International Journal of Applied Electromagnetics and Mechanics 15 (2001/2002) 39–46
IOS Press

Dynamic thresholding for automated analysis of bobbin probe eddy current data

H. Shekhar*, R. Polikar, P. Ramuhalli, X. Liu, M. Das, L. Udpa and S.S. Udpa
Materials Assessment Research Group, Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011, USA

Abstract. An automated algorithm is presented for the analysis and classification of eddy current bobbin probe data obtained from nuclear power plant steam generator tubes (SGT). This algorithm attempts to find a balance between the two seemingly conflicting requirements of SGT eddy current data analysis, namely, detecting and identifying as many of the actual defects as possible, whilst limiting the number of false alarms to a minimum. Initial results presented in this paper look very promising.

1. Introduction

Steam generator tubes (SGT) in nuclear power plants are routinely inspected using eddy current techniques. Although a variety of probes are employed for inspecting steam generator tubes, the most commonly employed method involves the use of multifrequency differential bobbin type probes. Such probes are very simple and rugged. The structural integrity of the tube is assessed by analyzing the impedance changes of the probe as it travels within the tube. The inspection is usually carried out using four different frequencies to extract additional information concerning the state of the tube. Unfortunately the probe is sensitive to the presence of a number of other structures within the steam generator such as tube support plates, conductive deposits, tube sheets, etc. in addition to flaws present in the tube. The multifrequency measurements allow selective suppression of signal artifacts introduced by such benign structures as well as garner additional information concerning the flaws [1]. The inspection of steam generators involves the testing of several hundreds of kilometers of tubing. The data generated by the inspection procedure is very large. Since manual interpretation of the data is extremely cumbersome and time consuming, the industry has considerable interest in automating the analysis procedure. Although a number of automatic analysis systems have been proposed over the years, their performance have fallen short of levels obtained using manual procedures. Typical problems associated with automatic procedures include high false alarm rates and or poor detection/characterization performance.

This paper describes anautomatic procedure for analyzing multifrequency eddy current probe data derived from the inspection of nuclear steam generator tubes. The procedure uses a multi-step method to accomplish the task. The first step involves a filtering process to minimize the noise contained in the signal. The second step involves a two-stage process for identifying segments of the signal stream that represent defects. The third stage involves the classification of signals. Initial results obtained using this approach indicate that it is possible to build an automated system capable of meeting the seemingly conflicting the need for high detection and low false alarm rate.

*Corresponding author: Hemabh Shekhar; E-mail: hemabh@iastate.edu.

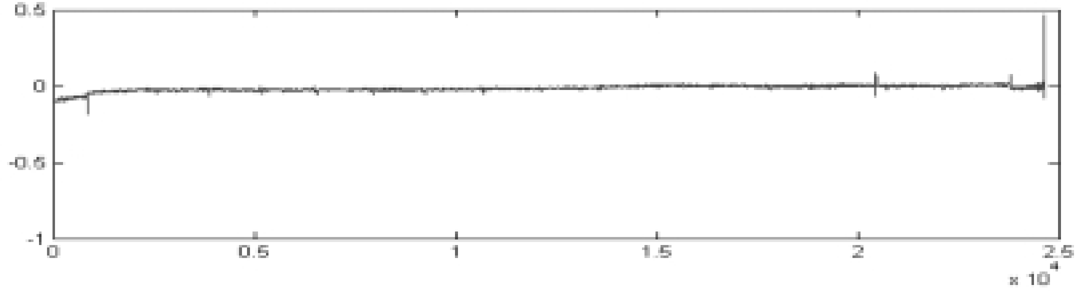

Fig. 1. Signal from a steam generator tube, data length (23450).

