XAI for signal analysis of guided wave testing at pipe elbows

Toshihiro YAMAMOTO 1, Takashi FURUKAWA 1, Hideo NISHINO 2

1 Japan Power Engineering and Inspection Corporation, Yokohama, Japan
Phone: +81 45 511 1376, Fax: +81 45 511 2750; e-mail: yamamoto-toshihiro@japeic.or.jp
2 Tokushima University, Tokushima, Japan

Abstract
Guided wave testing can provide an efficient screening method for local wall thinning of piping owing to its long inspection range and its ability to inspect pipes with limited access. However, there is a problem in signal interpretation of guided wave testing. While the interpretation of echo signals in guided wave testing is relatively simple for straight pipes, it becomes much more difficult when inspected piping includes an elbow because wave propagation becomes more complicated. This study investigates a signal analysis method that utilizes deep learning to evaluate local wall-thinning at an elbow based on echo signals in guided wave testing using multiple frequencies. To this end, we considered the configuration of deep learning system to analyze echo signals in guided wave testing and utilized Shapley additive explanations (SHAP), one of the explainable AI (XAI) techniques, to provide reasoning for prediction results obtained by deep learning models.

Keywords: guided wave testing, pipe elbow, local wall thinning, deep learning, XAI

1. Introduction

1.1 Background
Guided wave testing can provide an efficient screening method for local wall thinning of piping owing to its long inspection range and its ability to inspect pipes with limited access (at height, buried, covered with insulation, etc.). However, there is a problem in signal interpretation of guided wave testing. Although the interpretation of echo signals to detect flaws in guided wave testing is relatively simple for straight pipes, geometric irregularities such as elbows, tees and reducers make this interpretation much more difficult because these irregularities disturb the propagation of guided waves. Meanwhile, in actual application, local wall thinning such as liquid droplet impingement (LDI) erosion and flow-accelerated corrosion (FAC) often occurs around an elbow. Therefore, how to interpret echo signals is an important issue in using guided wave testing for evaluation of local wall thinning of piping that includes an elbow.

1.2 Our past study
In our past study [1], we conducted laboratory experiments to apply guided wave testing to piping specimens that include an elbow as shown in Fig. 1. These piping specimens are made of aluminum alloy, and their size is 50A Schedule 40, which means 60.5 mm in outer diameter and 3.9 mm in thickness. A dry-coupled piezoelectric sensor system was used to transmit and receive guided waves. This sensor system consists of two sets of eight transducer elements to be aligned circumferentially around the outer surface of a pipe. One set was placed 485 mm before the inlet of the elbow as a transmitter to generate the fundamental torsional guided wave mode T(0, 1), and the other set was placed 30 mm before the transmitter as a receiver.

Figure 2 shows typical received signals when a piping specimen has no flaw. In Fig. 2, there are three clusters of waves, which are here named signals 1, 2 and 3 from left to right. Signal 1 is the wave packet propagating directly from the transmitter to the receiver. Signal 2 is several wave packets reflected around the elbow. Signal 3 is the wave packet reflected by the lower end of the piping specimen. A signal variation due to a flaw at the elbow overlaps signal 2.

We prepared 12 piping specimens for 12 positions of a flaw as shown in Fig. 3. A carbide cutting bit (9.5 mm in diameter) was used to make a U-shaped gutter on the outer surface of the elbows as a simulated flaw (Fig. 4). Each piping specimen has a single flaw at one of the 12
positions. The depth of each flaw was increased up to 2 mm. The echo signals obtained with
the receiver were recorded for every 0.25 mm increase of the depth.

Figure 5 shows the signal variations observed around signal 2 with the increasing depth of
the flaw when the frequency of the guided waves is 50 kHz and the flaw position is 1. The
received signals were normalized so that the amplitude of signal 1 becomes 1000. As shown in
the upper graph of Fig. 5, the signal variations are small compared with the baseline signals.
The baseline signals are defined as the received signals for a specimen without a flaw. To extract
only flaw signals, the baseline signals were subtracted from all the received signals as shown
in the bottom graph of Fig. 5.
1.3 The aim of this study
In our past study [1], we concluded

- High flaw detectability can be obtained throughout the elbow by guided wave testing using multiple frequencies of guided waves.
- The flaw position can be estimated from the frequency at which a strong flaw signal is obtained.

To utilize these characteristics, received signals obtained by guided wave testing using multiple frequencies should be comprehensively analyzed, but this analysis is rather complicated.

The aim of this study is to use deep learning to make the analysis of received signals in guided wave testing more efficient and sophisticated. This paper introduces the study that includes

- FEM simulation was conducted to prepare a number of training data.
- The analysis of received signals was automated by deep learning system.
- The ability to provide reasoning for prediction results by deep learning models was investigated by using SHAP.

