Combined Artificial Neural Network and Least Squares Method for Exploring Relationships between Welding Conditions and Weld Characteristics

Houichi Kitanoa,* and Terumi Nakamuraa

a Welding and Joining Technology Group, Research Center for Structural Materials, National Institute for Materials Science, Ibaraki, Japan

In this study, a method was developed for the automatic generation of input–output relationship models that clearly indicate the role of input factors in determining the outputs. This proposed method was used to model the relationships between the welding conditions and penetration dimensions of TIG welds. The models developed using the proposed method clearly and explicitly indicate the effects of the welding current, arc length, and welding speed on the penetration dimensions. In addition, the errors of the predicted weld penetration dimensions obtained using the developed model were less than 3%.

Key Words: Machine learning, RF5, Method of least squares, Empirical formulae, Weld penetration shape

1. INTRODUCTION

Recently, machine learning techniques, such as artificial neural networks (ANNs), have been used to explore the relationships between manufacturing processes and performance with the ultimate goal of optimizing the manufacturing processes. There have also been some research into the use of machine learning techniques to model the relationship between welding conditions and the characteristics of welded parts, such as the weld bead shape, the penetration shape, the weld deformation, and the residual stress distribution [1–6]. With the developed model in the previous researches, it is possible to predict the characteristics of welded parts and propose appropriate welding conditions for the desired characteristics of weld parts. However, there are problems with models developed using the major machine learning techniques such as ANNs, which was used in previous researches. The main problem is that such models are black boxes and difficult to interpret. Namely, the input–output relationship is not explicitly defined in these models. Therefore, it is difficult to understand, for example, how the weld penetration depth changes in direct response to independent changes in the welding current, voltage, and speed.

On the other hand, the conventional regression method of exploring the input–output relationship is the generation of empirical formulae. The input–output relationship is generally clear and explicit in such empirical formulae. Therefore, the use of empirical formulae makes it easy to understand the effects of the input factors on the outputs. However, this approach requires the manual collection and organization of a large amount of experimental data relevant to the physical mechanisms of the objective phenomena. Therefore, it is difficult to generate empirical formulae describing complex phenomena influenced by many factors.

In this study, a method was developed for the automatic generation of input–output relationship models that clearly indicate the role of the input factors in determining the outputs. The proposed method is a combination of a law discovery method called Rule Extraction from Facts version 5 (RF5) and the method of least squares. RF5 is an ANN-based method that can output an approximate model of the input–output relationship in the form of a polynomial equation as a result of learning without any prototype formulae [7]. In models developed using the proposed combination of the two methods, the role of input factors is clear and easy to interpret. Namely, the method has the advantages of both previous machine learning techniques and conventional regression method to make empirical formulae. The method can develop models in which input–output relationships are clear and explicit in contrast to the model developed by previous machine learning techniques. In addition, the model can be developed automatically in contrast to the conventional regression method. The method was then applied to development of a model of the relationship between welding conditions and weld penetration shapes of tungsten inert gas (TIG) welds. The relationship between welding conditions and weld penetration shapes of TIG welds is relatively simple because heat transfer by droplets does not occur. Therefore, the relationship is appropriate to evaluate of the applicability of proposed method. Then, the developed models obtained as learning results using the proposed method and their accuracy are discussed.
2. MODELING BY A COMBINATION OF RF5 AND THE METHOD OF LEAST SQUARES

2.1 RF5 Method

RF5 can be used to obtain input–output relationships approximated as polynomial equations without requiring the preparation of appropriate prototype functions. Let \( \mathbf{x} \) be an \( n \)-dimensional input vector and \( y \) be a target value corresponding to \( \mathbf{x} \). In the RF5 method, it is assumed that the input–output relationships can be approximated as polynomial equations expressed as

\[
y = c_0 + \sum_{i=1}^{n} c_i \prod_{j=1}^{n} x_j^{w_{ij}}, \tag{1}
\]

where \( c_0, c_i, \) and \( w_{ij} \) are real numbers and \( h \) is a natural number. An appropriate value can be added to each input element of \( \mathbf{x} \) to convert all inputs to positive values. Therefore, Eq. (1) is equivalent to

\[
y_t = c_0 + \sum_{i=1}^{h} c_i \exp(\sum_{j=1}^{n} w_{ij} \ln(x_j)). \tag{2}
\]

