A GA–ANN HYBRID MODEL FOR PREDICTION AND OPTIMIZATION OF CO₂ LASER-MIG HYBRID WELDING PROCESS

S. Chaki¹* and S. Ghosal²

¹Department of Automobile Engineering, MCKV Institute of Engineering, 243, G.T. Road (N), Liluah, Howrah-711204, West Bengal, India  
²Email: sudiptochaki@hotmail.com  
Phone: +91 33 2654 9315/17; Fax: +91 33 2654 9318

ABSTRACT

The paper presents a hybrid model of an Artificial Neural Network (ANN) and Genetic Algorithm (GA) for modeling of a hybrid laser welding process. This model is employed for the prediction and optimization of penetration depth with corresponding process parameters. A single program developed for the purpose initially establishes an optimized ANN architecture using a Back-Propagation Neural Network (BPNN), with Bayesian regularization (BR). This trained ANN is then used in conjunction with GA to find optimum parameters for the process. An experimental dataset obtained from published literature has been employed to study the effect of process input parameters (arc power, focal distance from the workpiece surface, torch angle and the distance between the laser and the welding torch) in CO₂ laser–MIG hybrid welding on the depth of penetration for 5005 Al–Mg alloy, and has been used in the present work for the purpose of training, testing and optimization of the GA–ANN model. The results indicate that the GA–ANN model can predict the output with reasonably good accuracy (mean absolute % errors of 0.7198%) and can optimize the process parameters with a negligible computational time of 100.09 s. The proposed approach is envisaged for application in a multi-variable complex problem with reasonable accuracy for prediction and optimization of operational parameters. A 4-7-1 network trained using BPNN with the BR method has been found to show the best prediction capability and a maximum penetration depth of 3.84 mm has been obtained during optimization.

Keywords: GA–ANN hybrid model; CO₂ laser-MIG hybrid welding; back-propagation neural networks.

INTRODUCTION

In recent years, the CO₂ laser–MIG hybrid welding method has emerged as a popular tool in the shipbuilding, transport and aerospace industries [1]. This method combines the advantages of laser beam welding, such as high welding speed, deep weld penetration and minimal distortion by heat, with the large gap-bridging ability and cost-effectiveness of conventional GMAW processes to produce a more efficient welding tool [2-4]. However, as the operational parameters and weld properties have a complex relationship, a proper computer-based model is needed for online control of weld properties and optimization of the process parameters. Apart from some conventional statistical analysis [5-8], little effort has been observed so far on soft computing
technique-based modeling and optimization of process parameters. In the absence of such modeling, the present research work has developed a hybrid model (GA–ANN) involving an artificial neural network (ANN) and genetic algorithm (GA) for the prediction and optimization of the penetration depth of CO₂ laser–MIG hybrid welding.

ANN is a relatively new computational tool that has been applied in a wide spectrum of problems for modeling complex non-linear relationships among multiple input variables and the output(s) [9-12]. The multilayer feed forward neural network (MFFNN), also known as the ‘universal function approximator’ [13], is attractive to researchers for such parametric modeling due to its inherent capability to approximate the underlying function of a given data set to any arbitrary degree of accuracy. Researchers have recently used the back-propagation neural network (BPNN) algorithm, a popular variant of MFFNN [14], for modeling of many laser and allied application processes such as laser cutting [15], CO₂ laser welding [16] and electric discharge machining (EDM) [17, 18]. ANN has also been employed to model heat transfer and pressure drop prediction in an in-line flat tube bundle [9], for forward kinematics solution of a 6-6 SPU parallel manipulator [19], in the investigation of fossil fuel and liquid biofuel blend properties [12], and for grinding of ductile cast iron using water-based SiO₂ nanocoolant [10]. M.N. Yahya et al. [11] have combined a gray-level co-occurrence matrix (GLCM) and a neural network (NN) to identify the absorption coefficient and dimensions of rooms. But, for a small and noisy experimental dataset, BPNN with the Bayesian regularization (BR) technique [20-22] is found to show better prediction capability compared to LM or gradient-descent BPNN. However, application of BR in laser and allied fields of research has not been observed by the authors so far. As the present problem deals with a relatively small number of experimental datasets, BPNN with BR is thought to be a suitable algorithm for ANN modeling.

