Radial basis function neural network model based prediction of weld-plate distortion due to pulsed metal inert gas welding

Sukhomay Pal, Surjya K. Pal* and Arun K. Samantaray

Department of Mechanical Engineering, Indian Institute of Technology Kharagpur, 721302, India

Abstract: Welding shrinkage and distortion affect the shape, dimensional accuracy and the strength of the finished product. This work concerns the prediction of welding distortion in a pulsed metal inert gas welding (PMIGW) process. Six different types of radial basis function network models have been developed to predict the distortion of welded plates. Six process parameters, namely pulse voltage, back-ground voltage, pulse duty factor, pulse frequency, wire feed rate and the welding speed; along with the root mean square (RMS) values of two sensor signals, namely the welding current and the voltage signals, are used as input variables of these models. The angular distortion and the transverse shrinkage of the welded plate are considered as the output variables. Inclusion of sensor signals in the models, as developed in this work, results in better output prediction.

Keywords: PMIGW; Distortion prediction; Radial basis function network; Response surface methodology

Nomenclature

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$bV$</td>
<td>Back-ground voltage (in Volts)</td>
</tr>
<tr>
<td>$V_p$</td>
<td>Pulse voltage (in Volts)</td>
</tr>
<tr>
<td>$f$</td>
<td>Pulse frequency (in Hz)</td>
</tr>
<tr>
<td>$P_f$</td>
<td>Percentage pulse duty factor</td>
</tr>
<tr>
<td>$v_w$</td>
<td>Wire feed rate (in m/min)</td>
</tr>
<tr>
<td>$v_i$</td>
<td>Table feed rate (in mm/s)</td>
</tr>
<tr>
<td>$\mu_i$</td>
<td>Root mean square (RMS) value of recorded current signal (in Volts)</td>
</tr>
<tr>
<td>$\mu_v$</td>
<td>Root mean square (RMS) value of recorded voltage signal (in Volts)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Transverse shrinkage</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Angular distortion (in degrees)</td>
</tr>
<tr>
<td>$X_i$</td>
<td>$i^{th}$ input pattern</td>
</tr>
</tbody>
</table>

* Corresponding author. Tel.: +91 3222 282996; fax: +91 3222 255303. E-mail address: skpal@mech.iitkgp.ernet.in (S.K. Pal).
1. Introduction

In an arc welding process, the welding zone is locally heated by the welding arc. The temperature distribution in the weld and the surrounding base material are non-uniform and the temperature also changes with welding time. During the welding cycle, non-uniform thermal stresses resulting from plastic thermal strains produce internal forces which result in welding distortion, shrinkage, bending, etc. Welding shrinkage and distortion affect the shape, dimensional accuracy and strength of a finished product. Therefore, the prediction and control of welding shrinkage and distortion have become an important issue in modern manufacturing industries.

Welding distortion in a simple butt joint for a rectangular plate with single side weld with each part having free expansion of the joint, can be divided into four types, namely, 1) transverse shrinkage, 2) longitudinal shrinkage, 3) angular distortion, and 4) bending distortion. These distortions are schematically shown in Fig. 1. Note that the longitudinal shrinkage and bending distortion are generally small. There has been lots of research to predict weld distortion using analytical and numerical models \(^1\)-\(^8\), but these models are inaccurate and do not consider the actual welding condition.

Pulsed metal inert gas welding (PMIGW), as a major joining processes, has following inherent advantages: deep penetration, smooth weld bead, high welding speed, large metal deposition rate, lower spatter, lower distortion and shrinkage, and lesser probability
of porosity and fusion defects. In an automated manufacturing environment, the total quality index of a product depends on the quality of output from every sub-process in the production chain and obviously, welding is one of them. Consequently, different technologies are needed precisely predict the weld quality at different operating conditions. Weld quality can be measured directly or indirectly. Direct methods are visual inspection and vision sensing of the weld puddle; indirect methods are arc sensing, infrared sensing, radiographic sensing, inductive sensing, arc sound sensing, acoustic emission sensing and ultrasonic sensing. All the above mentioned methods use measurements of some signal(s) to correlate with the weld quality. Among the various sensors used, arc sensors, i.e., current and voltage sensors, are considered to be the most reliable, simple and competitive.