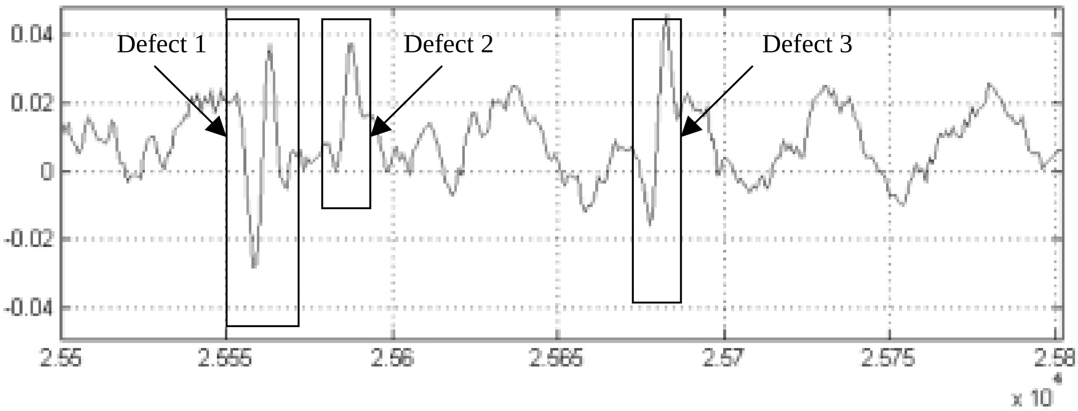

Fig. 2. Defect signals in a tube.

2. Neyman-Pearson (N–P) detector based dynamic thresholding

Eddy current bobbin data is characterized by significant amount of noise, making its analysis a difficult task. Furthermore, most defects are typically less than 0.5” long, which results in very few data points buried in long streams of data generated during the course of inspection of the entire tube. For example, a typical defect signal may contain fewer than 10 data points, if a sampling rate of 1 kHz is employed. The signal stream from the entire tube, in contrast, typically consists of over 25000 data points. Short defect indications obscured in large volumes of noisy data make the bobbin coil eddy current data analysis an extremely laborious and challenging task. Figure 1 shows a typical eddy current signal from a steam generator tube. The figure depicts the imaginary component of the complex impedance in the mix channel, where the 200 kHz and 400 kHz signals have been combined to selectively eliminate the support plate signals. Figure 2 shows an example of defect signals embedded in noise in a section of data.

The figures indicate that a preprocessing scheme that can identify and eliminate a majority of the noisy segments of the signal that do not bear any information would be very beneficial. In addition, the scheme must retain any signals due to defects. This paper describes a scheme to reduce the number of data points to be subsequently analyzed by a classification algorithm. The scheme uses a Neyman-Pearson (N-P) detector to compute a dynamic or adaptive threshold value that separates data points into one of two classes – noise and potential defect indications.

2.1. Neyman-pearson detector theory

A Neyman-Pearson detector maximizes the probability of detecting a signal (in presence of noise) for a given probability of false alarm (PFA). This objective can be accomplished by using a hypothesis testing approach. The hypotheses are given by

$$H_0 : X(n) = W(n) \tag{1a}$$

$$H_1 : X(n) = S(n) + W(n), n = 1, 2, \ldots, N \tag{1b}$$

where, $X(n)$ is the measured signal, $S(n)$ is the defect signal, W(n) is the additive noise and N is length of the measured signal. The $N - P$ detector uses the log likelihood ratio to compute a test statistic that is used for threshold computation. The log-likelihood ratio is specified as

$$T(X) = \log \left[\frac{P[X|H_1]}{P[X|H_0]} \right] \tag{2}$$

where, $P[X|H_1]$ is the PDF of X given H_1 and $P[X|H_0]$ is the PDF of X given H_0. Assuming that the noise is additive white Gaussian, with zero mean and variance σ_w^2, and the signal is distributed as a multivariate Gaussian, with covariance C_s and mean μ_s:

$$W \sim N(0, \sigma_w^2) \tag{3a}$$

$$S \sim N(\mu_s, C_s) \tag{3b}$$

The test statistic can then be formulated using equations (2), (3a) and (3b) as [2]:

$$T(X) = X^T (C_s + \sigma_w^2 I)^{-1} \mu_s + \frac{1}{2\sigma_w^2} X^T \left[C_s (C_s + \sigma_w^2 I)^{-1} \right] X \tag{4}$$

This test statistic is then compared to a threshold (λ to determine which of the two hypotheses is true:

$$\text{If} T(X) \leqslant \lambda, H_0 \text{is true} \tag{5a}$$

$$\text{If} T(X) > \lambda, H_1 \text{is true} \tag{5b}$$

The threshold (λ is obtained using a constraint on the probability of false alarm (PFA) [3]

$$PFA = \int_{\lambda}^{\infty} P[T(X)|H_0] dT \tag{6}$$

Equation (6) uses the $T(X)$ distribution of noise to compute the threshold λ for a given value of PFA.