Genetic algorithms are derivative-free stochastic optimization methods based loosely upon the concept of natural selection and natural genetics [23]. GA has become popular due to its lesser tendency to get trapped in local minima and is generally expected to find a global solution because it works with a population instead of a single point as in traditional optimization methods. Researchers have already employed GA with a user-defined objective function for optimization of the process parameters in laser cutting [24], laser welding [25], laser forming [26], and in other advanced machining processes including ultrasonic machining (USM), abrasive jet machining (AJM), water jet machining (WJM) and abrasive-water jet machining (AWJM) [27]. Ghosal and Chaki [28] have employed an ANN–Quasi Newton hybrid model for estimation and optimization of depth of penetration in hybrid CO₂ laser-MIG welding. Recently, Kumar et al. [29] have employed an integrated GA–Taguchi model for parametric optimization of the submerged arc welding process. In the present work, a hybrid model of GA–ANN is proposed and implemented for prediction and optimization of operational parameters in any given domain of experimental parameters. The uniqueness of the model is that it does not require formulation of a separate objective function for GA optimization as it uses a trained ANN to calculate the value of the objective function. The hybrid methodology combines a single hidden layer BPNN with BR and GA in a single program for prediction and optimization of the penetration depth of CO₂ laser–MIG welding.
GA–ANN HYBRID MODEL FOR PREDICTION AND OPTIMIZATION

The GA–ANN hybrid model developed in the present work can be employed for training, prediction and optimization of the operational parameters of an experiment by running a single program in the MATLAB 7.0 environment. Two separate data files are used for storing training and testing data of ANN. The program first employs a user-defined ANN training algorithm to perform the task of training and subsequent testing to assess the function approximation and prediction capability of a particular ANN architecture. A number of ANN architectures are studied in this way by randomly varying the number of hidden layer neurons, and the architecture with the maximum prediction accuracy is considered as the best ANN. Next, the program calls the subroutine of GA and starts iteration with the initial population. As no defined objective function (theoretically connecting input parameters and output) exists for the experimental dataset, the program control switches from the GA subroutine to the ANN module in the main program and employs the best ANN, as determined earlier, to generate the value of the objective function corresponding to the initial population generated by the GA. The program control then switches back to the GA subroutine and the cycle continues up to the point of appropriate convergence. The program completes the training, prediction and optimization in a single run. A schematic diagram for the present method is given in Fig. 1. The working of ANN and GA for the present problem is furnished in some detail as follows.

Figure 1. The logical flow diagram of the GA–ANN hybrid model.
WORKING OF ANN

Dataset

The present work has employed the experimental data obtained from the work of Casalino [7], where a 3 kW CO₂ laser–MIG welding setup has been used for producing 20 butt joints through the random combination of input parameters in 5005 aluminum alloy (0.6% magnesium). A schematic diagram of the CO₂ laser–MIG welding setup is shown in Fig. 2, which is a combination of a CNC programmable CO₂ laser system and a MIG welding gun. The input controllable parameters are power (P in W), focal distance (F in mm) from the workpiece surface, torch angle (A in deg.) and the distance between the laser and the welding torch (S in mm). As the main objective of the hybrid welding is to increase the weld penetration, the depth of weld penetration (D in mm) is measured as output to assess the effect of the input parameters on it. The range and levels of experimental input data and corresponding output are given in Table 1.