Equation (2) can be regarded as a three-layer feed-forward neural network, where \( h \) is the number of neurons in the hidden layer, \( w_{ij} \) is the weight between the \( j \)-th input and the \( i \)-th hidden neuron, \( c_i \) is the weight between the \( i \)-th hidden neuron and the output, and \( c_0 \) is the bias of the output neuron. The hidden neurons have no biases. After the value of \( h \) is determined, the input–output relationships are approximated as the polynomial equation written in Eq. (1) by optimizing \( w_{ij}, c_i, \) and \( c_0 \).

2.2 Combination of RF5 and Method of Least Squares

The application of RF5 yields the input–output relationships in the form of the polynomial equation in Eq. (1). The number of terms of the equation is \( h + 1 \). Each term is calculated as the product of \( c_i \) and \( x_j \) raised to the power \( w_{ij} \). However, when RF5 is applied to develop a model of an input–output relationship describing an actual phenomenon, the role of each input factor cannot be readily understood, because the exponents \( w_{ij} \) are real numbers. Therefore, in this study, it is assumed that the input–output relationships can be approximated by the following polynomial form:

\[
y = c'_0 + \sum_{i=1}^{h} c'_i \prod_{j=1}^{n} x_j^{b_{ij}/a_{ij}}, \tag{3}
\]

where \( a_{ij} \) is a natural number less than or equal to 10, \( b_{ij} \) is an integer, and \( c'_0 \) and \( c'_i \) are real numbers. The upper limit of \( a_{ij} \) increases the readability of the effects of the input factors on the output in the developed input–output relation. \( a_{ij} \) and \( b_{ij} \) are determined by determining the value of \( b_{ij}/a_{ij} \) that is closest to the value of \( w_{ij} \) obtained by RF5. After \( a_{ij} \) and \( b_{ij} \) are determined, \( c'_0 \) and \( c'_i \) are adjusted to fit the input–output relationship in Eq. (3) to those obtained using the learning dataset by applying the method of least squares.

The proposed procedure for the development of input–output relationship models is as follows:

Step 1. Set the value of \( h \) and determine the number of terms in the final polynomial approximation.

Step 2. Use RF5 to compute the relationship for the learning dataset that optimizes \( w_{ij}, c_i, \) and \( c_0 \).

Step 3. Calculate \( a_{ij} \) and \( b_{ij} \) by determining the value of \( b_{ij}/a_{ij} \) closest to the value of \( w_{ij} \) optimized by RF5.

Step 4. Apply the method of least squares to calculate the values of \( c'_0 \) and \( c'_i \) that fit the input–output relationship in Eq. (3) to the relationship obtained using the learning dataset.

Step 5. Check if the error of the predicted value obtained by the developed input–output relationship model is sufficiently small.

Step 6. Stop if the error is sufficiently small; otherwise, return to Step 1 and increase the value of \( h \).

To check the error, any general index of the error, such as the mean absolute error or root mean square error, can be used. When initial value of \( h \) is set to 1, an approximate polynomial equation with fewest number of terms can be obtained as the initial results of Step 4. The input–output relationship is most clear in the equation. If the prediction error of the equation is not sufficiently small, the value of \( h \) can be increased to obtain more accurate polynomial equation as written in Step 6.

