Summary of my PhD project "Using Artificial Neural Network Modeling to design and Develop a Novel Manufacturing Process"

I've used neural network modeling to improve and industrialize a novel welding technique Indeed, I developed predictive models using a relatively novel approach which was the integration of neural network modeling and design of experiments (DOE) using the Taguchi method to optimize the process parameters. It resulted in boosting the production performance by 9 times at only a small fraction of the cost of available techniques in the market.

Some of my PhD achievements:

Journal paper # 1

Manufacturing and Materials Processing

MDPI

Article

Effect of Tool Design and Process Parameters on Lap Joints Made by Right Angle Friction Stir Welding (RAFSW)

Mahboubeh Momeni and Michel Guillot *

PI2/REGAL Research Team, Department of Mechanical Engineering, Laval University, Québec, QC G1V 0A6, Canada
* Correspondence: mguillot@gmc.ulaval.ca; Tel.: +1-(418)-998-6549 or +1-(418)-656-3343

check for

Journal paper # 2

The International Journal of Advanced Manufacturing Technology (2018) 99:3077–3089
https://doi.org/10.1007/s00170-018-2672-8

ORIGINAL ARTICLE

CrossMark

Development of friction stir welding technique at right angle (RAFSW) applied on butt joint of AA6061-T6 aluminum alloy

Mahboubeh Momeni[1] · Michel Guillot[2]

Received: 8 May 2018 / Accepted: 5 September 2018 / Published online: 19 September 2018

Journal paper # 3

https://doi.org/10.1007/s00170-020-05764-2

ORIGINAL ARTICLE

Check for updates

Implementation of right angle friction stir welding (RAFSW) to assemble the side panels of truck box

Mahboubeh Momeni[1] · Michel Guillot[2]

Received: 3 February 2020 / Accepted: 13 July 2020 / Published online: 12 August 2020

An overview of the modeling process in the PhD project

Background

The relationship between the tool design, process parameters, and mechanical properties of the FSW joints is nonlinear and extremely complex. Thus, it is unlikely to have effective analytical expressions to describe these relations. To optimize the tool design and process parameters, experiments are too costly and time-consuming. Artificial intelligence (AI) systems such as artificial neural network (ANN) modeling can be employed to develop the FSW technique while the needed number of experiments and time is reduced, considerably. The excellent capability of neural networks (NNs) makes them a great choice to learn a nonlinear mapping from training data sets, recognize the patterns between input and output parameters without any prior assumptions about their nature, and to generalize the results. ANN modeling can be employed to model different kinds of manufacturing techniques such as FSW processes. It is capable of predicting FSW forces and joint mechanical properties made at different process parameters without considering simplification and assumptions for the process

Theory

A multi-layer perceptron network generally consists of N inputs, M outputs, and L-1 hidden layers and an output layer, as shown in Figure. Adding layers to the perceptron network brings extremely great modeling power. Indeed, the single-layer perceptron only allows modeling of relationships in which variables are linearly separable. When two layers (a hidden layer and an output layer) are used, it is possible to model hypersurfaces. For such a two-layer network, the number of neurons in the hidden layer must be large enough to correctly represent the desired shape. The heuristic below indicates the way of carrying out learning by back-propagation according to the method of the delta [123].

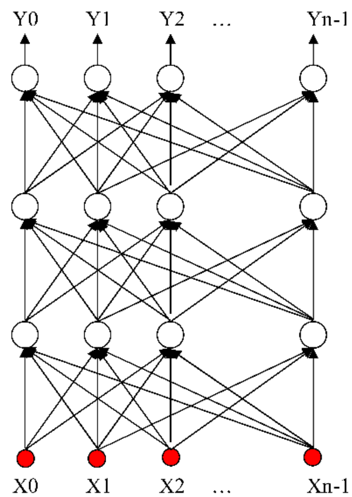

Figure1: The schematic of multilayer perceptron

The learning algorithm of a multi-layered perceptron (back-propagation) is as follows:

1) Initialize the weights and deviations

w(0) = small random values generally between -1 and 1.

q(0) = small random values generally between -1 and 1.