![Schematic diagram of experimental setup](image)

**Figure 2. Schematic diagram of experimental setup [7].**

**Table 1. Range and levels of experimental parameters.**

<table>
<thead>
<tr>
<th>Factors</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Minimum value</th>
<th>Maximum value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental inputs</td>
<td>S (mm)</td>
<td>5</td>
<td>10</td>
<td>20</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>P (W)</td>
<td>900</td>
<td>1050</td>
<td>1200</td>
<td>900</td>
</tr>
<tr>
<td></td>
<td>F (mm)</td>
<td>0</td>
<td>2.5</td>
<td>3.5</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>A (°)</td>
<td>45</td>
<td>60</td>
<td>60</td>
<td>45</td>
</tr>
<tr>
<td>Experimental output</td>
<td>D (mm)</td>
<td>1.32</td>
<td></td>
<td>3.26</td>
<td></td>
</tr>
</tbody>
</table>

**ANN Training and Testing**

In the present work, the training algorithm using a single hidden layer BPNN with BR consists of the ANN architecture with four neurons for input parameters in the input layer, one hidden layer and one neuron for output in the output layer, as given in Figure 3. The numbers of hidden layer neurons are varied randomly and altogether six...
different single hidden layer architectures have been trained and subsequently tested in the present study. Based on the prediction capability, the best of the ANN architectures is selected and used for the latter part of the program. Out of 20 sets of experimental data [7], 16 datasets have been selected for ANN training purposes and the remainder are used for ANN testing. The training and testing datasets are furnished in Table 2. For the present problem, the input \(X\) and output \(T\) vector fed to the input and output node for ANN training and subsequent testing are given by

\[
X = [S \ P \ F \ A]
\]

\[
T = [b] \text{(1)}
\]

Figure 3. Architecture of the back-propagation network.

It is observed during some pilot training sessions that, beyond 300 epochs (iterations), error decreases only asymptotically. Thus, any further increase in epoch number would increase the training time with very little reduction in error. So, in the present study, each network is trained over 300 epochs only. All the neurons in the input, hidden and output layers bear weighted connections. Activation functions for the hidden and output layers are sigmoidal and linear respectively. Traditional BPNN minimizes the error, which is the difference between the network output \(O\) and desired output value \(T\). However, in order to improve the generalization (or prediction) capability, BR minimizes \(\Phi\), which is a linear combination of the sum of the squared errors (SSE) and the sum of the squared weights (SSW).

A brief working of the ANN training module with BR is given below:

**Initialize:** Weights by generating random number,

Normalize input and output dataset between 0 and 1:

\[
X_{\text{nor}} = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}}, \quad T_{\text{nor}} = \frac{T - T_{\text{min}}}{T_{\text{max}} - T_{\text{min}}}
\]

(1)

where \(X_{\text{max}}, X_{\text{min}}\) and \(T_{\text{max}}, T_{\text{min}}\) are maximum and minimum real values of input and output.
Set regularization parameters $\alpha=0$ and $\beta=1$.
Set number of hidden layer neurons ($n$), error tolerance and training epoch ($K$).

Iterate:

I. Compute output of hidden layer ($Z_j$):
Activation value ($N_j$) for $j$-th neurons of the hidden layer:

$$N_j = \mathbf{w}_j^T \mathbf{x}, \quad j=1,2,\ldots,n, \quad \mathbf{w}_j = [w_{j1} \ w_{j2} \ \cdots \ w_{jm}]$$

$$Z_j = f(N_j) = \frac{1}{1+e^{-N_j}}$$

II. Compute output signal ($O_j$) in similar way to linear activation function

III. Compute sum of the square errors (SSE):

$$SSE = \sum (T-O)^2$$

IV. Compute sum of the squared weights (SSW):

$$SSW = \sum w_j^2$$

V. Update regularization parameters $\alpha$ and $\beta$

VI. Weights are updated by using Bayes’ rule of conditional probability.

VII. Compute objective function ($\Phi$):

$$\Phi = \beta \times SSE + \alpha \times SSW$$

Continue (until $\Phi^{k+1} < \Phi^k$ < error tolerance)

Table 2. Experimental dataset for ANN training and testing.