2.2. Special cases

The test statistic, $T(x)$, can be simplified under the following assumptions:
a. If signal is zero mean, i.e., $\mu_s = 0$, then the test statistic is modified as

$$T(X) = \frac{1}{2\sigma_w^2} X^T \left[C_s(C_s + \sigma_w^2 I)^{-1}\right] X \tag{7}$$

b. If the signal is white with non-zero mean, i.e., $C_s = {\sigma_s}^2 I$, then $T(x)$ is given by

$$T(X) = \frac{X^T N \mu_s}{({\sigma_s}^2 + {\sigma_s}^2)} + \frac{\sigma_s^2}{2\sigma_w^2({\sigma_s}^2 + \sigma_s^2)} X^T X \tag{8}$$

c. If signal is both zero-mean and white [4,5], then

$$T(X) = \frac{\sigma_s^2}{\sigma_w^2(\sigma_s^2 + \sigma_s^2)} X^T X \tag{9}$$

As discussed above, the signal consists mostly of noise points with a few defect signal points. In order to capture the test statistics of a defect signal more accurately, a sliding window of length M given by equation (10), is used.

$$\mathbf{y} = \mathbf{y}(n) = \left[X\left(n - \frac{M}{2} + 1\right), X\left(n - \frac{M}{2} + 2\right), \ldots, X(n), X(n+1), \ldots, X\left(n + \frac{M}{2}\right)\right] \tag{10}$$

Using the signal within the sliding window as the measured signal, the test statistic is given by

$$T[y(n)] = \frac{\sigma_s^2}{\sigma_w^2(\sigma_s^2 + \sigma_s^2)} y^T y \tag{11}$$

The test statistic $(T[y(n)])$ can be computed for each point n. A signal point is labeled as an indication if $T[y(n)]$ is greater than the threshold λ.

2.3. N-P detector implementation

The raw bobbin coil data must first be preprocessed before the N-P detector can be applied. The preprocessing steps filter the data to remove any biases and trends that may be present. The first preprocessing step is to remove the trend and bias from the signal. An N-point discrete cosine transform (DCT) [6] based zero-phase high pass filter is used for this purpose, where N is the signal length. Zero phase filtering is required to ensure that no shifts in the signal are introduced. The cutoff frequency for the high-pass filter was chosen (a trial and error method) to be around 50 Hz.

The model used for this problem assumes that the data comes from a mixture of two zero mean white Gaussian distributions, one representing noise and the other representing both defect signal and noise. Hence, the test statistic for the N-P detector is given by [**?**]. The test statistic requires the computation of (σ_w^2 and (σ_s^2, the variances of the noise and the signal. These are computed from the histogram of the data. A sliding window of length M is used to compute the value of $T[y(n)]$ at each point and the desired probability of false alarm (PFA) is then selected (for example, 0.05). The threshold value, λ, is computed based on this PFA, as the (100-PFA)$^{\mathrm{th}}$ percentile of the distribution of $T[y(n)]$. All data points whose test statistic values are greater than λ are marked as potential defect signal points while all other data points are discarded as noise points.

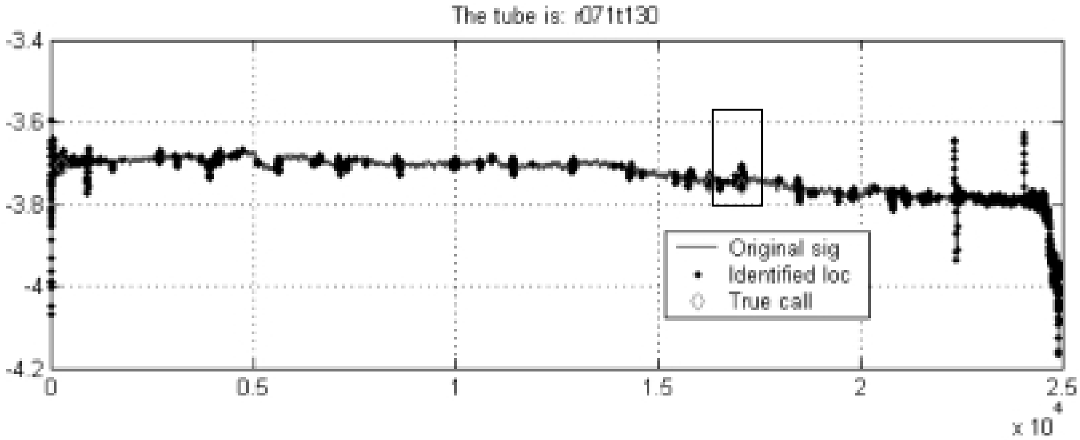

Fig. 3. Bobbin coil eddy current data showing result of dynamic thresholding.