3. WELD PENETRATION SHAPE PREDICTION BY PROPOSED METHOD

3.1 Learning Dataset

The penetration shapes of TIG welds reported by Narang et al. [8] were used as the learning dataset in the proposed method. The material used in the experiment was structural steel, the chemical composition of which is shown in Table 1. The length, width, and thickness of the test pieces were 180, 65, and 8 mm, respectively. The measured penetration width and depth of the weld bead obtained from test pieces welded under various welding
Equation (2) can be regarded as a three-layer feed-forward neural network, where
\[ h \] is the number of neurons in the hidden layer, \( \ell \) is the value of the input–output relationship models for bead width and depth of penetration are as follows:

\[ BW = 1.16 \frac{I^3}{\ell^3} + 1.39, \tag{4} \]

\[ DP = 0.46 \frac{I^3}{\ell^3} + 0.43, \tag{5} \]

where \( BW \) is the bead width, \( DP \) is the depth of penetration, \( I \) is the welding current, \( Al \) is the arc length, and \( v \) is the welding speed. In Eqs. (4) and (5), the values of exponents, coefficients and constant term correspond to \( h, a, c \), and \( c_0 \) optimized in Step 1 to 4 in Section 2.2 respectively. The predicted and experimental results for the bead width \( BW \) and penetration depth \( DP \) are shown in Fig. 1. The mean absolute error \( \varepsilon \) was used as the index of the error of the predicted results with respect to the experimental results. The mean absolute error \( \varepsilon \) is defined as follows:

\[ \varepsilon = \sum \frac{|y - \hat{y}|}{n} \]

Table 1 Chemical composition of structural steel (wt%) [8].

| C  | Si  | Mn  | P  | S  | Ni  | Cr | Fe |
|----|-----|-----|----|----|-----|----|----|
| 0.16 | 0.178 | 0.45 | 0.18 | 0.07 | 0.13 | 0.016 | Bal. |

Table 2 Penetration dimensions [8].

| No. | Current (A) | Arc length (mm) | Welding Speed (mm/s) | Bead width (mm) | Depth of penetration (mm) |
|-----|-------------|-----------------|----------------------|----------------|--------------------------|
| 1   | 55          | 2               | 15                   | 5.46           | 1.59                     |
| 2   | 55          | 2               | 30                   | 4.71           | 1.25                     |
| 3   | 55          | 2               | 45                   | 4.16           | 1.04                     |
| 4   | 55          | 2.5             | 15                   | 5.77           | 1.76                     |
| 5   | 55          | 2.5             | 30                   | 4.93           | 1.38                     |
| 6   | 55          | 2.5             | 45                   | 4.46           | 1.18                     |
| 7   | 55          | 3               | 15                   | 6.09           | 1.91                     |
| 8   | 55          | 3               | 30                   | 5.03           | 1.42                     |
| 9   | 55          | 3               | 45                   | 4.55           | 1.23                     |
| 10  | 55          | 2               | 15                   | 6.12           | 1.99                     |
| 11  | 75          | 2               | 30                   | 5.13           | 1.39                     |
| 12  | 75          | 2               | 45                   | 4.59           | 1.16                     |
| 13  | 75          | 2.5             | 15                   | 6.59           | 2.06                     |
| 14  | 75          | 2.5             | 30                   | 4.26           | 1.5                      |
| 15  | 75          | 2.5             | 45                   | 4.85           | 1.32                     |
| 16  | 75          | 3               | 15                   | 7.07           | 2.18                     |
| 17  | 75          | 3               | 30                   | 5.45           | 1.65                     |
| 18  | 75          | 3               | 45                   | 5.16           | 1.45                     |
| 19  | 95          | 2               | 15                   | 6.65           | 2.17                     |
| 20  | 95          | 2               | 30                   | 5.38           | 1.51                     |
| 21  | 95          | 2               | 45                   | 4.75           | 1.23                     |
| 22  | 95          | 2.5             | 15                   | 7.19           | 2.23                     |
| 23  | 95          | 2.5             | 30                   | 6.16           | 1.63                     |
| 24  | 95          | 2.5             | 45                   | 5.2            | 1.32                     |
| 25  | 95          | 3               | 15                   | 7.64           | 2.51                     |
| 26  | 95          | 3               | 30                   | 6.31           | 1.74                     |
| 27  | 95          | 3               | 45                   | 5.11           | 1.41                     |
 where \( f_k \) is the predicted value, \( y_k \) is the actual value, and \( n \) is the number of data points. The mean absolute error \( \varepsilon \) of the bead width \( BW \) and depth of penetration \( DP \) calculated from Eq. (6) were 2.88% and 2.95%, respectively. Increase in \( h \) could be a cause of overfitting of the developed model because there is error in the data set used for learning. Therefore, return to Step 1 in Step 6 did not performed. Equations (4) and (5) represent the role of each input factor and allow the effects of the welding current, arc length, and welding speed on the bead width \( BW \) and penetration depth \( DP \) to be readily understood. From Eq. (4), it can be understood that the bead width \( BW \) is proportional to the welding current raised to the power of 1/2, the arc length raised to the power of 3/8, and the welding speed raised to the power of 3/8. From Eq. (5), it also can be understood that the penetration depth \( DP \) is proportional to the welding current raised to the power of 5/9, the arc length raised to the power of 1/2, and the welding speed raised to the power of 3/5. In this way, the effect of input factors on the outputs can be understood quickly and mathematically. For example, Eq. (4) indicates that welding current must be quadrupled to double the bead width.