<table>
<thead>
<tr>
<th>Exp. No.</th>
<th>Training input data</th>
<th>Desired training output</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S (mm)</td>
<td>P (W)</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>1200</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>1050</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>1200</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>900</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>900</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>900</td>
</tr>
<tr>
<td>7</td>
<td>20</td>
<td>1050</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>1050</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>1200</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>1050</td>
</tr>
<tr>
<td>11</td>
<td>10</td>
<td>1050</td>
</tr>
<tr>
<td>12</td>
<td>5</td>
<td>900</td>
</tr>
<tr>
<td>13</td>
<td>20</td>
<td>900</td>
</tr>
<tr>
<td>14</td>
<td>20</td>
<td>900</td>
</tr>
<tr>
<td>15</td>
<td>20</td>
<td>1200</td>
</tr>
<tr>
<td>16</td>
<td>20</td>
<td>1200</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Testing input data</th>
<th>Desired testing output</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>
Upon convergence, the ANN training module returns the training sum of squared errors (SSE) and the final weights for the hidden layer and output layer as output, which constitute the trained ANN architecture. The trained network generates an approximate function from the input dataset and the trained ANN is further worked with some test input data to test its prediction capability. The ANN output thus produced is compared with corresponding known experimental output to determine the prediction error. The prediction capability of the particular architecture is measured by the SSE. The architecture that shows the highest accuracy in prediction is considered as the best ANN. This best ANN then interfaces with the GA subroutine of the model for optimization (Figure 2).

WORKING OF GA

In the present model, GA is introduced after completion of ANN training and testing. During GA optimization initially a population is generated randomly, and is further modified over iterations by subsequent operations like objective function evaluation from ANN, fitness scaling to obtain the fitness function value, selection of above average population members and generation of new population members using crossover and mutation. The computation stops if there is no further improvement in the best fitness value for a certain number of consecutive iterations (or reproduction of new generations) cycles. This termination number of iterations is called ‘stall generations’. The optimization problem can be formulated as:

Maximize $D(S, P, F, A)$

Subject to the constraints:

\[ 5 \leq S \leq 20 \]
\[ 900 \leq P \leq 1200 \]
\[ 0 \leq F \leq 3.5 \]
\[ 45 \leq A \leq 60 \]  

As during ANN training the input parameters are normalized within a range of 0 and 1, the range of constraining variables becomes 0–1 for the purpose of optimization. Moreover, as the algorithm was meant for finding the minima of a function, it is converted into the present maximization problem by minimizing the objective function $-f = D(S, P, F, A)$. Thus, after incorporating these modifications the optimization problem is formulated as:

Minimize $-D(S, P, F, A)$

Subject to the constraints: $0 \leq S, P, F, A \leq 1$  

GA in the present model works according to the following steps:

Set GA Parameters: Population size ($N$), crossover fraction ($p$), mutation rate ($m$), stall generation number ($n$) and maximum number of generations ($t_{max}$)

Initialize Population: Randomly generate the initial population of N individuals or strings. Each string contains four substrings that indicate the constraining variables.
Iterate:

I) Objective Function Evaluation:
Values of substrings in every string of the initial population are sent to ANN module. The trained ANN uses the substring values as input to predict output parameter and returns predicted output as objective function value to GA module.

II) Fitness Scaling:
Convert the objective function values to the scaled values within a range, known as fitness function values ($f_i$).

III) Selection or Reproduction:
Identify good (above average) solutions in a population, generate multiple copies of good solutions and eliminate bad solutions from population proportionately by keeping population size constant. If the average fitness of all population members is $f_{avg}$, a solution with a fitness $f_i$ gets an expected $f_i/f_{avg}$ number of copies in the mating pool.

IV) Crossover:
Randomly pick up two parent strings from the mating pool and swap the part of strings between two randomly selected crossover points to generate two new strings or children with crossover fraction or probability $p$.