3. Results

The dynamic thresholding algorithm was applied to differential bobbin coil data collected from steam generator tubes. Each coil in the differential arrangement was 0.06" wide and the two coils were separated by a distance of 0.06". The excitation frequencies used were 200 kHz, 400 kHz and 600 kHz. In addition, data was also collected at a low frequency (35 kHz) and a mix channel, combining the data at 200 kHz and 400 kHz, was created to eliminate signals from external support structures, enhancing the signals from any defects that may occur near these structures.

The results of the dynamic thresholding algorithm are shown in Figs 3–8. In all cases, a sliding window of length 5 was used to compute the test statistic. The PFA used in all cases was 0.25. The solid line in Fig. 3 shows the raw bobbin coil eddy current data from a steam generation tube. The data stream consists of a total of 24929 data points. The dots represent the data points that were marked by the algorithm as potential defect signals. The algorithm identified a total of 9053 data points as potential defect indications, thereby achieving a reduction of 63.68%. The diamonds indicate the location of four defects in this data set. Figure 4 is an enlarged view of two defect signals in the tube. The box in Fig. 3 shows the corresponding section on the complete data set.

Similar results are shown in Figs 5 and 6 for a second signal. In this case, the algorithm achieved a total reduction of 61.89% (9049 data points identified as potential defect indications out of a total of 23746 data points in the data set). The algorithm also detected all 4 defects present in the data. Similarly, a third data set showed a reduction of 65.47% (Figs 7 and 8). However, the algorithm missed one defect signal. The enlarged view in Fig. 8 shows that the defect missed had extremely low amplitude.

The above algorithm was applied to data from a total of 64 tubes (which had a total of 149 defects and 8 dents) and the performance of the algorithm is summarized in Table 1. The first column gives the PFA used for analysis and the distribution of tubes with defect and no defects. The second column shows the minimum data reduction achieved over all the defect/non-defect tubes while the third column shows the maximum data reduction achieved over the all the defect/non-defect tubes. These results indicate that the dynamic thresholding algorithm achieves an average data reduction of around 60% over all the tubes. This is a significant reduction in the amount of data that needs to be analyzed by a subsequent classification algorithm. In addition, the table indicates that the algorithm was able to detect almost all the defects and dents present in the dataset. In fact, only one defect and one dent were missed in a total of two tubes.

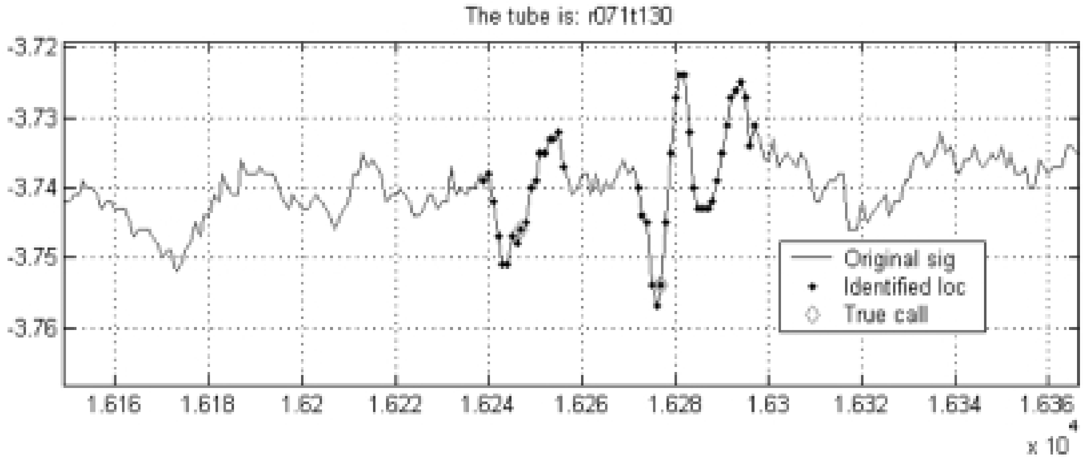

Fig. 4. Zoomed in view of a defect signal in the tube shown in figure 3.