![Diagram](image)

Fig. 1. Different types of distortions of a rectangular welded plate (not to scale).

Arc sensors monitor the change in one or more of the electrical parameters of the arc, i.e., current and/or voltage. Researchers have proposed arc sensing techniques for seam tracking in arc welding processes and online monitoring and control of the welding processes.

Most of the research on arc signal based weld quality monitoring analyze arc stability and weld quality by time and frequency domain methods. Arc sensing techniques have also been used to determine droplet detachment. In these works, statistical qualifiers of arc signals are compared to some preset nominal values.

Deng et al. developed a three dimensional thermal elastic-plastic finite element model to predict distortion and shrinkage. Tseng and Chou studied the effect of
shielding gas on angular distortion. Weld plate distortion was predicted through an artificial neural network (ANN) model by Lightfoot et al.\textsuperscript{19-20} considering the standard deviation as the measure of the actual and the predicted distortions.

Multilayer neural network is especially suitable for model based supervision of uncertain systems. Andersen et al.\textsuperscript{21} pioneered the application of neural networks in modeling of arc welding processes. Since then, many researchers have modeled arc welding processes by using various kinds of ANN models. Cook et al.\textsuperscript{22} used back propagation neural network (BPNN) models for welding process modeling and control. ANN model has been developed to select welding parameters for required output specifications\textsuperscript{23}. ANN model gives better predictions of back-bead geometry in metal inert gas welding (MIGW) as compared to multiple regression analysis model\textsuperscript{24}. A RBF neural network (RBFN) model was used for modeling and optimizing a MIG welding process\textsuperscript{25}.

However, very few attempts have been made to correlate the arc signals to the weld quality by using ANN models. Most of the above mentioned works in literature either use only the process parameters to correlate with the resultant weld quality or use both process parameters and sensor signals, but use different inefficient statistical methods which do not properly account for process uncertainties. Note that having same process parameter setting does not always result in the same output quality, the variations could be significant (as shown in the experimental results given in this work) due to various process uncertainties and external disturbances. This is why, sensor signals may be used to further improve the prediction accuracy and this work is primarily concerned with investigations in this direction. In the limited number of works reported in this direction, Ohshima et al.\textsuperscript{26} proposed a neuro-arc sensor model to simultaneously detect deviation and height of the torch. Quero et al.\textsuperscript{27} used current signal and ANN technique to monitor the weld quality.

In this work, authors have developed a model to predict the welding distortion by six different types pre-trained radial basis function network (RBFN) models, by utilizing the statistical properties of the signals taken during the welding process as inputs in addition to the usual process parameters to refine the model. The trainings of RBFN models were
done by using part of available experimental data, which were obtained from response surface method and thereafter, the remaining experimental data were used to test and compare the performance of various RBFN models.

2. Experiments

2.1. Specimen preparation

In this work, a set of two mild steel specimens, each having dimension of 125 mm × 100 mm × 8 mm, were used as the workpiece. Optical emission spectroscopy was done to find out the chemical composition of the base metal (Table 1). These specimens were prepared with V-shaped groove having the groove angle, the root face and the root gap of 30°, 2 mm and 2 mm, respectively. Thereafter, 53 pairs of such specimens were prepared and then their faces were cleaned by a surface grinder. To make a butt weld joint, two plates were tack welded at the two ends as shown by points A and B in Fig. 2. Transverse shrinkage were measured at $L_1$, $L_2$ and $L_3$ locations (Fig. 2) by using vernier caliper and their average was calculated. If shrinkage is $\Delta L$ and $L$ is length at any measured location, then transverse shrinkage is defined as

$$\rho = \frac{1}{3} \left( \frac{\Delta L_1}{L_1} + \frac{\Delta L_2}{L_2} + \frac{\Delta L_3}{L_3} \right)$$

(1)