V) Mutation:
Generate new strings known as mutation children by small random changes in the individuals in the modified population after reproduction and crossover with mutation rate of $m$.

Continue (until, stall generation $>$ n)

GA operators as explained and corresponding set values / options for the present study are furnished below:

Size of populations: 100
Number of stall generations: 30
Fitness function: Rank scaling
Selection function: Roulette wheel
Crossover function: Two point
Crossover fraction: 0.8
Mutation function: Uniform
Mutation rate: 0.04

Normalized minimum value of objective function ($D$) thus obtained and corresponding optimum operational input parameters ($S$, $P$, $F$ and $A$) are post-processed to get de-normalized or actual dimensional values of the physical variables concerned.

RESULTS AND DISCUSSION

Performance of ANN Training and Testing

Experimental datasets arbitrarily categorized as input and output with regard to training and testing of ANN are normalized and sent to six different ANN architectures. The performance of the network architectures in terms of training and testing efficacy is given in Table 3. Prediction capability being the primary objective of a trained ANN, it is felt that the performance of a particular ANN during testing with test data should be the yardstick for selecting the best ANN architecture. It is clear from Table 3 and Fig. 4
that the best testing performance is shown by the 4-7-1 architecture with test SSE 2.02E-03. Therefore, the 4-7-1 architecture is considered as the best ANN for the present problem due to its superior prediction capability, and it is employed further for interfacing with GA for optimization.

Table 3. Training and testing performance of different network architectures using BPNN with BR

<table>
<thead>
<tr>
<th>Sl. no.</th>
<th>Hidden layer neurons</th>
<th>Time elapsed (s)</th>
<th>Training SSE</th>
<th>Testing SSE</th>
<th>MSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>4.750</td>
<td>1.66E-04</td>
<td>2.75E-03</td>
<td>1.10E-02</td>
<td>4.74E-02</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>4.985</td>
<td>4.94E-05</td>
<td>3.24E-03</td>
<td>1.29E-02</td>
<td>5.14E-02</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>4.531</td>
<td>1.27E-08</td>
<td>5.06E-04</td>
<td>2.02E-03</td>
<td>2.03E-02</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>4.250</td>
<td>4.65E-02</td>
<td>4.55E-03</td>
<td>2.52E-02</td>
<td>1.81E-02</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>5.015</td>
<td>3.65E-01</td>
<td>3.55E-02</td>
<td>1.42E-01</td>
<td>1.71E-01</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>5.047</td>
<td>3.65E-01</td>
<td>3.55E-02</td>
<td>1.42E-01</td>
<td>1.71E-01</td>
</tr>
</tbody>
</table>

Figure 4. Variation of SSE with hidden layer neurons during ANN testing.

The prediction capability of a particular ANN is assessed by calculating absolute % error in prediction for every test data after corresponding de-normalization, as follows:

\[
\text{Absolute % error in prediction} = \frac{\text{Experimental output} - \text{Predicted output in ANN}}{\text{Experimental output}} \times 100 \quad (4)
\]

During testing with test data it is observed from Table 4 that the 4-7-1 network can reduce mean absolute % error during prediction of the penetration depth by up to 0.7198%, while the corresponding maximum absolute % error is 1.1270%. Therefore,
the order of magnitude of the errors indicates the capability of ANN to predict cutting quality with fairly good accuracy for a set of unknown input parameters. This good prediction capability coupled with the very negligible training time has justified the use of ANN as a powerful tool for estimation of the penetration depth of CO₂ laser–MIG welding for any operational parameter setting.

