \( M \) and \( S \), as will be shown in the next section.
CALIBRATION OF PROBE-FLAW INTERACTION MODEL AND SNR DEFINITION

Two types of air-core probes were tested and modeled: one was designed to operate at 10 kHz and the other at 200 kHz. The coil dimensions and other probe parameters are shown in Fig. 3.

A probe-flaw interaction model was used to calculate the flaw signal produced by each probe for a variety of flaw sizes. The model output will be denoted by \( V_{fm} \); it corresponds to \( V(f) \) in Eq. (1). Because the transfer function for the eddy-current measurement system at Martin Marietta is unknown, the model output must be calibrated to relate it to the measured voltage produced by the probe for the same flaw size. During the calibration procedure, \( V_{fm} \) was computed for a single specific flaw size for each probe. Then the probe was used to measure the background voltage as well as the peak voltage produced by the known flaw. Using the

*Calculations using a nonuniform-illumination model were performed at Stanford University by S. Ayter and M. Riaziat.
measurements, estimates were obtained for the mean background voltage, $M_b$, and the mean peak output voltage, $M_p$. Assuming that $M_p = M_f + M_b$, where $M_f$ = mean flaw-induced voltage, the calibration factor was obtained from the ratio $(M_p - M_b)/V_{fm} = M_f/V_{fm}$. Then, regardless of the flaw size used in the modeling calculations, all model estimates of the flaw-induced voltage were multiplied by the same calibration factor to obtain an approximation of $M_f$.

The SNR is the critical factor that determines the system's flaw-detection capabilities. For our purposes, SNR is defined by

$$\text{SNR} = \frac{(M_p - M_b)/S_b}{M_f/S_b}^2 = \frac{[M_f/S_b]^2}{10 \log(\text{SNR})}.$$  \hspace{1cm} (2)

In the text, SNR values are given in decibels, which is equal to 10 log(SNR). To calculate SNR, $M_b$ and $S_b$ are estimated from the measurements of background voltage after performing a square root transformation. For the flaw-induced signal, $M_p$ can be estimated from measurements, or $M_f$ can be obtained from the calibrated model output. Using $S_b$ in Eq. (2) assumes that the pdf of $V(q,f)$ has the same shape as that of $V(q)$.  

Fig. 2. Experimental Distribution and Fitted Gaussian Probability Density Function of the Square Root of the Background Voltage for a 10-kHz Air-Core Probe
RESULTS

Operating Characteristic

The operating characteristic (OC) of the flaw-detection system consists of a graph that shows the POD that can be achieved for a given PFR, and vice versa. Fig. 4 shows two pdfs obtained using the simulation model to estimate $M_f$ for a crack depth of 0.040 in., and the measured background voltage to define $M_b$ and $S_b$; the SNR is 11.6 dB. The separation and spread of the two pdfs determine the POD and PFR for each threshold. Fig. 5 shows the OC associated with Fig. 4; the staircase appearance of the extreme left of the OC is caused by the numerical integration procedure used to calculate POD and PFR. In Fig. 5, low values of threshold voltage yield high POD and PFR, and vice versa. The figure shows the fundamental limitations faced by the probe for the given crack depth. For example, a POD of 90% has an associated PFR of 1%, which cannot be changed. Thus, the OC provides a complete description of the probe's detection capabilities for a given material, crack size, and operating conditions.
Fig. 4. Probability Density Function for 10-kHz Air-Core Probe at 100 kHz. (Crack depth = 0.040 in., Ti-6-4 Material, SNR = 11.6 dB.)

Fig. 5. Operating Characteristic of 10-kHz Probe at 100 kHz. (Crack depth = 0.040 in., Ti-6-4 Material, SNR = 11.6 dB.)
Comparison of Air-Core Probes

The detection performance of the two probes was compared for various crack depths and a fixed PFIR of 0.10%. Fig. 6 plots POD as a function of crack depth for the two probes. The point labeled "experiment" corresponds to measured data, the values for the other points were estimated using the simulation model. The figure shows that the 10-kHz probe performs well only for crack depths greater than 40 mils. By contrast, the 200-kHz probe shows nearly perfect detection for crack depths as small as 30 mils; its performance for depths smaller than 30 mils was not explored because model data were not available at the time. It is clear from Fig. 6 that the 200-kHz probe outperforms the 10-kHz probe.

