Improving Stability of Welding Model with ME-ELM

Jianxiong Ye, Han Ye, Zhigang Li, Xingling Peng, Jinlan Zhou and Bo Guo

Abstract Welding shape is important in evaluating welding quality, but accurate predictive model is hard to achieve, because welding is a complex nonlinear process, and the sampled data are inevitably contaminated. Extreme learning machine (ELM) is used to construct a single-hidden layer feedforward network (SLFN) in our study, for improving stability of welding model, M-estimation is combined with ELM and a new algorithm named ME-ELM is developed; researches indicate that it works more effective than BP and other variants of ELM in reducing influence, furthermore, it can improve the model's anti-disturbance and robustness performance even if the data are seriously contaminated. Real TIG welding models are constructed with ME-ELM, by comparing with BP, multiple nonlinear regression (MNR), and linear regression (LR), conclusions can be gotten that ME-ELM can resist the interference effectively and has the highest accuracy in predicting the welding shape.

Keywords Welding shape · Welding model · ME-ELM algorithm Stability

1 Introduction

There are many kinds of welding methods, such as resistance welding, braze welding, gas metal arc welding (GMAW), tungsten inert gas (TIG) welding, flux cored arc welding (FCAW), submerged arc welding (SAW), etc., as a kind of

J. Ye (✉) · X. Peng · J. Zhou · B. Guo
Jiangxi Province Key Laboratory of Precision Drive & Control, Nanchang Institute of Technology, Nanchang, China
e-mail: jxlpyjx@163.com

H. Ye
Jiangxi Aeronautical Institute, Nanchang, China

Z. Li
School of Mechanics Engineering, East China Jiao Tong University, Nanchang, China

S. Chen et al. (eds.), *Transactions on Intelligent Welding Manufacturing*, Transactions on Intelligent Welding Manufacturing,
https://doi.org/10.1007/978-981-10-7043-3_4

hot-working process, welding is widely used in building of marine engineering, railways, and electrical power plants etc., it is reported that more than 2/3 of the steel need welding before utilization [1].

The mechanical property after welding is not only determined by composition of metal, but also by the shape of welding joint [2, 3], the desired welding shape relies on many factors, such as welding speed, wire feeding speed, welding current, welding gas flow rate, so it is difficult to construct the model between the shape and so many welding parameters. Till now, there are at least four kinds of modeling methods: multi-nonlinear regression (MNR), response surface methodology (RSM), Taguchi method, and ANN nonlinear mapping [4–7]. Shi et al. [8] used MNR to predict the bead geometry in wet FCAW, and sensitive analysis is performed later, this method is also used on SAW to predict the pips bead shape and realize online control [3]; Palani and Morgan [9] used RSM to develop a model predicting welding joint shape in FCAW; Taguchi is popularly used and has various forms, Tarng et al. [10] and Biswas et al. [11] applied grey-Taguchi and PCA-Taguchi to predict the bead shape in SAW. However, because welding shape relates to many factors, all methods above can not work efficiently and effectively. ANN and other similar intelligent calculation methods are widely used now [12–15], but from the view of mathematics, ANN and its variants still have to face several issues like time-consuming, over-fitting, or local minima, it is meaningful to find out new measures to build welding model.

Based on ELM, some hybrid methods are supposed in our research. SA, GS, and GA are combined with normal ELM to find out better network structures, ME-ELM is suggested for reducing training data noise to enhancing the model stability and accuracy.

The rest of this paper is arranged as follows. Section 2 introduces the principle of basic ELM, points out relevant problems relating accuracy and stability; Sect. 3 focus on enhancing model stability, ME-ELM is introduced in detail, tests on specific complicated functions indicate that this algorithm has the ability to refrain the adverse effect of noise; Sect. 4 provides models on real TIG welding, besides the method proposed in this paper, BP and MNR are also used to create welding models, the residual errors are compared as well as that of LR(linear regression) which has been published in references [27] already; in the end, conclusions are presented in Sect. 5.

2 Introduction of Basic ELM

2.1 Principle of ELM

ELM is a kind of feedforward neural network, it has two types of structure, named as multilayer structure and single-layer structure. Single-layer structure is an SLFN, it has an input layer, an output layer, and only one hidden layer, each hidden neuron has an activation function, the functions may be same or different, just as shown in

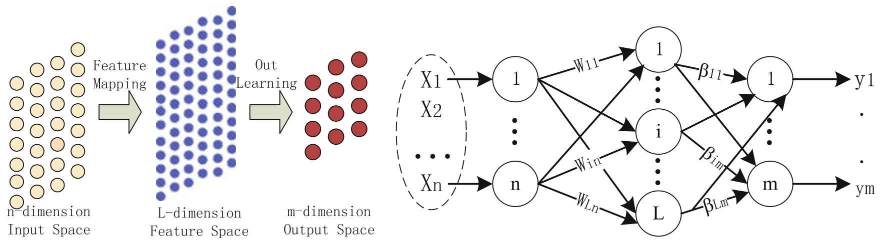

Fig. 1 Network structure of ELM

Fig. 1, $[x_1, x_2, \ldots, x_N]$ means input data, $[y_1, y_2, \ldots, y_N]$ means output data, $W = [W_1, W_2, \ldots, W_N]$ and $\beta = [\beta_1, \beta_2, \ldots, \beta_N]$ are input matrix and output matrix. Hidden layer maps data from input space to feature space with input matrix W, and then convert them into result space with output matrix β, it is clear that they play an important role in model performance.