Table 4. Performance of 4-7-1 networks during testing with testing input data

<table>
<thead>
<tr>
<th>Exp. no.</th>
<th>Experimental output</th>
<th>Predicted output with ANN testing</th>
<th>Absolute error</th>
<th>Absolute % error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.26</td>
<td>3.2368</td>
<td>0.0232</td>
<td>0.7117</td>
</tr>
<tr>
<td>2</td>
<td>3.07</td>
<td>3.0442</td>
<td>0.0258</td>
<td>0.8404</td>
</tr>
<tr>
<td>3</td>
<td>1.85</td>
<td>1.8463</td>
<td>0.0037</td>
<td>0.2000</td>
</tr>
<tr>
<td>4</td>
<td>2.52</td>
<td>2.5484</td>
<td>0.0284</td>
<td>1.1270</td>
</tr>
<tr>
<td>Mean absolute % error</td>
<td></td>
<td></td>
<td></td>
<td>0.7198</td>
</tr>
</tbody>
</table>

Optimization with GA

In the present study, the best ANN with 4-7-1 architecture is used for computing the objective function value, i.e. the penetration depth of CO₂ laser–MIG welding, during iterations of GA optimization. The search ended after 124 generations when the ‘stall generation’ criterion was reached. Fig. 5 shows the plot of the best function value in each generation versus the generation number as well as the convergent nature of the problem. A summary of the results of the optimization is given in Table 5. The output of GA optimization provides the maximum penetration depth of 3.84 mm. This can be achieved with corresponding values of input operational parameters S, P, F and A with values of 5.04 mm (level 1), 1198.979 W (level 3), 3.49 mm (level 3) and 45.06° (level 1) respectively, which can be viewed as one possible combination of level values (Table 1). However, the available experimental dataset (Table 2) does not include this combination of input parameter values. But welding carried out with these parameters is expected to produce a joint with the maximum penetration. But the depth of penetration produced (Table 3) in GA optimization is far better than that obtained through experimentation (Table 2). Therefore, it can be said that the model is capable of optimizing the operational parameters of an experimental study and can be employed further for online implementation. The total computational time of the GA–ANN model, combining training, prediction and optimization, is also found to be 100.09 seconds only on a desktop Pentium IV, 3 GHz and 512 MB PC.

Table 5. Results of GA optimization.

<table>
<thead>
<tr>
<th>Output parameter</th>
<th>Values and unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimum penetration (D)</td>
<td>3.835 mm</td>
</tr>
<tr>
<td>Operational parameters</td>
<td></td>
</tr>
<tr>
<td>Distance between the laser and the welding torch (S)</td>
<td>5.044 mm</td>
</tr>
<tr>
<td>Power (P)</td>
<td>1198.979 W</td>
</tr>
<tr>
<td>Focal distance from the workpiece surface (F)</td>
<td>3.494 mm</td>
</tr>
<tr>
<td>Torch angle (A)</td>
<td>45.06°</td>
</tr>
</tbody>
</table>
A GA–ANN hybrid model for prediction and optimization of CO₂ laser-MIG hybrid welding process

CONCLUSIONS

In the present work a GA–ANN hybrid model is developed and employed for prediction and optimization of the penetration depth of a CO₂ laser-MIG butt welding process for 5005 Al–Mg alloy. During ANN training of the model, of the different ANN architectures used, 4-7-1 has been found to be the most efficient for the BPNN with BR method and is employed for subsequent GA optimization. The ANN model is found to show reasonable accuracy (mean absolute % errors of 0.7198%) during prediction of output. Results for an optimization study to maximize penetration corroborate well the levels of experimental input parameter values. Moreover, it is observed that the optimized penetration depth (3.84 mm) is considerably greater than the maximum penetration depth obtained (3.26 mm) in available experimental data. However, its validation would need additional experimentation. The total computational time of the GA–ANN model is negligible (100.09 s). Though the hybrid model has been developed for single objective optimization, it can be easily modified for solving multi-objective optimization problems as well.

ACKNOWLEDGMENTS

The authors would like to thanks to MCKV Institute of Engineering and Jadavpur University for financial assistance and laboratories facilities.

REFERENCES


A GA–ANN hybrid model for prediction and optimization of CO$_2$ laser-MIG hybrid welding process