The angular distortion can be calculated by using the coordinates in vertical direction before and after welding. The vertical coordinates, as shown in Fig. 2 by solid circles, were measured using dial indicator (Mitutoyo Corp., model: ID-S1012). The values at $D_{bi}$ were measured with respect to the points $D_{ai}$ (reference point) by clamping the left edge of the plate. The mean vertical displacement is obtained by taking measurements at five such points:

$$D = \frac{1}{5} \sum_{i=1}^{5} (D_{bi} - D_{ai})_{t=0} - (D_{bi} - D_{ai})_{t=\infty},$$

(2)

where $D_{ai}$ & $D_{bi}$ ($i = 1...5$) are the vertical displacement at $i^{th}$ point shown in Fig. 2, and $t=0$ and $t=\infty$ refer to the pre-welded state and the state after sufficient time has elapsed such that the welded-plate has cooled down to the environmental temperature.
The angular distortion is given by

\[ \theta = \tan^{-1}\left( \frac{2D}{L} \right) \]  

(3)

where \( L = 200 \text{ mm} \) is the length of each specimen.

Table 1. Chemical composition of base metal (in weight percentage)

<table>
<thead>
<tr>
<th>C</th>
<th>Si</th>
<th>Mn</th>
<th>P</th>
<th>S</th>
<th>Ni</th>
<th>Cr</th>
<th>Cu</th>
<th>Fe</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.139</td>
<td>0.151</td>
<td>0.499</td>
<td>0.075</td>
<td>0.044</td>
<td>0.024</td>
<td>0.019</td>
<td>0.056</td>
<td>98.993</td>
</tr>
</tbody>
</table>

Fig. 2. Locations of shrinkage and distortion measurement and tack weld.

2.2. Equipment

A Fronius make welding machine is used in the present study. It has a constant voltage source, Transarc 500 and the control unit is VR131 type. A schematic diagram of the experimental setup is shown in Fig. 3.

The welding torch or welding gun (model AW502) was mounted on a fixed arm. The motor-driven carriage has a variable speed in the range of 1 mm s\(^{-1}\) to 16 mm s\(^{-1}\). Copper coated mild steel wire of 1.2 mm diameter is used in the experiment. The shielding gas (Argon) was supplied in a regulated manner at a constant flow rate of 15.0 liters/min, from a constant pressure source of 10 kgf/cm\(^2\).

A Hall Effect current transducer (LEM, model LT 500S) was used to monitor the welding current. Moreover, potential difference (scaled down in 1:11 ratio) was sensed between the workpiece and the contact tip. The analog outputs from these sensors were
converted into digital signals by an A/D card (Measurement computing corporation, model: PCI-DAS 4020/12) fitted to an IBM PC through sampling at 10 kHz.

![Schematic arrangement of the experimental setup.](image)

**Fig. 3.** Schematic arrangement of the experimental setup.

2.3. Experimental procedure

Response surface method\(^2^8\) was used to understand the effect of different process parameters on the weld bead geometry. Three levels, six factors and half fraction central composite experimental design with nine center points was performed. This design requires fifty-three experimental runs. A commercially available software package, MINITAB\(^2^9\), was used to setup the design matrix. The design matrix is shown in Table 2.