2) Present the data X_i at the network input and estimate the current output Y_{est}. In the first hidden layer adjacent to the inputs X_i, we calculate:

$$I_j = \sum_i w_{ij} X_i \quad (.1)$$

$$O_j = a_j = \frac{1}{\left[1 + \exp\left(-\left(I_j + \Theta_j\right)\right)\right]} \quad (.2)$$

In the intermediate hidden layers:

$$I_j = \sum_i w_{ij} O_i \quad (.3)$$

$$O_j = a_j = \frac{1}{\left[1 + \exp\left(-\left(I_j + \Theta_j\right)\right)\right]} \quad (.4)$$

In the output layer:

$$I_j = \sum_i w_{ij} O_i \tag{.5}$$

$$Y_{est\ j} = a_j = \frac{1}{\left[1 + \exp\left(-\left(I_j + \Theta_j\right)\right)\right]} \tag{.6}$$

3) Evaluate the necessary correction values in the network from the output error. These delta values become: In the output layer:

$$\delta_{Sortie\ \mathrm{j}} = \varepsilon Y_j \left(1 - Y_j\right) \tag{.7}$$

$$\varepsilon = Y_{des\ \mathrm{j}} - Y_j \tag{.8}$$

In the intermediate layers:

$$\delta_j = O_j (1 - O_j) \sum_k \delta_k w_{kj} \tag{.9}$$

4) Adjust the synaptic coefficients according to the equation:

$$w_{ij}(n+1) = w_{ij}(n) + \eta \delta_j O_i + \alpha \left(w_{ij}(n) - w_{ij}(n-1)\right) \tag{.10}$$

Where O_i will take the value X_i on the layer of neurons adjacent to the inputs.

5) Adjust the deviations according to the equation:

$$\theta_j(n+1) = \theta_j(n) + \eta \delta_j + \alpha \left(\theta_j(n) - \theta_j(n-1)\right) \tag{.11}$$

6) Repeat steps 2 to 5 for all the data in the training file.

7) Repeat steps 2 to 6 for as many passes as necessary to reduce the sum of the errors in the squared training data below a prescribed value.

Hidden and output layers with “sigmoid” transfer function were used to predict. The sigmoid transfer function was:

$$F(x) = \frac{1}{1 + e^{-x}} \tag{.12}$$

Where x is the weighted sum of the input [123].

Several factors affect the prediction accuracy of the backpropagation neural network models such as the architecture of the network, momentum coefficient, and learning rate of the model [73, 74, 123]. The factors to compare the different architectures are root-mean-squared error (RMSE), maximum error, mean relative error (MRE), and mean absolute error (MAE) [76, 123]. Furthermore, the relationship between the experimental data and predicted values by the developed ANN models has been studied by calculating the amount of correlation coefficient (R^2) using linear regression analysis. When the correlation coefficient is close to one, it indicates a close relationship between the experimental data and predicted data by the developed ANN models [73, 76, 77].

Mean relative error:

$$MRE = \frac{1}{N}\sum_{i}^{N}\left(\frac{|A_i - Y_i|}{A_i}\right)\times 100 \qquad (.13)$$

Root-mean squared error:

$$RMSE = \left(\frac{1}{N}\sum_{i}^{N}(A_i - Y_i)^2\right)^{1/2} \qquad (.14)$$

Mean absolute error:

$$MAE = \frac{1}{N}\sum_{i}^{N}|A_i - Y_i| \qquad (.15)$$

Maximum error:

$$Max.E = \mathbf{max}(|A_i - Y_i|) \qquad (.16)$$

Absolute fraction of variance:

$$R^2 = 1 - \left(\frac{\sum_{i}^{N}(A_i - Y_i)^2}{\sum_{i}^{N} Y_i^2}\right) \qquad (.17)$$

Where A_i, Y_i, N, and i are experimental value, predicted value, total number of experimental data, and trial number, respectively. The formulas are from Ref.[76]

General Methodology

Nowadays, Taguchi method is extensively used to optimize manufacturing processes [72]. Taguchi design of experiments (DOE) is very effective to minimize the needed number of experiments to investigate the effect of input process parameters on the output results. After conducting the experiments based on the DOEs, artificial neural network (ANN) modeling, instead of analysis of variances (ANOVA), has been used to model complex relationships between inputs and outputs, in this study [72]. Taking advantage of Taguchi method along with high predication capability of ANN, one would be able to predict the relationship between inputs and outputs through minimized number of experiments.