Another view of the relative performance of the two probes is shown in Fig. 7, which plots the SNR in decibels as a function of crack depth. Both probes have equivalent SNRs for the largest depth. However, the superior performance of the 200-kHz probe is evidenced by its ability to retain a high SNR of about 15 dB for the smaller crack depths, whereas the SNR of the 10-kHz probe decreases rapidly with decreasing crack depth.

Fig. 7 suggests that the SNR is the most important figure of merit that describes the relative performance of the probes. Additional evidence for this conclusion is shown in Figs. 8 and 9, which show the no-flaw and flaw pdfs of the two probes for the same flaw size. The figures are based on measured rather than model-derived data. The figures show that the two probes have different signal and noise characteristics. For example, the 10-kHz probe (Fig. 8) has lower mean background and higher mean signal-plus-background voltage than the 200-kHz probe (Fig. 9). However, the latter is less noisy, as indicated by the narrowness of its pdf. The net result is that for this flaw size, both probes have an equivalent SNR.

DISCUSSION AND CONCLUSIONS

The analysis has shown that a statistical detection model provides a meaningful basis for comparing eddy-current systems and/or probes. Together, the OC and SNR completely describe the system's capabilities, and any comparisons should be based on these characteristics. It follows that sensitivity is an incomplete criterion for comparing performance.

We used a probe-flaw interaction model successfully to investigate the ability of different air-core coil probes to detect small flaws. Such models are essential for extrapolating the statistical calculations to small flaws and for identifying the factors that govern probe performance.
The analysis of measured background voltage showed that the raw voltage has a Gamma pdf and that the square root of the voltage has a Gaussian pdf. Hence, the SNR in the square root domain can be used as a unique measure of probe effectiveness.

The detection criterion used in this work consisted of a simple threshold exceedance using square-root-transformed voltage output. Thus, no attempt was made to process the system output to enhance
Fig. 7. Comparison of Signal-to-Noise Ratio for Two Air-Core Probes

SNR. Future work should be done to develop signal-processing algorithms for improving SNR, and hence detection performance. For example, previous experience suggests that band-passing the output voltage should enhance the SNR.

ACKNOWLEDGMENTS

The authors wish to thank Messrs. W. Rummel, R. Schaller, R. Muthart, and F. Ross at Martin Marietta in Denver, Colorado for providing the eddy-current probes, the flawed samples, and the measured data. This work was sponsored by the Center for Advanced Nondestructive Evaluation, operated by Ames Laboratory, USDOE, for the Defense Advanced Research Projects Agency and the U.S. Air Force Materials Laboratory under Contract W-7405-ENG-82.
Fig. 8. Probability Density Functions for 10-kHz Probe.
(Crack depth = 0.052 in., Ti-6-4 Material, SNR = 19 dB.)

Fig. 9. Probability Density Functions for 200-kHz Probe.
(Crack depth = 0.052 in., Ti-6-4 Material, SNR = 19 dB.)
REFERENCES


3. T. G. Kincaid, College of Engineering, Boston University, Boston, Massachusetts, private communication (June 1983).


DISCUSSION

G. Birnbaum (National Bureau of Standards): I had a question for Al. Is it really safe to consider building filters with regard to small signal detection in the presence of a large amount of noise without really understanding the purpose of taking the square root of the voltage to get a Gaussian distribution? For example, the results could be an artifact of a particular experimental setting. So it seems to me that that point needs to be settled before one can safely march on.

A.J. Bahr: That's true. However, there are some algorithms that are not as sensitive to the noise distribution as others. That somewhat fits, again, along those lines, but I think more work definitely is required to establish the general characteristics of the eddy current noise, understand it. What I shared was meant to suggest that we have some methodology that may help us understand that.