3. RBFN modeling of distortion of the PMIGW process

The advantages of ANN models are that models are developed from the experimental database without making any simplifying assumptions; and in many cases, they even outperform their classical counterparts. Various types of ANN are used for modeling\(^1^9\-^2^0, ^2^3\-^2^5\). In this work, RBFN model has been developed and used. RBFN models train extremely fast with less chance falling into local minima and they require fewer training samples\(^3^0\). This technique often proves to be more accurate than other network models\(^3^1\).
<table>
<thead>
<tr>
<th>Exp. No</th>
<th>$V_b$</th>
<th>$V_p$</th>
<th>$f$</th>
<th>$P_f$</th>
<th>$v_w$</th>
<th>$v_i$</th>
<th>$\mu_i$</th>
<th>$\mu_v$</th>
<th>$\theta$</th>
<th>$\rho \times 10^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17</td>
<td>34.6</td>
<td>130</td>
<td>50</td>
<td>11</td>
<td>3.76</td>
<td>1.3023</td>
<td>2.5549</td>
<td>1.28921</td>
<td>1.781</td>
</tr>
<tr>
<td>2</td>
<td>17</td>
<td>34.6</td>
<td>130</td>
<td>50</td>
<td>9</td>
<td>3.76</td>
<td>1.1415</td>
<td>2.7449</td>
<td>1.45771</td>
<td>1.798</td>
</tr>
<tr>
<td>3</td>
<td>14</td>
<td>30</td>
<td>182</td>
<td>35</td>
<td>11</td>
<td>2.456</td>
<td>1.5672</td>
<td>1.7755</td>
<td>0.643405</td>
<td>0.163</td>
</tr>
<tr>
<td>4</td>
<td>17</td>
<td>34.6</td>
<td>130</td>
<td>65</td>
<td>9</td>
<td>3.76</td>
<td>1.2652</td>
<td>2.8668</td>
<td>0.775737</td>
<td>1.727</td>
</tr>
<tr>
<td>5</td>
<td>14</td>
<td>30</td>
<td>182</td>
<td>65</td>
<td>11</td>
<td>5.635</td>
<td>1.2566</td>
<td>2.3814</td>
<td>1.307208</td>
<td>1.208</td>
</tr>
<tr>
<td>6</td>
<td>20</td>
<td>30</td>
<td>182</td>
<td>35</td>
<td>7</td>
<td>2.456</td>
<td>1.0122</td>
<td>2.4273</td>
<td>0.354238</td>
<td>2.387</td>
</tr>
<tr>
<td>7</td>
<td>20</td>
<td>39</td>
<td>182</td>
<td>65</td>
<td>7</td>
<td>2.456</td>
<td>1.246</td>
<td>2.3827</td>
<td>1.386287</td>
<td>3.242</td>
</tr>
<tr>
<td>8</td>
<td>17</td>
<td>34.6</td>
<td>130</td>
<td>50</td>
<td>9</td>
<td>3.76</td>
<td>1.1717</td>
<td>2.7386</td>
<td>1.352307</td>
<td>1.523</td>
</tr>
<tr>
<td>9</td>
<td>14</td>
<td>30</td>
<td>80</td>
<td>35</td>
<td>7</td>
<td>2.456</td>
<td>1.0072</td>
<td>2.2308</td>
<td>1.377016</td>
<td>2.088</td>
</tr>
<tr>
<td>10</td>
<td>17</td>
<td>34.6</td>
<td>80</td>
<td>50</td>
<td>9</td>
<td>3.76</td>
<td>1.1516</td>
<td>2.7705</td>
<td>1.026849</td>
<td>1.974</td>
</tr>
<tr>
<td>11</td>
<td>17</td>
<td>34.6</td>
<td>182</td>
<td>50</td>
<td>9</td>
<td>3.76</td>
<td>1.1434</td>
<td>2.6927</td>
<td>1.857914</td>
<td>1.900</td>
</tr>
<tr>
<td>12</td>
<td>30</td>
<td>80</td>
<td>35</td>
<td>11</td>
<td>2.456</td>
<td>1.4692</td>
<td>1.9772</td>
<td>1.706364</td>
<td>1.208</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>20</td>
<td>39</td>
<td>80</td>
<td>65</td>
<td>11</td>
<td>2.456</td>
<td>1.1052</td>
<td>2.9427</td>
<td>1.420071</td>
<td>1.911</td>
</tr>
<tr>
<td>14</td>
<td>17</td>
<td>34.6</td>
<td>80</td>
<td>50</td>
<td>9</td>
<td>3.76</td>
<td>1.0516</td>
<td>2.7334</td>
<td>0.767717</td>
<td>2.268</td>
</tr>
<tr>
<td>15</td>
<td>17</td>
<td>34.6</td>
<td>130</td>
<td>50</td>
<td>7</td>
<td>3.76</td>
<td>1.1786</td>
<td>2.6165</td>
<td>1.02033</td>
<td>1.244</td>
</tr>
<tr>
<td>16</td>
<td>14</td>
<td>30</td>
<td>80</td>
<td>35</td>
<td>11</td>
<td>5.635</td>
<td>0.9973</td>
<td>2.2885</td>
<td>0.353819</td>
<td>1.010</td>
</tr>
<tr>
<td>17</td>
<td>17</td>
<td>34.6</td>
<td>130</td>
<td>50</td>
<td>9</td>
<td>3.76</td>
<td>1.1945</td>
<td>2.7313</td>
<td>1.284363</td>
<td>1.459</td>
</tr>
<tr>
<td>18</td>
<td>14</td>
<td>30</td>
<td>80</td>
<td>65</td>
<td>11</td>
<td>2.456</td>
<td>0.979</td>
<td>2.6119</td>
<td>1.167035</td>
<td>1.953</td>
</tr>
</tbody>
</table>