The working process is as follows: Given training samples (x_i, t_i), input vector and output vector are $x_i = (x_{i1}, x_{i2}, \ldots, x_{in}) \in R^n$ and $y_i = (y_{i1}, y_{i2}, \ldots, y_{im}) \in R^m$, if the number of hidden neurons is L and activation function is $g(x_i)$, we may have:

$$y_j = \sum_{i=1}^{L} \beta_i g(w_i x_j + b_i),\ j = 1, \ldots, N \tag{1}$$

where $w_i = (w_{i1}, w_{i2}, \ldots, w_{in})$ indicates weights from input row to ith hidden neuron, $\beta_i = [\beta_{i1}, \beta_{i2}, \ldots, \beta_{im}]$ represents rights from ith hidden layer to output layer.

Equation (1) can be summarized as:

$$Y = H\beta \tag{2}$$

β means output matrix, H is input matrix and it can be expressed:

$$\begin{aligned} & H(w_1, \ldots, w_L, b_1, \ldots, b_L, x_1, \ldots, x_N) \\ & = \begin{bmatrix} g(w_1 \cdot x_1 + b_1) & \ldots & g(w_L \cdot x_1 + b_L) \\ \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots \\ g(w_1 \cdot x_N + b_1) & & g(w_L \cdot x_N + b_L) \end{bmatrix}_{N \times L} \end{aligned} \tag{3}$$

There is a conclusion in [16] that for a stand SLFNs which has n input neurons, m output neurons, and L hidden layer neurons, given N distinct observations $\{x_i, y_i\}$, if the activation function $g: R \to R$ is infinitely differentiable in any interval, then we can randomly get the w_i and b_i according to any continuous probability distribution and have the result that the input matrix H is definitely invertible and β can be analytically calculated out based on least square solution. So we have Eq. (4).

$$\hat{\beta} = H^{\dagger}T = \min_{\beta}((H\beta - Y) = 0) \tag{4}$$

Judgments can be obtained from Eq. (4) that different $H^{\dagger}$ leads to different network output, and more, different number of hidden neurons needs different methods of calculating $H^{\dagger}$ [17], such as singular value decomposition, orthogonal projection, and iterative methods [18, 19], so how to solve $H^{\dagger}$ is important, a popular and efficient closed-form solution is:

$$\hat{\beta} = \begin{cases} H^T(I/C + HH^T)^{-1}T, & if \quad N \leq L \\ (I/C + H^TH)^{-1}H^TT, & if \quad N > L \end{cases} \tag{5}$$

where C is a parameter used for controlling the trade-off between the training error and norm of output weights [20, 21], it can improve network accuracy significantly.

2.2 Main Problems in Modeling of ELM

Equation (4) discloses the reason why ELM has fast training speed, but reveals two problems with accuracy and stability. The first one is about the number L and the rights of input matrix W, if we can choose proper value of them, the model accuracy will be improved; the second problem is about matrix β, as the discussion above, β is essentially decided by H, different H results in different β, but random W and traditional solving of $H^{\dagger}$ can not guarantee a better β, especially when the training data is contaminated, so we need a new way for better β.

3 Methods for Improving Model Stability

3.1 Design of ME-ELM

Equations (4–5) illustrate that the simulation results are greatly affected by training data, noisy data will inevitably lead to poor performance. M-estimator is a kind of robust estimator which is good at drawing out a reliable conclusion from bad data, especially outliers. So, M-estimator is tried to combine with ELM to decrease the noise influence, by adopting estimation function and least square criterion, output matrix parameters are adjusted during iterations [22], this way is embedded into training algorithm of ELM and is called ME-ELM.

Considering the training target $(H\beta - Y) = 0$, formula (4) can be expressed as follows:

$$\hat{\beta} = (H^T PH)^{-1} H^T PT \tag{6}$$

where P is an adjusting matrix, it can reduce outliers influence by changing its values adaptively. If samples have no bias, all related coefficients in P will be 1, which implies ME-ELM works in the same way as ELM; if part of samples are moderately polluted, the corresponding coefficients in matrix P will be less than 1, to weaken the noise influence; for samples with gross errors, the relative coefficients tend to be zeros, therefore the bad impacts can be decreased greatly. So P plays an important role here, its values are regulated by estimation function ψ, for clear description, statistic function (x) is introduced as first.