In this research, feed-forward neural networks with backpropagation algorithm has been used to model the relationship between the tool geometry and the process parameters on the downward axial force and the fracture force of the joints. Several factors affect the prediction accuracy of the backpropagation neural network models such as architecture of the network, momentum coefficient, and learning rate of the model [73, 74]. An ANN model with too small architecture can results in insufficient degree of freedom; and too large network causes to over fit the data. Thus, there is an optimized number of neurons and hidden layers to have a reliable ANN model [75]. The factors to compare the different architectures are root-mean squared error (RMSE), maximum error, mean relative error (MRE), and mean absolute error (MAE) [76]. Furthermore, the relationship between the experimental data and predicted values by the developed ANN models has been studied by calculating the amount of correlation coefficient (R^2) using linear regression analysis. When the correlation coefficient is close to one, it indicates a close relationship between the experimental data and predicted data by the developed ANN models [73, 76, 77].

In this research, the ANN modeling is done using a software developed by Prof. Michel Guillot. The author does not have any contribution in the development of this software and have only used it to analyze the data obtained from the experiments. The input parameters to study are the tool design parameters and process parameters. The outputs are downward axial force and either the strength or the failure force of the joint. Some details regarding the theory of the ANN modeling and a case study on the developed ANN models in this research is provided in appendix.

A case study: details of developing ANN models

In this part, the details of developing the ANN models used in chapter 3 is presented. The experimental data obtained in the previous part of this paper were utilized to train artificial neural networks in this section. 36 experiments are separated from 40 experiments as training data for the ANN models; and the 4 remaining tests were utilized as validation tests to evaluate the predictability and the accuracy of the developed ANN models. Feed-forward neural networks with backpropagation algorithm were used to model the relationship between the tool geometry and the process parameters on the downward axial force and the fracture force of the joints. The effect of the architecture of the network, momentum coefficient, and learning rate of the model on the accuracy of the neural network models is studied to find the best ANN modeling factors. An ANN model with too small architecture can result in insufficient degree of freedom, and too large network causes to overfit the data. Thus, there is an optimized number of neurons and hidden layers to have a reliable ANN model [75]. In this research, several numbers of neurons in one hidden layer, 4 to 10 neurons, were used to find the optimal architecture with minimal prediction errors. Table A.1 shows the errors for the validation data for some single hidden layer ANNs with different numbers of neurons in the hidden layer. Accordingly, it was found that a single hidden layer ANN with 8 neurons yields the lowest amount of error without overfitting. Since it shows very good predictability with errors less than 5%, it was selected for the prediction of fracture forces in this part of the research. The same strategy for modeling downward axial forces leads to the selection of 10-8-1 architecture as it provides acceptable predictivity within a small range of errors. **Error! Reference source not found.** illustrates the optimal architecture of the ANN model for both fracture force and downward axial force. **Error! Reference source not found.** indicates the details regarding these models. The learning rate and momentum both was 0.5. The procedure to select them is indicated in pages number 95 and 96 in chapter 5. **Error! Reference source not found.** shows the input and output for trained and validation data for fracture force using a 10-8-1 ANN model. **Error! Reference source not found.** indicates calculated errors and correlation coefficients for trained, validation, and the entire set of data for fracture forces modeled by the 10-8-1 ANN model.