Table 2. Design matrix of experimental run with corresponding responses.
3.1. RBFN structure

Broomhead and Lowe\textsuperscript{32} were the first to exploit the potential of RBFN. RBFN mainly consists of three layers having different tasks for different layers. The input layer receives the information from an external source. In this case, it is an eight dimensional input vector (\( X \)), where each dimension represents an input parameter. The second layer is the hidden layer consisting of \( H \) number of basis functions. The main characteristic feature of RBF is that its response increases or decreases monotonically with distance from a central point. Various functions have been used as basis functions or activation functions, among them Gaussian function is the most popular. There are no connection weights between the input layer and the hidden layer. The hidden layer transforms the input space into a higher dimensional hidden space in a nonlinear manner which, according to Cover’s theorem\textsuperscript{33}, is likely to increase the linear separability of the input pattern. The output layer is generally linear, but may be nonlinear\textsuperscript{34} and it is terminated at the external receptor node(s). In this work, both linear and nonlinear activation functions are used in the output layer.

3.2. Forward pass calculation

\textit{Output of hidden neuron}: The output of the \( j^{th} \) hidden node is given as

\[
R_j^i(n) = e^{-\frac{\|X_i - C_j(n)\|^2}{2\sigma_j(n)^2}},
\]

(4)

where variables are declared in the nomenclature, \( \|X_i - C_j(n)\|^2 \) is a Euclidean norm and the dimension of \( C_j \) is same as \( X \).

\textit{Final output from an output neuron}: The output of \( k^{th} \) output neuron at \( n^{th} \) iteration for \( i^{th} \) input pattern was calculated using three types of activation functions:

\[
O_k^i(n) = \begin{cases} S_k^i(n), & \text{linear;} \\ \frac{S_k^i(n)}{\sum_{j=1}^{H} R_j^i(n)}, & \text{pseudo-linear;} \\ \frac{1}{1 + e^{-S_k^i(n)}}, & \text{sigmoid;} \end{cases}
\]

(5)

and

\[
S_k^i(n) = \sum_{j=0}^{H} w_{jk} R_j^i(n).
\]

(6)
3.3. Training algorithms

The training of the network was done by minimization of mean square error (MSE)

\[ MSE = \xi(n) = \frac{1}{2NM} \sum_{i}^{N} \sum_{k}^{M} (T_{i}^{k} - O_{i}^{k}(n))^{2}. \]  

(7)

Two types of training algorithms were used in the present work:

1. Half training: in this case, the hidden layer parameters, i.e., center vector and width of Gaussian function were updated using k-mean clustering technique and connection weights of output layer were adjusted through back propagation algorithm. The training of the two layers was done in two different time scales. First, the hidden layer and then the output layer were updated as follows:

\[ C_{j}(n+1) = \begin{cases} C_{j}(n) + \eta_{1}(X_{j} - C_{j}(n)), & \text{if } C_{j} \text{ is the winning center for } i^{th} \text{ input pattern;} \\ C_{j}(n), & \text{otherwise.} \end{cases} \]  

(8)

\[ \sigma_{j}(n+1) = \frac{d_{\max}(n)}{\sqrt{2H}}, \quad j = 1 \ldots H, \]  

(9)

\[ w_{jk}(n+1) = w_{jk}(n) - \eta_{2} \frac{\partial \xi(n)}{\partial w_{jk}(n)}, \]  

(10)

where learning rates ($\eta_{1}$ and $\eta_{2}$) are in the range 0 to 1.