$$Q(\beta) = \sum_{i=1}^{N} \rho(e_i) = \sum_{i=1}^{N} \rho(T_i - H_i \beta_i) \tag{7}$$

where $Q(x)$ is optimization objective function, solutions to Eq. (7) are called M-estimators:

$$\hat{\beta} = \arg\min_{\beta} (\sum_{i=1}^{N} \rho(T_i - H_i \beta_i)) \tag{8}$$

Define estimation function $\psi(x) = \frac{\partial \rho(x)}{\partial \beta}$, so the minimum β is:

$$\frac{\partial Q(\beta)}{\partial \beta} = 0 \Rightarrow \sum_{i=1}^{N} \psi(T_i - H_i \beta) H_i = 0 \tag{9}$$

There are several popular estimation functions ψ, similar results can be gotten in terms of efficiency and deviation with one of them [23], Hurb function is expressed as Eq. (10) [24].

$$\psi(x) = \begin{cases} x & |x| \le k \\ k & |x| > k \end{cases}, \quad \rho(x) = \begin{cases} x^2/2 & |x| \le k \\ k|x| - k^2/2 & |x| > k \end{cases}, \quad k = 1.345 \tag{10}$$

Algorithm of ME-ELM can be designed as follows:

Step 1 Determine the network structure, acquire original value with normal ELM: $\beta_0 = H^{\dagger} T$ and $e_0 = T - H^{\dagger} \beta_0$

Step 2 Setting initial parameters, such as adjusting factor $k = 1.345$ and error variable $\varepsilon = 1$

Step 3 Iteration process:

while ((ε <1e04) or (N<100))

(a) Standardizing e_i as $e_i = e_i / s = 0.6745 e_i / \mathrm{med}(|e_i|)$, $\mathrm{med}(|e_i|)$ is the middle of $|e_i|$.

(b) Adjusting $W_i = \frac{\psi(u_i)}{u_i}$ and calculate $\hat{\beta}^{(i)} = \left(H^T W_i H\right)^{-1} H^T W_i T$ with Huber function.

(c) Renewing variables: $e_i = T - H\hat{\beta}^{(i)}$, $N = N+1$, $\varepsilon = \left|\hat{\beta}^{(i)} - \hat{\beta}^{(i-1)}\right|$

$$e_i = T - H\hat{\beta}^{(i)}, \quad N = N+1, \quad \varepsilon = \left|\hat{\beta}^{(i)} - \hat{\beta}^{(i-1)}\right|$$

End while

3.2 Experiments for Stability

To verify the capability of ME-ELM in enhancing stability, some popular algorithms are used for comparison, including ELM, ELM-C, B_ELM, and BP, the networks are constructed based on noisy samples and tested with non-noise data, then stability performance can be distinguished by RMSE and DEV. SinC function is defined as follow:

$$y(x) = \begin{cases} \sin x/x & x \neq 0 \\ 1 & x = 0 \end{cases} \tag{11}$$

First, 5000 groups data are generated randomly, then white Gaussian noise is added to independent variables, by selecting different noise distribution range of [−0.2 0.2], [0 2], and [−2 2], three batches of 5000 training data are prepared, after network models are constructed by various algorithms, stability performance can be checked out with noise-free data. Comparisons are carried out between ME-ELM and some other algorithms, such as ELM-C, BP, the number of hidden neurons is fixed as 20, to examine the universal property of ME-ELM, several linear and nonlinear multivariable function are tested, results of $y = \exp(x_1/2)\, x_1 + \sin(x_2)$ are also listed in Table 1.

In addition, to avoid rank deficient problem in ME-ELM, random minor value can be added to the adjusted value on the base of Eq. (10), which is expressed as follows.

$$\psi(x) = \begin{cases} x & |x| < k(k > 0) \\ k + 0.1 * rank(0) & |x| > k(k > 0) \end{cases} \tag{12}$$

Obvious difference in stability performance can be seen from Fig. 2, where both ELM [16] and ME-ELM are used to model SinC function when the training data have been added Gaussian noise, and the noise distribution interval is [0 2].

It is clear from the comparison above that ME-ELM has good ability of noise reduction, it can produce a better model which has much better performance than ELM, ELM-C, and B-ELM.

Table 1 Testing error comparison between different models

Function	Testing error	ELM	ELM−C	B-ELM	BP	ME-ELM	Noise distribution
$y = \sin c(x)$	RMSE	0.002498	0.075118	0.365096	0.03426	**0.000121**	[−0.2 0.2]
	Dev	0.000044	−0.002815	0.105855	−0.000025	**−0.000001**	
	RMSE	0.102556	0.123228	0.329231	0.109186	**0.002475**	[0 2]
	Dev	−0.100852	−0.098573	0.005053	−0.102048	**0.000630**	
	RMSE	0.020395	0.074426	0.3352543	0.037857	**0.001051**	[−2 2]
	Dev	0.002487	0.004764	0.072820	0.002093	**−0.000035**	
$y = \exp(x_1/2)*$ $x_1 + \sin(x_2)$	RMSE	0.002307	0.056295	0.073086	0.006358	**0.000030**	[−0.2 0.2]
	Dev	0.0000078	0.000599	−0.033326	0.000113	**0**	
	RMSE	0.101063	0.115764	0.486288	0.111284	**0.000576**	[0 2]
	Dev	0.099678	−0.100685	−0.415225	−0.102691	**−0.000111**	
	RMSE	0.016717	0.057668	0.093592	0.045291	**0.000105**	[−2 2]
	Dev	−0.000287	−0.001082	0.074013	−0.000445	**0.000013**	