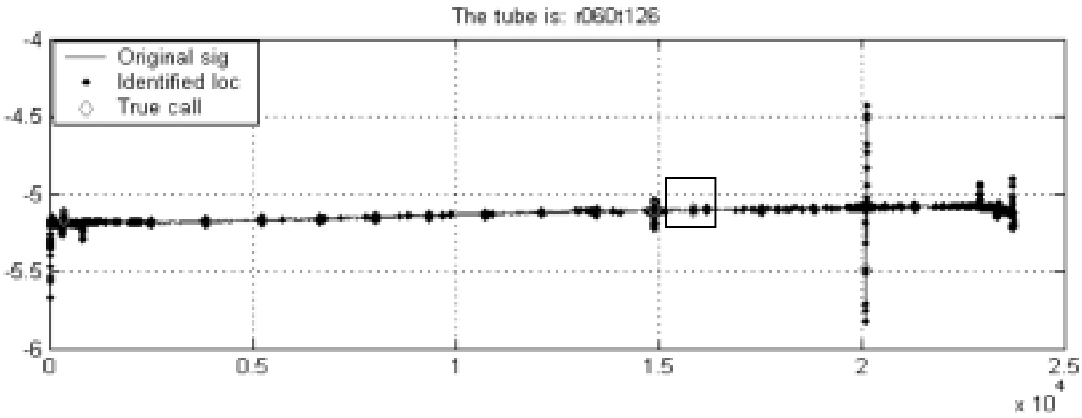

Fig. 5. Bobbin coil eddy current data showing result of dynamic thresholding.

Table 1
Data reduction obtained using the dynamic thresholding algorithm

PFA = 0.25	Minimum reduction over all tubes (%)	Maximum reduction over all tubes (%)	# of dents missed	# of defects missed	# of tubes with missed dents/defects
34 tubes (dents+defects)	58.49	67.75	1	1	2
30 tubes (No defects)	61.77	67.72	–	–	–

4. Conclusions

A dynamic thresholding algorithm based on Neyman-Pearson hypothesis testing was applied to signal reduction in bobbin coil eddy current data from steam generator tubes. The results indicate that the algorithm significantly reduces the number of data points that need to be further analyzed, while detecting almost all the defects/dents present in the database. The signal level of the defect missed was not large enough to be identified by the algorithm with the PFA chosen (0.25). The data set reduction of around

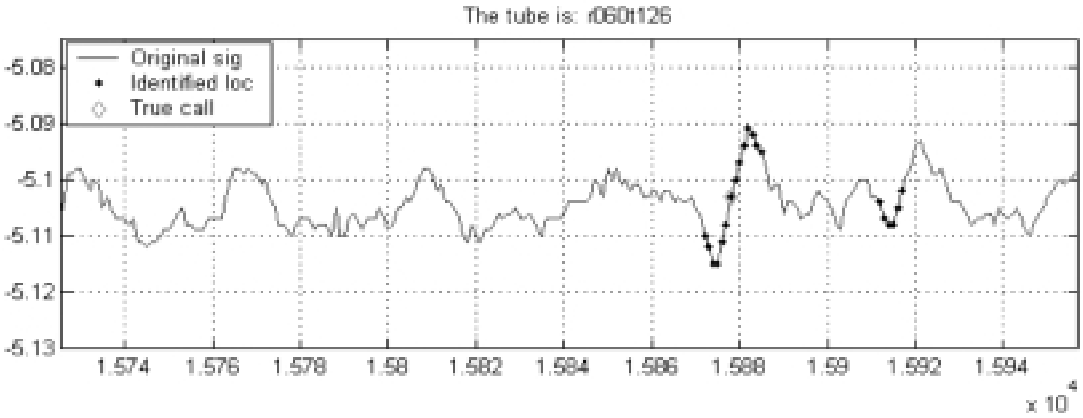

Fig. 6. Zoomed in view of a defect signal in the tube shown in figure 5.