2. Full training: In this case, all free parameters (center vector, width of Gaussian function and weights between hidden and output layers) were updated through back propagation algorithm as follows:

\[ C_{j}(n+1) = C_{j}(n) - \eta_{3} \frac{\partial \xi(n)}{\partial C_{j}(n)}, \]  

(11)

\[ \sigma_{j}(n+1) = \sigma_{j}(n) - \eta_{4} \frac{\partial \xi(n)}{\partial \sigma_{j}(n)}, \]  

(12)

\[ w_{jk}(n+1) = w_{jk}(n) - \eta_{5} \frac{\partial \xi(n)}{\partial w_{jk}(n)}, \]  

(13)

where the learning rates ($\eta_{3}$, $\eta_{4}$ and $\eta_{5}$) are in the range 0 to 1.

Six different cases have been considered to predict the welding distortion, as discussed below.

**Case 1:** Linear activation function for the output layer with half training algorithm.

**Case 2:** Pseudo linear activation function for the output layer with half training algorithm.
Case 3: Sigmoid activation function for the output layer with half training algorithm.

Case 4: Linear activation function for the output layer with full training algorithm.

Case 5: Pseudo linear activation function for the output layer with full training algorithm.

Case 6: Sigmoid activation function for the output layer with full training algorithm.

4. Experimental results

The transverse shrinkage, angular distortion and RMS values of welding current and voltage signals, corresponding to different experiments, are shown in Table 2. The recorded actual welding current and voltage signals corresponding to the experimentally obtained maximum and minimum angular distortions (Exp. Nos. 7 and 26, respectively) are given in Fig. 4(a) and Fig. 4(b), respectively.

![Fig.4: Part of the arc signals corresponding to (a) maximum (Exp. No. 7) and (b) minimum (Exp. No. 26) angular distortions.](image)

Note that in Exp. No. 7 (Fig.6(a)), there is higher voltage, current and frequency as compared to Exp. No. 26 (Fig.6(b)). However, these qualitative states do not necessarily correlate to the output quality. The current signal in Fig.6(a) is approximately proportional to the voltage signal, whereas the signals in Fig.6(b) do not show this trend. This means that the arc in Exp. No. 7 is more stable as compared to that in Exp. No. 26.

The experimental outputs have brought out the fact that having same process parameters (Exp. Nos. 2, 8, 21, 24, 25, 28, 38, 46, 47) does not always result in the same output quality. Note that for these experiments with same process parameters, the variation in actual angular distortion and transverse shrinkage are found to be 21.43% and
22.69%, respectively. Although the same process parameters setting were used in the mentioned experiments, the arc signals showed major discrepancies (in stability of the arc), thereby producing grossly varying output quality. This is why sensor signals (current and voltage, in this case) were used as inputs in the developed models with the hope that it would rightly predict the output quality.

5. Prediction of weld bead geometry

From the arc signal analysis, it is well understood that arc signals can be used to detect welding distortion. Therefore, in the present work, the influence of these signals on the weld quality was investigated by considering them in the input of the RBFN models.

5.1. Prediction of welding distortion by using RBFN model

The RBFN models were trained by using first 45 input-output pairs (Table 2) in a batch mode. The initial RBF centers were chosen randomly from the input space, weight values were chosen randomly between 0 to 1, and the bias value at the input and output layers were taken as zero and one, respectively. All the input and output variables were normalized between 0.1 and 0.9. The training objective was the mean square error (MSE) minimization through training algorithms as mentioned in section 3.2. The performance of a RBFN depends on the dimension of hidden space, learning rate and momentum coefficient. Therefore, several combinations were tried out to choose an optimal combination. In this work, the number of RBFs was varied from 5 to 25, and learning rate was varied between 0.05 to 0.5. After training the network, eight remaining dataset were used to test the network performance. The best performance in training and testing for six different cases mentioned in section 3.3 are shown in Table 3.