4. CONCLUSIONS

In this study, a method was developed for the generation of input–output relationship models that clearly indicate the role of input factors in determining the output. The proposed method is a combination of RF5 and the method of least squares. In this method, input–output relationships are approximated as polynomial expressions without requiring the preparation of appropriate prototype functions. The proposed method was then applied to the development of relationship models between the welding conditions and the weld penetration dimensions for TIG welds. In this way, models of the relationships between the welding conditions and the weld penetration dimensions were developed. The mean absolute errors of the predicted weld penetration dimensions obtained using the developed model were less than 3%.

The proposed method can be applied to the development of input–output relationship models of more complex phenomena that are influenced by many factors. Models developed using the proposed method yield clear, explicit, and easily interpretable input–output relationships.

Reference

1) H. Fujii and K. Ichikawa: Estimation of Weld Properties by Bayesian Neural Network Part 1, Journal of the Japan Welding Society, 69-3 (2000), 44-47.
2) H. Fujii and K. Ichikawa: Estimation of Weld Properties by Bayesian Neural Network Part 2, Journal of the Japan Welding Society, 70-3 (2001), 23-27.
3) M. P. Lightfoot, G. J. Bruce, N. A. McPherson and K. Woods: The application of artificial neural networks to weld-induced deformation in ship plate, WELDING JOURNAL-NEW YORK-, 84-2 (2005), 23-30.
4) R. Dhas and S. Kumanan: Weld residual stress prediction using artificial neural network and. Fuzzy logic modeling, Indian Journal of Engineering & Materials Sciences, 18 (2011), 351-360.
5) A. Iqbal, S. M Khan and H. S. Mukhtar: ANN assisted prediction of weld bead geometry in gas tungsten arc welding of HSLA steels, Proceedings of the World congress on engineering, 1 (2011), 6-8.
6) L. Yu, K. Saida, M. Mochizuki, K. Nishimoto, S. Hirano and N. Chigusa: Neural network-based HAZ hardness prediction system for multi-layer welding extended its application to various steels, Quarterly Journal of the Japan Welding Society, 35-4 (2017), 179-193.
7) K. Saito and R. Nakano: Law Discovery using Neural Networks, Proceedings of the 15th International Joint Conference on Artificial Intelligence, San Francisco (1997), 1078-1083.
8) H. K. Narang, U. P. Singh, M. M. Mahapatra and P. K. Jha: Prediction of the weld pool geometry of TIG arc welding by using fuzzy logic controller, International Journal of Engineering, Science and Technology, 3-9 (2011), 77-85.