Bold numbers are the best ones with least errors in different trials

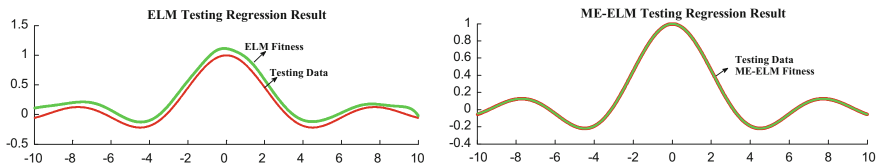

Fig. 2 Testing results of ELM and ME-ELM

4 Study of TIG Welding Modeling and Welding Shape Prediction

Welding is a complex time-variant process and welding joint shape relates to many variables [25, 26], furthermore, joint shape is susceptible to interference, even the fluctuation of electric grid may result in the changing of geometry, and all the training data will inevitably be contaminated, so how to guarantee the model performance is very important.

4.1 Experimental Design and Data Acquisition

To design experiments, the contributing factors and relative levels should be determined at first, and then experimental design matrix will be arranged in orthogonal method, the work is often planned as follows:

1. Identification of important process parameters.
2. Finding the upper and lower limits with different levels of the parameters.
3. Confirm design matrix according to orthogonal table
4. Conducting the experiments as per the design matrix, if needed, repeating the specific experiment.
5. Specimen preparation, if necessary, measuring the bead shape on different samples.
6. Treating data with filtering and recording these responses.

For convenient comparison, TIG welding data published in [27] are used, it is also used in [28] where the TIG welding variables include welding speed (S), wire speed (WS), cleaning percentage (CP), welding current (C) and arc gap (G), and weld bead shape parameters comprise front height (FH), front width (FW), back height (BH), and back width (BW), shown in Table 2.

Table 2 Experimental data [27]

Number	Input					Output			
	Speed/ (cm.min−1)	Wire speed/ (cm.min−1)	Cleaning/ (%)	Gap/ (mm)	Current/ (V)	Front height/ (mm)	Front width/ (mm)	Back height/ (mm)	Back width/ (mm)
1	24	1.5	30	2.4	80	−0.149	6.09	0.672	5.664
2	24	1.5	30	3.2	80	0.027	6.411	0.412	5.197
3	24	1.5	70	2.4	80	−0.179	7.432	0.593	7.058
4	24	1.5	70	3.2	80	−0.306	7.287	0.63	6.895
5	24	2.5	30	2.4	80	0.155	6.676	0.743	5.96
6	24	2.5	30	3.2	80	0.099	6.824	0.803	5.732
7	24	2.5	70	2.4	80	−0.129	7.009	0.878	6.989
8	24	2.5	70	3.2	80	−0.077	7.46	0.82	7.809
9	24	1.5	30	2.4	95	−0.017	8.664	0.437	8.75
10	24	1.5	30	3.2	95	−0.25	8.782	0.593	9.993
11	24	1.5	70	2.4	95	−0.553	9.757	0.852	9.993
12	24	1.5	70	3.2	95	−0.42	10.374	0.736	10.687
13	24	2.5	30	2.4	95	−0.345	9.783	0.965	10.237
14	24	2.5	30	3.2	95	−0.043	8.803	0.654	9.076
15	24	2.5	70	2.4	95	−0.134	9.75	0.798	9.465
16	24	2.5	70	3.2	95	−0.168	10.348	0.708	10.193
17	24	1.5	30	2.4	110	−0.599	11.348	0.805	11.679
18	24	1.5	30	3.2	110	−0.745	11.491	1.1	11.848
19	24	1.5	70	2.4	110	−0.254	11.237	0.47	12
20	24	1.5	70	3.2	110	−0.683	12.946	0.945	13.921
21	24	2.5	30	2.4	110	−0.232	9.338	0.866	10.611
22	24	2.5	30	3.2	110	−0.557	12.348	1.139	12.403
23	24	2.5	70	2.4	110	−0.623	11.767	1.128	12.86
24	24	2.5	70	3.2	110	−0.617	12.533	1.084	13.346
25	35	1.5	30	2.4	80	0.123	5.355	0.245	4.104

(continued)

Table 2 (continued)