2. Signal analysis of guided wave testing at pipe elbows

2.1 FEM simulation
To prepare a number of training data, FEM simulation was performed with ComWAVE, which has been developed by ITOCHU Techno-Solutions Corporation. Figure 6 shows the shape model that was created for the FEM simulation to reproduce the laboratory experiments described in [1]. To reduce the computation time, the element size for meshing was made rather coarse and set to 0.5 mm.
Figures 7–9 show the flaw signals obtained from the experiments and the simulations when the frequency is 50 kHz and the flaw position is 1, 2 and 3 (These positions are aligned in the axial direction of the elbow). The baseline subtraction was performed for these flaw signals. Although the step size for the increasing flaw depth is 0.25 mm in the experiments, the step size is 0.5 mm in the simulations because the element size is 0.5 mm. In addition, the flaw depth was increased up to 3 mm instead of 2 mm in the simulations to increase the variety of the input data for deep learning.

As shown in Figs. 7–9, the flaw signals do not agree very well quantitatively between the experiments and the simulations perhaps partly because their test conditions are not exactly the same. However, on the whole, their general trends are the same. In their trends, the deeper the flaw depth, the larger the amplitude. Also, when the flaw position moves toward the outlet of the elbow, the flaw signal is delayed.

2.2 Deep learning system
The configuration of deep learning system was considered for prediction of flaw information based on flaw signals of guided wave testing.

As the input data for deep learning, the flaw signals obtained for the three frequencies 30, 40 and 50 kHz were connected as shown in Fig. 10 for each combination of the flaw positions and depths. The data set includes these series of the connected flaw signals for the 12 flaw positions and the six levels of the flaw depths. Furthermore, these 72 input data was increased to 720 by adding 10 different patterns of Gaussian noise ($\mu = 0, \sigma = 1$) to each series of the connected flaw signals. These input data were split into 460 training data, 116 validation data and 144 test data.

Figure 11 shows the function to be implemented by deep learning models. The deep learning models receive input data and output predicted values of the flaw depth and position. Table 1 describes the specification of the deep learning system. Because the flaw depth and position are essentially continuous values, to obtain the prediction results as continuous values, predictions were made by solving a regression problem not a classification problem. The flaw positions are separately represented by the axial position (1, 2, 3 or 4) and the circumferential position (1, 2 or 3) as shown in Fig. 12 so that adjacent positions are labeled with consecutive numbers.
Figure 7  Comparison of flaw signals (flaw position 1)

(a) Experiments
(b) Simulations

Figure 8  Comparison of flaw signals (flaw position 2)

(a) Experiments
(b) Simulations

Figure 9  Comparison of flaw signals (flaw position 3)

(a) Experiments
(b) Simulations

Figure 10  Form of input data for deep learning
Table 1  Specification of the deep learning system

<table>
<thead>
<tr>
<th>Libraries</th>
<th>TensorFlow + Keras (written in Python)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category</td>
<td>Fully connected neural network</td>
</tr>
<tr>
<td>Intermediate layers</td>
<td>4 layers (512 nodes for each)</td>
</tr>
<tr>
<td>Dropout</td>
<td>20% for each intermediate layer</td>
</tr>
<tr>
<td>Mini-batch gradient descent</td>
<td>32 in batch size</td>
</tr>
<tr>
<td>Loss function</td>
<td>Mean square error</td>
</tr>
<tr>
<td>Optimizer</td>
<td>RMSprop</td>
</tr>
<tr>
<td>Early stopping</td>
<td>The error exceeds the interim minimum for 200 epochs in a raw</td>
</tr>
</tbody>
</table>

2.3 Signal analysis using SHAP

Explainable AI (XAI) can provide reasoning for prediction results obtained by an AI. Shapley additive explanations (SHAP) are one of the XAI techniques and an explanation technique for machine learning using the notion of Shapley values in game theory [2]. Shapley values are used to distribute profit fairly among the “players” based on their contributions in a cooperative game. When SHAP is used for machine learning, “players” are regarded as the features of input data, and “profit” is regarded as prediction results so as to know how each feature of the input data contributes to the prediction results. We used SHAP to analyze prediction results for guided wave testing because

- SHAP has general versatility and can be applied to various machine learning techniques including deep learning.
- SHAP has been implemented as a Python package and can be used in combination with scikit-learn or TensorFlow [3].

Figure 13 shows an example of analysis of the prediction results by using SHAP. In this case, the axial position of a flaw was predicted, and both the predicted axial position and the true axial position is 1. Although we also made predictions of the circumferential position of a flaw and the flaw depth, the explanations about them are omitted in this paper due to limitations of
space. In Fig. 13, the upper graphs show an example of the input data, and the lower graphs show the SHAP values corresponding to the input data. The horizontal axis represents the data index. Each SHAP value corresponds to the input data element at the same data index. In the examples given in this paper, the envelop curve of each series of the connected flaw signals was obtained and used as the input data. In a regression problem, a predicted value can be broken down into the sum of the SHAP values plus the mean of the predicted values for all the test data. Therefore, positive SHAP values raise the predicted value, and negative SHAP values lower the predicted value. Especially, for prediction of the axial position of a flaw, a large predicted value implies the flaw is closer to the outlet of the elbow, and a small predicted value implies the flaw is closer to the inlet of the elbow. In Fig. 13, the SHAP values show the predicted value was determined mainly based on the particular input data elements that form three humps, each of which appears at the flaw signal for each frequency. The first hump (30 kHz) discreetly advocates the flaw is nearer to the outlet of the elbow. The second and third humps (40 and 50 kHz) advocate the flaw is nearer to the inlet of the elbow.