A sample of training data

Sample No.	Sample Code	PL (mm)	SD (mm)	SGD (mm)	PBD (mm)	PA (°)	PLD (mm)	V (mm/min)	w (rpm)	PD (mm)	C	DAF (N)	FF (N)
1	1-a	1.8	8.5	0.1	4	20	0.25	1400	3500	1.8	1	2600	2994
2	1-b	1.8	8.5	0.1	4	20	0.25	1400	3500	1.85	1	3060	4453
3	2-a	1.8	9.5	0.1	4	20	0.45	2000	5000	1.8	2	2500	3390
4	2-b	1.8	9.5	0.1	4	20	0.45	2000	5000	1.85	2	3120	3937
5	3-a	1.8	10.5	0.25	4.8	24	0.25	1400	3500	1.8	2	2900	5373
6	3-b	1.8	10.5	0.25	4.8	24	0.25	1400	3500	1.88	2	4130	5422
7	4-a	1.8	11.5	0.25	4.8	24	0.45	2000	5000	1.8	1	2810	2847
8	4-b	1.8	11.5	0.25	4.8	24	0.45	2000	5000	1.88	1	4320	5218
9	5-a	2.2	8.5	0.1	4.8	24	0.25	2000	5000	2.2	2	3210	3510
10	5-b	2.2	8.5	0.1	4.8	24	0.25	2000	5000	2.25	2	3590	3390
11	6-a	2.2	9.5	0.1	4.8	24	0.45	1400	3500	2.2	1	3420	5218
12	6-b	2.2	9.5	0.1	4.8	24	0.45	1400	3500	2.25	1	4000	5609
13	7-a	2.2	10.5	0.25	4	20	0.25	2000	5000	2.2	1	3200	4502
14	7-b	2.2	10.5	0.25	4	20	0.25	2000	5000	2.28	1	3840	5756
15	8-a	2.2	11.5	0.25	4	20	0.45	1400	3500	2.2	2	3310	2709
16	8-b	2.2	11.5	0.25	4	20	0.45	1400	3500	2.28	2	4530	3292
17	9-a	2.6	8.5	0.25	4	24	0.45	1400	5000	2.6	2	2970	4083
18	9-b	2.6	8.5	0.25	4	24	0.45	1400	5000	2.68	2	3240	3314
19	10-a	2.6	9.5	0.25	4	24	0.25	2000	3500	2.6	1	4100	3358
20	10-b	2.6	9.5	0.25	4	24	0.25	2000	3500	2.68	1	4720	4746
21	11-a	2.6	10.5	0.1	4.8	20	0.45	1400	5000	2.6	1	3170	6107
22	11-b	2.6	10.5	0.1	4.8	20	0.45	1400	5000	2.65	1	3780	5640
23	12-a	2.6	11.5	0.1	4.8	20	0.25	2000	3500	2.6	2	4480	2171
24	12-b	2.6	11.5	0.1	4.8	20	0.25	2000	3500	2.65	2	4700	2300
25	13-a	3	8.5	0.25	4.8	20	0.45	2000	3500	3	1	4320	4982
26	13-b	3	8.5	0.25	4.8	20	0.45	2000	3500	3.08	1	4680	6557
27	14-a	3	9.5	0.25	4.8	20	0.25	1400	5000	3	2	2930	2123
28	14-b	3	9.5	0.25	4.8	20	0.25	1400	5000	3.08	2	3360	2042
29	15-a	3	10.5	0.1	4	24	0.45	2000	3500	3	2	3934	2136
30	15-b	3	10.5	0.1	4	24	0.45	2000	3500	3.05	2	4390	2056
31	16-a	3	11.5	0.1	4	24	0.25	1400	5000	3	1	3040	3581
32	16-b	3	11.5	0.1	4	24	0.25	1400	5000	3.05	1	3750	4141
33	17-a	1.8	8.5	0.1	3.5	20	0.25	1700	4250	1.8	1	2970	3118
34	17-b	1.8	8.5	0.1	3.5	20	0.25	1700	4250	1.835	1	3320	3982
35	18-a	1.75	12	0.3	5	25	0.5	1800	4500	1.75	2	2760	5454
36	18-b	1.75	12	0.3	5	25	0.5	1800	4500	1.83	2	4320	5961
37	19-a	1.75	12	0.3	4.5	25	0.35	2200	5000	1.75	1	3340	2224
38	19-b	1.75	12	0.3	4.5	25	0.35	2200	5000	1.83	1	4600	4043
39	20-a	1.75	11.5	0.3	4.5	25	0.45	2400	5500	1.75	2	3020	4328
40	20-b	1.75	11.5	0.3	4.5	25	0.45	2400	5500	1.83	2	4000	4582

The optimizing process of Neural Network model's architecture

Table 2: The calculated errors for the validation data predicted by some single hidden layer ANNs with different number of neurons in the hidden layer.