Number	Input					Output			
	Speed/ (cm.min−1)	Wire speed/ (cm.min−1)	Cleaning/ (%)	Gap/ (mm)	Current/ (V)	Front height/ (mm)	Front width/ (mm)	Back height/ (mm)	Back width/ (mm)
26	35	1.5	30	3.2	80	0.108	5.173	0.34	3.418
27	35	1.5	70	2.4	80	−0.044	5.833	0.51	4.875
28	35	1.5	70	3.2	80	−0.09	5.831	0.502	5.082
29	35	2.5	30	2.4	80	0.251	5.656	0.557	4.37
30	35	2.5	30	3.2	80	0.23	5.562	0.593	3.948
31	35	2.5	70	2.4	80	0.18	5.711	0.45	5.085
32	35	2.5	70	3.2	80	0.12	5.85	0.626	4.989
33	35	1.5	30	2.4	95	−0.213	6.348	0.458	5.874
34	35	1.5	30	3.2	95	−0.19	6.992	0.447	6.74
35	35	1.5	70	2.4	95	−0.152	7.163	0.464	6.994
36	35	1.5	70	3.2	95	−0.213	7.25	0.504	7.019
37	35	2.5	30	2.4	95	−0.164	7.288	0.715	6.724
38	35	2.5	30	3.2	95	−0.113	6.966	0.746	6.433
39	35	2.5	70	2.4	95	−0.107	7.055	0.696	7.24
40	35	2.5	70	3.2	95	−0.018	7.549	0.591	7.166
41	35	1.5	30	2.4	110	−0.575	8.337	0.766	8.763
42	35	1.5	30	3.2	110	−0.267	8.605	0.506	8.58
43	35	1.5	70	2.4	110	−0.385	9.109	0.672	9.652
44	35	1.5	70	3.2	110	−0.564	9.67	0.743	9.952
45	35	2.5	30	2.4	110	−0.556	8.756	1.011	8.853
46	35	2.5	30	3.2	110	−0.188	9.442	0.666	9.614
47	35	2.5	70	2.4	110	−0.309	9.015	0.784	9.041
48	35	2.5	70	3.2	110	−0.318	9.297	0.785	9.47
49	46	1.5	30	2.4	80	0.357	4.982	0.001	2.255
50	46	1.5	30	3.2	80	0.168	4.898	0.277	2.998

(continued)

Table 2 (continued)

Number	Input					Output			
	Speed/ (cm.min−1)	Wire speed/ (cm.min−1)	Cleaning/ (%)	Gap/ (mm)	Current/ (V)	Front height/ (mm)	Front width/ (mm)	Back height/ (mm)	Back width/ (mm)
51	46	1.5	70	2.4	80	0.088	5.02	0.281	3.302
52	46	1.5	70	3.2	80	0.09	4.423	0.42	3.172
53	46	2.5	30	2.4	80	0.39	4.78	0.062	1.33
54	46	2.5	30	3.2	80	0.487	4.992	0.139	1.6
55	46	2.5	70	2.4	80	0.38	5.231	0.397	2.817
56	46	2.5	70	3.2	80	0.394	5.337	0.378	3.041
57	46	1.5	30	2.4	95	−0.321	5.847	0.44	5.332
58	46	1.5	30	3.2	95	−0.152	5.704	0.386	5.35
59	46	1.5	70	2.4	95	−0.155	5.967	0.445	5.415
60	46	1.5	70	3.2	95	−0.09	5.892	0.399	5.319
61	46	2.5	30	2.4	95	−0.236	5.984	0.696	5.531
62	46	2.5	30	3.2	95	0.067	6.03	0.575	5.636
63	46	2.5	70	2.4	95	−0.075	5.562	0.816	4.835
64	46	2.5	70	3.2	95	0.138	6.546	0.575	6.285
65	46	1.5	30	2.4	110	−0.217	6.092	0.359	6.419
66	46	1.5	30	3.2	110	−0.339	7.335	0.619	7.52
67	46	1.5	70	2.4	110	−0.249	7.719	0.492	7.706
68	46	1.5	70	3.2	110	−0.396	7.633	0.458	7.601
69	46	2.5	30	2.4	110	−0.01	6.396	0.536	6.197
70	46	2.5	30	3.2	110	0.074	6.863	0.484	6.072
71	46	2.5	70	2.4	110	−0.201	7.052	0.658	7.48
72	46	2.5	70	3.2	110	−0.358	7.759	0.798	7.917

4.2 Models Performance Comparison for TIG

For simplicity, only the comparison between ME-ELM and BP, MNR, and LR are given out. Welding data are normalized at first and then separated into two groups: 56 records are used for training and the rest 16 records for testing. To begin with ME-ELM, the number of middle neurons is set as 40 by SA at first, and then, GA is used to get proper input matrix, M-estimation is used at last to calculate the output matrix. BP network is created with 40 middle layer numbers and four output numbers, active function and output function are "tansig" and "purelin", the comparison is shown in Fig. 3.