<table>
<thead>
<tr>
<th>Case No.</th>
<th>Best architecture</th>
<th>MSE of training</th>
<th>MSE of testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>8-19-2</td>
<td>0.02775347</td>
<td>0.014697</td>
</tr>
<tr>
<td>Case 2</td>
<td>8-16-2</td>
<td>0.00138553</td>
<td>0.006212</td>
</tr>
<tr>
<td>Case 3</td>
<td>8-20-2</td>
<td>0.01322712</td>
<td>0.009621</td>
</tr>
<tr>
<td>Case 4</td>
<td>8-18-2</td>
<td>0.0054645</td>
<td>0.0062315</td>
</tr>
<tr>
<td>Case 5</td>
<td>8-17-2</td>
<td>0.0024156</td>
<td>0.00848368</td>
</tr>
<tr>
<td><strong>Case 6</strong></td>
<td><strong>8-16-2</strong></td>
<td><strong>0.00050538</strong></td>
<td><strong>0.00208832</strong></td>
</tr>
</tbody>
</table>

Among all the six cases, case 6, i.e. sigmoid activation function for the output layer with full training algorithm, was found to be the best. This is in agreement with the
observation made by Taghi et al.\textsuperscript{34}, where it has been shown that full-training algorithm has superior prediction performance.

The testing performance of the best case (case 6) is shown in Table 4, where the predicted values and the percentage errors in distortion are shown for the 8 test data.

Table 4. RBFN predicted distortion with percentage error (for 8-16-2 architecture).

<table>
<thead>
<tr>
<th>Exp. No.</th>
<th>predicted $\theta$</th>
<th>predicted $\rho$</th>
<th>% error in prediction of $\theta$</th>
<th>% error in prediction of $\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td>46</td>
<td>1.379323</td>
<td>0.001602</td>
<td>3.49</td>
<td>3.19</td>
</tr>
<tr>
<td>47</td>
<td>1.26436</td>
<td>0.00144</td>
<td>3.88</td>
<td>-9.45</td>
</tr>
<tr>
<td>48</td>
<td>1.445513</td>
<td>0.001721</td>
<td>-8.76</td>
<td>-3.96</td>
</tr>
<tr>
<td>49</td>
<td>0.657345</td>
<td>0.001441</td>
<td>5.41</td>
<td>4.47</td>
</tr>
<tr>
<td>50</td>
<td>0.742358</td>
<td>0.00171</td>
<td>-2.83</td>
<td>5.44</td>
</tr>
<tr>
<td>51</td>
<td>0.85653</td>
<td>0.001966</td>
<td>5.67</td>
<td>-1.98</td>
</tr>
<tr>
<td>52</td>
<td>1.220676</td>
<td>0.001481</td>
<td>8.11</td>
<td>2.42</td>
</tr>
<tr>
<td>53</td>
<td>0.8685</td>
<td>0.001883</td>
<td>6.37</td>
<td>3.13</td>
</tr>
</tbody>
</table>

Based on the results given in Tables 4 and 5, the following observations can be made.

1) The RBFN architecture 8-16-2, with learning rate of 0.05 and sigmoid activation function for the output layer with full training algorithm gives the lowest prediction error. The average prediction error on angular distortion and transverse shrinkage are 5.56% and 4.25%, respectively.

2) The best training algorithm is full training, because it utilizes the full capacity of the network. However, note that the convergence rate is smaller in all full training algorithm cases.

3) Sigmoid activation function in the output layer is better than the other two.

6 Conclusions

Welding distortion in a PMIGW process is predicted in this work from process parameters and measured arc signals by using six different radial basis function network models. A series of experiments were carried out by applying response surface method, which evenly distributed the process parameters over the parameter space in the operating range. The experimentally obtained data were then used to train and test the different types of RBFN models of various architectures. The testing phase revealed that a RBFN model of specific training case and architecture is well-suited for prediction of weld distortion in a PMIGW process as compared to other training cases and architectures.
Furthermore, the experimental outputs have brought out the fact that having same process parameters does not always result in the same output quality. This is why, arc signals were used as inputs in the developed models and then it was found that the resulting model gave good prediction over the considered domain. Thus, it may be concluded that in the industrial scenario, in addition to giving the desired process parameter settings, online monitoring of process through available sensors may be beneficial to ensure proper weld quality.

References