10-4-1				
Experimental data (N)	predicted value (N)	error (N)	error (%)	error^2
3982	4020	38	1.0%	1444
2709	3044	335	12.4%	112225
4746	4816	70	1.5%	4900
2123	2537	414	19.5%	171396
				289965
validation data	MAX E	4816		
	RMSE	269		

10-5-1				
Experimental data (N)	predicted value (N)	error (N)	error (%)	error^2
3982	4132	150	3.8%	22500
2709	2940	231	8.5%	53361
4746	4648	98	2.1%	9604
2123	2068	55	2.6%	3025
				88490
validation data	MAX E	231		
	RMSE	149		

10-6-1				
Experimental data (N)	predicted value (N)	error (N)	error (%)	error^2
3982	4124	142	3.6%	20164
2709	2848	139	5.1%	19321
4746	4606	140	2.9%	19600
2123	1957	166	7.8%	27556
				86641
validation data	MAX E	166		
	RMSE	147		

10-7-1				
Experimental data (N)	predicted value (N)	error (N)	error (%)	error^2
3982	4005	23	0.6%	529
2709	2771	62	2.3%	3844
4746	4570	176	3.7%	30976
2123	1924	199	9.4%	39601
				74950
validation data	MAX E	199		
	RMSE	137		

10-8-1				
Experimental data (N)	predicted value (N)	error (N)	error (%)	error^2
3982	4156	174	4.4%	30276
2709	2810	101	3.7%	10201
4746	4876	130	2.7%	16900
2123	2025	98	4.6%	9604
				66981
validation data	MAX E	174		
	RMSE	129		

10-9-1				
Experimental data (N)	predicted value (N)	error (N)	error (%)	error^2
3982	3953	29	0.7%	841
2709	2825	116	4.3%	13456
4746	4862	116	2.4%	13456
2123	1892	231	10.9%	53361
				81114
validation data	MAX E	231		
	RMSE	142		

10-10-1				
Experimental data (N)	predicted value (N)	error (N)	error (%)	error^2
3982	3975	7	0.2%	49
2709	2931	222	8.2%	49284
4746	4991	245	5.2%	60025
2123	1910	213	10.0%	45369
				154727
validation data	MAX E	245		
	RMSE	197		

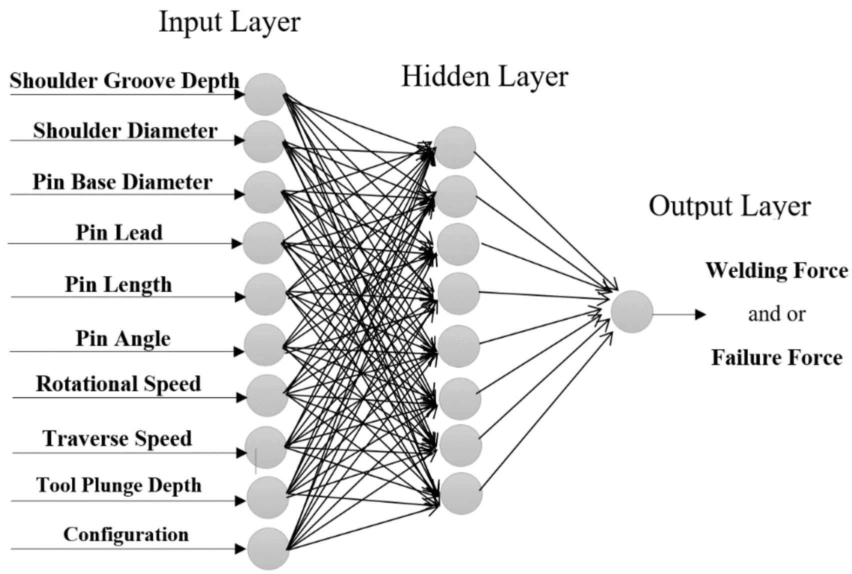

Figure 1: The architecture of ANN models in this paper. The architecture of the model for downward axial force and fracture force is 10-8-1 for both of them.