The MNR model is always better than linear regression [29], suppose the input variables are $X1$, $X2$, $X3$, $X4$, $X5$, the output variables indicating the bead shape are denoted by FH, FW, BH, and BW, the nonlinear regression forms can be given out as:

$$\left.\begin{aligned} FH &= g_1 x_1^{a_1} x_2^{a_2} x_3^{a_3} x_4^{a_4} x_5^{a_5} \\ FW &= g_2 x_1^{b_1} x_2^{b_2} x_3^{b_3} x_4^{b_4} x_5^{b_5} \\ BH &= g_3 x_1^{c_1} x_2^{c_2} x_3^{c_3} x_4^{c_4} x_5^{c_5} \\ BW &= g_1 x_1^{d_1} x_2^{d_2} x_3^{d_3} x_4^{d_4} x_5^{d_5} \end{aligned}\right\} \tag{13}$$

By proper treating of original data, it can be converted to:

$$\left.\begin{aligned} \lg(FH) &= G_1 + a_1 \lg x_1 + a_2 \lg x_2 + a_3 \lg x_3 + a_4 \lg x_4 + a_5 \lg x_5 \\ \lg(FW) &= G_2 + b_1 \lg x_1 + b_2 \lg x_2 + b_3 \lg x_3 + b_4 \lg x_4 + b_5 \lg x_5 \\ \lg(BH) &= G_3 + c_1 \lg x_1 + c_2 \lg x_2 + c_3 \lg x_3 + c_4 \lg x_4 + c_5 \lg x_5 \\ \lg(BW) &= G_4 + d_1 \lg x_1 + d_2 \lg x_2 + d_3 \lg x_3 + d_4 \lg x_4 + d_5 \lg x_5 \end{aligned}\right\}, G_i = \lg(g_i), \quad i = 1, 2, 3, 4 \tag{14}$$

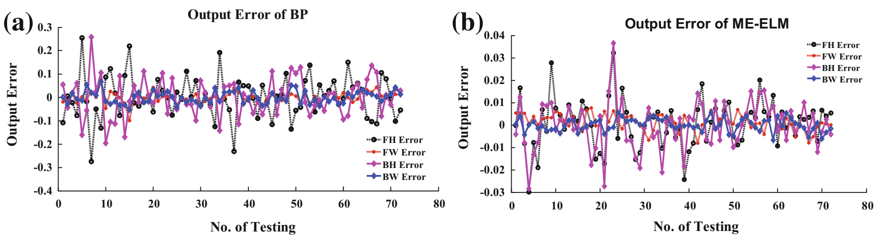

Fig. 3 Training errors of BP and ME-ELM

So the regression model can be achieved according to Least Square principle:

$$\left.\begin{array}{l} FH = 0.7341\left(\frac{x_1}{46}\right)^{0.5667}\left(\frac{x_2}{2.5}\right)^{0.3428}\left(\frac{x_3}{70}\right)^{-0.0874}\left(\frac{x_4}{3.2}\right)^{-0.0459}\left(\frac{x_5}{110}\right)^{-1.9754}-1 \\ FW = 7.8699\left(\frac{x_1}{46}\right)^{-0.6622}\left(\frac{x_2}{2.5}\right)^{0.0392}\left(\frac{x_3}{70}\right)^{0.0844}\left(\frac{x_4}{3.2}\right)^{0.1290}\left(\frac{x_5}{110}\right)^{1.4149} \\ BH = 0.7949\left(\frac{x_1}{46}\right)^{-1.2730}\left(\frac{x_2}{2.5}\right)^{0.7392}\left(\frac{x_3}{70}\right)^{0.3809}\left(\frac{x_4}{3.2}\right)^{0.6715}\left(\frac{x_5}{110}\right)^{2.3903} \\ BW = 7.6050\left(\frac{x_1}{46}\right)^{-1.0607}\left(\frac{x_2}{2.5}\right)^{-0.0473}\left(\frac{x_3}{70}\right)^{0.01877}\left(\frac{x_4}{3.2}\right)^{0.1341}\left(\frac{x_5}{110}\right)^{2.5845} \end{array}\right\} \quad (15)$$

Comparison results between MNR, LR, and ME-ELM are shown in Table 3, The better values are shown in bold font, it is clear that ME-ELM algorithm has much more better performance.

5 Discussion of ME-ELM in Underwater Welding

Studies of welding parameters optimization methods have been carried out for a long time, in which the welding model is very important, especially in underwater welding [30]. There are three kinds of welding methods for underwater welding, named as wet welding, dry welding, and semidry welding. Compared with other two methods, there are many bubbles and turbulent fluid accompany with the wet welding process, in the meantime, evaporation cooling has a great effect on the melt zone so as to lead to a bad welding performance, the horizontal resurfacing welding is shown in Fig. 4. So it is critical to refrain the noises for building the wet welding model, preliminary study suggested that ME-ELM can work its way effectively and detail results will be given out in further paper.