When the number of the data elements is large, it is difficult to grasp a general relation between the input data and the prediction results. Therefore, the number of the data elements was decreased from 960 to 24 so as to make the SHAP graph simpler and clarify which feature of the input data impacts on the prediction results. For example, the features 10 and 18 indicate the flaw is nearer to the inlet of the elbow in Fig. 13 (b).

Figure 14 shows the confusion matrices obtained with the deep learning model for prediction of the axial position of a flaw when the number of the input data elements is 960 and 24. The confusion matrix provides a visual representation of the predicted value versus the true value and represents the count of each combination of the predicted values and the true values to summarize the performance of a machine learning model. Decreasing the number of the input data elements makes the SHAP graph simpler and clarifies which feature of the input data impacts on the prediction results. For example, the features 10 and 18 indicate the flaw is nearer to the inlet of the elbow in Fig. 13 (b).
data elements makes the analysis more understandable and does not necessarily decrease the accuracy of the prediction.

2.4 Global explanation by SHAP

Figure 13 shows analysis results only for a single input data. SHAP can also provide a global explanation, which enables us to analyze the whole prediction logic of a machine learning model, by using a tool named “summary plot”. Figure 15 shows the summary plot, which uses all the test data, for prediction of the axial position of a flaw when the number of the input data elements is 24. In Fig. 15, the input features (input data elements) are arranged from top to bottom in order of decreasing impact, the color of the dots represents the feature value, and the dot position represents the SHAP value. In Fig. 16, the top 15 features are emphasized with reddish colors, which shows they are concentrated in the three humps of the input data.

Without going into detailed discussion, Fig. 15 indicates:

- When a flaw signal appears earlier, a flaw is located nearer to the inlet of the elbow; when a flaw signal appears later, a flaw is located nearer to the outlet of the elbow.
- The contribution of a flaw signal for 30 kHz to prediction results is smaller.
In the same way, the summary plot for prediction of the circumferential position of a flaw (omitted in this paper) indicates

- When a flaw signal has a large amplitude for 40 or 50 kHz, a flaw is located at the extrados of the elbow; when a flaw signal has a large amplitude for 30 kHz, a flaw is located at the intrados of the elbow.

It might be difficult to understand this tendency intuitively. Further explanation about this is provided in the next subsection. The summary plot for prediction of the flaw depth (omitted in this paper) indicates

- Basically, when a flaw signal has a larger amplitude, a flaw tends to be deeper.
- However, there are some exceptions.

### 2.5 Consideration of analysis results

In our past study [1], FEM simulation was conducted to obtain the distributions of maximum magnitude of displacement on the outer surface of the elbow of the same piping specimen as this study. Figure 17 shows these distributions for the frequencies from 20–60 kHz. These uneven distributions of the maximum magnitude are thought to be attributable to interference caused by guided waves at the elbow. Even if guided waves entering the elbow are coherent, the shape of the elbow gives rise to constructive and destructive interference of the guided waves. In this process, the wavelength of the guided waves changes how the guided waves interfere with each other.

These results indicate that the maximum magnitude tends to be higher at the intrados of the elbow for lower frequencies, while it tends to be higher at the extrados of the elbow for higher frequencies. Meanwhile, the experimental results in our past study [1] indicate that the flaw sensitivity tends to be higher at the intrados of the elbow for lower frequencies, while it tends to be higher at the extrados of the elbow for higher frequencies. This correspondence is understandable because large displacement is supposed to generate large waves reflected by a flaw. If you detect a flaw using 40 or 50 kHz, the flaw is supposed to be located at the extrados of the elbow. If you detect a flaw using 30 kHz, the flaw is supposed to be located at the intrados of the elbow. This explains the tendency derived from the summary plot for prediction of the circumferential position of a flaw.

### 3. Summary

Prediction of flaw information for guided wave testing was well implemented by deep learning models as a regression problem. In the deep learning models, the neural network seems to require at least four layers for good prediction perhaps because this signal analysis is rather complicated.

SHAP can give the ability to provide reasoning for prediction results of deep learning models, which are often called “black box”. To make SHAP analysis results understandable, the number of the input data elements is preferred to be as small as possible. Moreover, because even 24
input data elements are not small enough to make the analysis straightforward, a well-organized analysis method is desired.

In this study, both training data and test data for the deep learning system were obtained by FEM simulation. Toward practical use of this deep learning system, it should deal with prediction from input data obtained in actual measurements based on training data obtained by FEM simulation.

References