Visualization of hyperparameter tunning impact on error

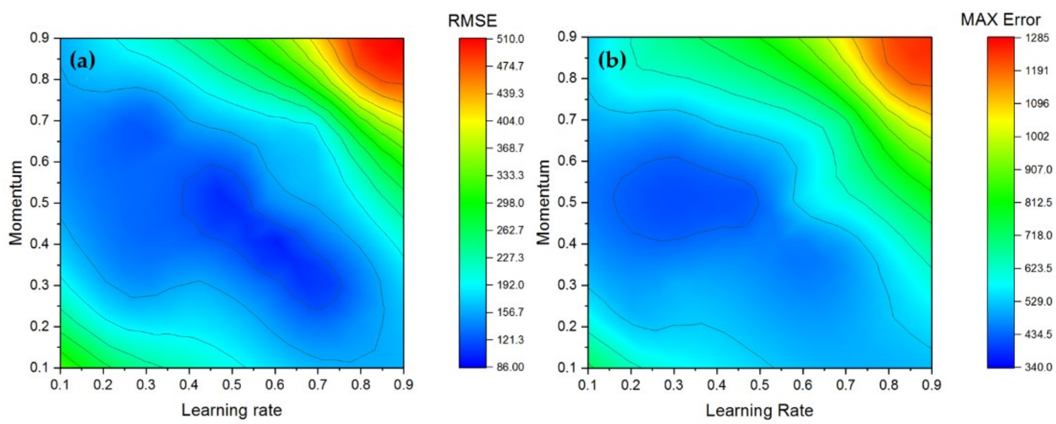

Figure 2: The effect of learning rate and momentum coefficient on (a) root-mean squared error (RMSE) and (b) maximum error when the architecture of the model is 10-8-1.

Statistical analysis of the optimized NN model's errors

Table 1: The RMSE and maximum error for the training data, the confirmation data, and the overall data when the architecture of the neural networks is 10-8-1.

Training Data				**Confirmation Data**				**All Data**			
Downward axial Force		Failure Force		Downward axial Force		Failure Force		Downward axial Force		Failure Force	
RMSE	**Max Error**	**RMSE**	**Max Error**	**RMSE**	**Max Error**	**RMSE**	**Max Error**	**RMSE**	**Max Error**	**RMSE**	**Max Error**
15.97	52	71.49	211	135.76	169	129.53	174	45.53	169	80.96	211

Table 2: The amount of different kind of errors for the developed ANN models.

	RSME	**MAE**	**MRE**	**Maximum error**
formula	$\left(\frac{1}{N}\sum_{i}^{N}(A_i - Y_i)^2\right)^{1/2}$	$\frac{1}{N}\sum_{i}^{N}\lvert A_i - Y_i\rvert$	$\frac{1}{N}\sum_{i}^{N}\left(\frac{\lvert A_i - Y_i\rvert}{A_i}\right) \times 100$	$\lvert A_i - Y_i\rvert$
Downward axial Force model	45.53	21.61	0.64%	169
Failure Force model	80.96	61.85	1.73%	211

(A_i, Y_i, N, and i are experimental value, predicted value, total number of experimental data, and trial number, respectively. ɪe formulas are from Ref.[76])

Table 3: The amount of correlation coefficient (R2) for the training data, confirmation data, and overall data for the developed ANN models.

$R^2 = 1 - \left(\frac{\sum_i^N (A_i - Y_i)^2}{\sum_i^N (Y_i)^2}\right)$	**R2 (For training data)**	**R2 (For confirmation tests)**	**R2 (For all data)**
Downward axial Force model	0.999981092	0.998622385	0.999846253
Failure Force model	0.999722461	0.998734858	0.999634318

(A_i, Y_i, N, and i are experimental value, predicted value, total number of experimental data, and trial number, respectively. The formula is from Ref.[76])

Table 4: The amount of experimental data, predicted data, and its error for the confirmation experiments.

Sample No.	**Measured downward axial Force (N)**	**Predicted downward axial Force (N)**	**Error of model for downward axial Force (N)**	**Measured failure force (N)**	**Predicted failure force (N)**	**Error of model for failure force (N)**
15	3310	3479	5%	2709	2810	3.7%
20	4720	4553	3.5%	4746	4876	2.7%
27	2930	3050	4.1%	2123	2025	4.6%
34	3320	3373	1.6%	3982	4156	4.4%

(Error of model = $\frac{\lvert A_i - Y_i\rvert}{A_i} \times 100$)