6 Conclusion

Methods for improving model accuracy and stability are studied in this paper, comparison results on benchmark problems, artificial functions, and real TIG welding process indicate that ME-ELM can work effectively. Further conclusions can be drawn as follows:

Table 3 Testing errors of MNR, LR [27] and ME-ELM for trials 1–16

Trial No.	Front height, FH /(mm)			Front width, FW /(mm)			Back height, BH/(mm)			Back width, BW/(mm)		
	MNR	LR [27]	ME-ELM	MNR	LR [27]	ME-ELM	MNR	LR [27]	ME-ELM	MNR	LR [27]	ME-ELM
1	−0.1624	**−0.0006**	0.0020	0.6293	−0.0415	**0.0019**	0.0099	**−0.0034**	0.00856	0.6206	0.1333	**0.0046**
2	−0.0903	0.0164	−0.0065	−0.5725	1.2334	**0.0003**	0.1509	**0**	−0.0027	0.5522	1.1617	**−0.0007**
3	−0.1112	0.0063	**0.0029**	0.2449	−0.1281	**0.0038**	−0.2356	**−0.0028**	0.0153	−0.496	0.2787	**0.0013**
4	**0.0023**	−0.0088	0.0158	−0.323	0.1066	**0.00130**	0.12986	**0.0056**	0.0130	1.2480	−0.2834	**0.0012**
5	−0.0074	0.0034	**0.0010**	−0.1452	−0.0647	**0.0003**	−0.1127	**−0.0033**	0.0070	−0.495	0.18	**−0.0010**
6	0.0708	−0.0121	**−0.0113**	0.4923	0.1534	**−0.0069**	−0.1654	0.0052	**−0.0030**	0.2487	−0.3809	**−0.0029**
7	0.0335	0.1452	**−0.0050**	0.2550	0.4801	**0.00163**	0.0299	**−0.1216**	−0.0024	−1.1512	0.5552	**−0.0040**
8	0.0214	**−0.0054**	−0.0299	−0.3990	0.0601	**−0.0055**	−0.0909	**0.0024**	−0.0248	0.3957	−0.1491	**−0.0063**
9	0.0660	**0.0015**	−0.0069	−0.7024	**0.0001**	−0.0066	−0.0604	**0.0044**	−0.0047	1.0689	−0.0704	**0.0027**
10	0.0053	**0.0041**	−0.0072	−0.1539	−0.0572	**−0.0007**	−0.2244	−0.0007	**−0.0002**	−0.8752	0.1368	**0.0023**
11	0.1061	−0.0118	**0.0075**	−1.0767	0.2149	**0.0006**	0.17868	0.0088	**−0.0015**	0.52997	−0.5185	**−0.0058**
12	−0.185	0.0077	**−0.0111**	0.6632	−0.2057	**−0.0047**	−0.06152	−0.013	**−0.0040**	0.04378	0.5619	**−0.0028**
13	0.1110	**0.0028**	−0.0103	−1.0788	−0.065	**0**	0.30235	−0.0058	**−0.0034**	1.61262	0.1894	**−0.0027**
14	−0.123	0.6169	**−0.0069**	−0.2866	−0.371	**−0.0030**	−0.0428	−0.0665	**0.0024**	−0.2102	−0.2039	**−0.0081**
15	0.1397	−0.0012	0.0207	−0.1012	−0.0114	**−0.0026**	0.37200	−0.002	**0.0086**	1.7290	0.0448	**0.0011**
16	0.0952	**−0.0014**	0.0023	0.1103	0.0397	**−0.0043**	−0.0031	0.0045	**0.0003**	−0.3124	−0.1223	**0.0006**

Bold numbers are the best ones with least errors in different trials

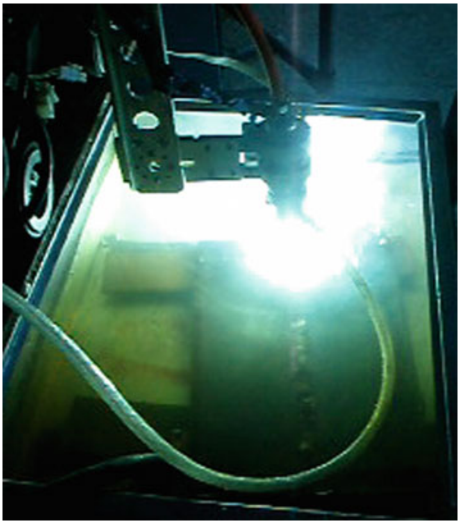

Fig. 4 Resurfacing of wet welding

- Regarding the ability of reducing data noise and improving the simulation accuracy, ME-ELM is better than BP, normal ELM and its corrective methods, such as ELM-C, B-ELM, it is suitable for constructing welding model.
- Parameter k of estimation functions in ME-ELM is very important. Small k lead to small rights for outliers, this will result in a strong suppression on influence. On the contrary, if k is set to be a big positive number, ME-ELM tends to be the normal ELM.
- MNR and LR are all prototype-based, their performance relies mainly on type assumption, generally speaking, and their simulation accuracy is inferior to that of ME-ELM.

Acknowledgements The work is partially founded by the Open Project Program of Jiangxi Province Key Laboratory of Precision Drive & Control (PLPDC-KFKT-201625), the National Natural Science Foundation of China (51665016), the JiangXi Province Science Foundation (20151BAB207047).