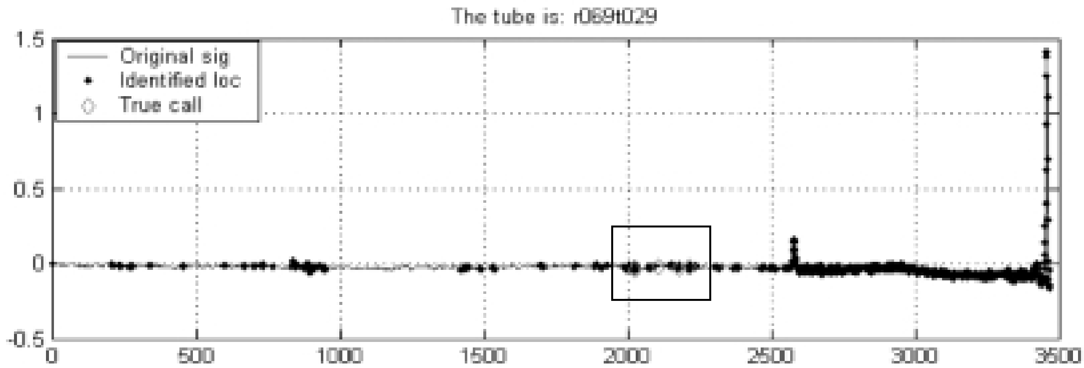

Fig. 7. Bobbin coil eddy current data showing result of dynamic thresholding.

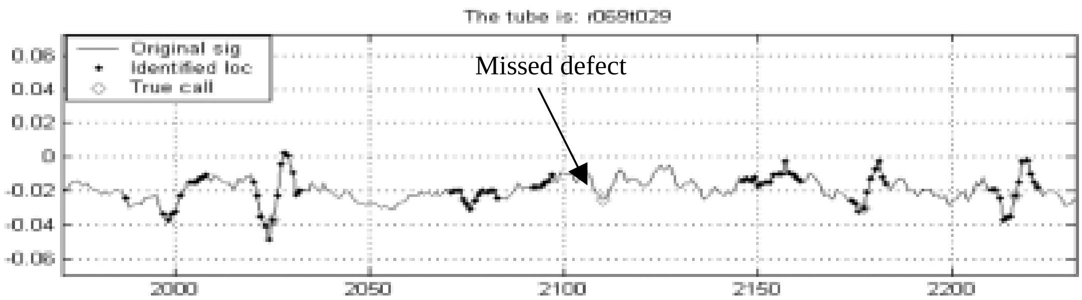

Fig. 8. Zoomed in view of a defect section in the tube shown in figure 7.

60% implies that the subsequent algorithms will have to analyze around 40% of the data. These results are significantly better than those offered by many of the other automated eddy current analysis schemes [7]. These results demonstrate the feasibility of the approach and encourage further evaluation of the scheme using additional field data.

Acknowledgement

This work was supported by the Electric Power Research Institute under contract WO-S533. The assistance of the Project Manager, Jim Benson, is gratefully acknowledged. The authors wish to thank Mr. Steve Brown for providing the data used in this study, as well as for his guidance.

References

[1] A.J. Levy, J.E. Oppenlander, D.M. Brudnoy, J.M. Englund, K.C. Loomis and A.M. Barsky, Dodger: an expert system for eddy current evaluation, *Materials Evaluation* **51** (1993), 34–44.
[2] M. Kay Steven, *Fundamentals of Statistical Signal Processing, Volume 2*, Prentice-Hall PTR: N.J. Englewood Cliffs, c1993–1998.
[3] D. Gibson Jerry and L. Melsa James, *Introduction to non parametric detection with applications*, IEEE press, NJ. Piscataway.
[4] N. McDonough Robert and A.D. Whalen, *Detection of Signals in Noise*, Academic Press: San Diego, CA, 1995.
[5] T.K. Moon and W.C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*, Prentice Hall, Upper Saddle River, NJ, 2000.
[6] A.V. Openheim and R.W. Shafer, *Discrete-Time Signal Processing*, Prentice Hall, Inc., Englewood Cliffs, NJ, 1990.
[7] Steve Brown and Jim Benson, personal communication, 2000.