References

1. Du B, Cheng Y, Li CF (2012) Analysis of global welding market. Chin Mech Manuf Digest 2:1–7
2. American Welding Society (2016) Welding handbook. Welding Handbook committee, Miami, Florida
3. Murugana N, Gunaraj V (2015) Prediction and control of weld bead geometry and shape relationships in submerged arc welding of pipes. J Mater Process Technol 168(3):478–487
4. Kim IS, Basu A, Siores E (1996) Mathematical models for control of weld bead penetration in the GMAW process. Int J Adv Manuf Technol 12(6):393–401

5. Rao PS, Gupta OP, Murty SSN et al (2009) Effect of process parameters and mathematical model for the prediction of bead geometry in pulsed GMA welding. Int J Adv Manuf Technol 45(5–6):496–505
6. Pal K, Pla SK (2011) Soft computing methods used for the modeling and optimization of gas metal arc welding. Int J Manuf Res 6(1):15–29
7. Murugan N, Parmar RS (1994) Effects of MIG process parameters on the geometry of the bead in the automatic surfacing of stainless steel. J Mater Process 41(4):381–398
8. Shi YH, Zhang ZP, Huang J (2013) Sensitivity model for prediction of bead geometry in underwater wet flux cored arc welding. Trans Nonferrous Met Soc China 23(7):1977–1984
9. Palani PK, Murugan N (2006) Development of mathematical models for prediction of weld bead geometry in cladding by flux cored arc welding. Int J Adv Manuf Technol 30(7–8):669–676
10. Tarng YS, Juang SC, Chang CH (2002) The use of grey based Taguchi methods to determine submerged arc welding process parameters in hard facing. J Mater Process Technol 128(1):1–6
11. Biswas A, Bhaumik S, Majumdar G et al (2011) Bead geometry optimization of submerged arc weld: exploration of weighted principal component analysis (WPCA). Appl Mech Mater 110:790–798
12. Satheesh M, Dhas JER (2014) Multi objective optimization of weld parameters of boiler steel using fuzzy based desirability function. J Eng Sci Technol Rev 7(1):29–36
13. Mondal P, Bose D (2015) Optimization of the process parameters for MIG welding of AISI 304 and IS 1079 using fuzzy logic method. Int Res J Eng Technol 2(8):483–488
14. Katherasan D, Elias JV, Sathiya P et al (2014) Simulation and parameter optimization of flux cored arc welding using artificial neural network and particle swarm optimization algorithm. J Intell Manuf 25:67–76
15. Subashini L, Vasudevan M (2012) Adaptive neuro-fuzzy inference system (ANFIS)-based models for predicting the weld bead width and depth of penetration from the infrared thermal image of the weld pool. Metall Mater Trans 43(2):145–154
16. Huang GB, Zhu QY, Siew CK (2006) Extreme learning machine: theory and applications. Neuro Comput 70:489–501
17. Frénay B et al (2010) Using SVMs with randomized feature spaces: an extreme learning approach. In: 18th European symposium on artificial neural networks, vol 2010, D-side publication, Bruges, pp 315–320
18. Huang GB, Ding X, Zhou H (2010) Optimization method based extreme learning machine for classification. Neuro Comput 74(12):155–163
19. Widrow B, Greenblatt A, Kim Y et al (2013) The no-prop algorithm: a new learning algorithm for multilayer neural networks. Neural Networks 37(1):182–188
20. Huang GB, Bai Z, Kasun LLC (2015) Local receptive fields based extreme learning machine. IEEE Comput Intell Mag 10(2):18–29
21. Huang GB (2014) An insight into extreme learning machines: random neurons, random features and kernels. Cogn Comput 6(3):376–390
22. Hampel FR, Ronchetti EM, Rousseeuw PJ et al (2011) Robust statistics: the approach based on influence functions. Wiley, Hoboken, New Jersey
23. Stigler SM (2010) The changing history of robustness. Am Stat 64(4):277–281
24. Street JO, Carroll RJ, Ruppert D (1988) A note on computing robust regression estimates via iteratively reweighted least squares. Am Stat 42(2):152–154
25. Ye JX (2013) Wet welding platform design based on ultrasonic sensor. Int J Appl Crypt 5 (2):707–714
26. Ye JX, Zhang H (2008) Automatic underwater arc welding seam-tracking system based on flux-core wire and rotating arc sensor. Trans China Weld 29(3):141–144
27. Juang SC, Tarng YS, Lii HR (1998) A comparison between the back-propagation and counter-propagation networks in the modeling of the TIG welding process. J Mater Process Technol 75(1–3):54–62

28. Nagesha DS, Datta GL (2010) Genetic algorithm for optimization of welding variables for height to width ratio and application of ANN for prediction of bead geometry for TIG welding process. Appl Soft Comput 10(3):897–907
29. Du JH (2010) A study on underwater welding seam forming and tracking based on rotating arc sensor. Dissertation, South China University of Technology
30. Ye JX, Zhang H (2009) Research of intellectual calculation in welding parameters optimization. Mater Rev 23(12):69–72