Levenberg-Marquardt, Bayesian-regularization, and scaled conjugate gradient algorithms for predicting surface roughness accuracy on side milling AISI 1045

By Amrifan Saladin Mohruni

WORD COUNT	3948	TIME SUBMITTED	28-APR-2023 07:38PM
		PAPER ID	99191812

RESEARCH ARTICLE | APRIL 24 2023

Levenberg-Marquardt, Bayesian-regularization, and scaled conjugate gradient algorithms for predicting surface roughness accuracy on side milling AISI 1045

M. Yanis ✉; A. Y. Budiman; A. S. Mohruni; ... et. al

AIP Conference Proceedings 2544, 020013 (2023)
https://doi.org/10.1063/5.0117323

Levenberg-Marquardt, Bayesian-Regularization, and Scaled Conjugate Gradient Algorithms for Predicting Surface Roughness Accuracy on Side Milling AISI 1045

M Yanis[1,a)], A Y Budiman[1,b)], A S Mohruni[1,c)], S Sharif[2,d)]
M A Suhaimi[2,e)] and H Dwipayana[3,f)]

[1]*Mechanical Engineering Department, Sriwijaya University, Kampus Unsri, Indralaya 30660, Ogan Ilir, South Sumatra, Indonesia.*
[2]*Faculty of Mechanical Engineering, Universiti Teknologi Malaysia, 81310 UTM-Skudai, Johor, Malaysia.*
[3]*Mechanical Engineering Department, Taman Siswa University, Palembang, South Sumatra, Indonesia.*

[a)]Corresponding author: yanis@unsri.ac.id
[b)]arieyudha2002@gmail.com
[c)]mohrunias@unsri.ac.id
[d)]safian@fkm.utm.my
[e)]azlansuhaimi@utm.my
[f)]hrd_dwipayana@yahoo.co.id

Abstract. Surface roughness quality is an important requirement for functional machine components such as considerations of wear, lubrication, corrosion, surface fatigue and friction. In machining, this is influenced by machining parameters and it is difficult to develop a theoretical model to describe machining efficiently and completely. In this study, prediction using Artificial Neural Network (ANN) was developed. The Levenberg Marquardt (LM), Bayesian Regularization (BR), and Scaled Conjugate Gradient (SCG) algorithms were compared for the AISI 1045 side milling data. Machining parameters consist of cutting speed, feeding rate, radial and axial depth of cut. The network was trained using structures with the number of neurons 1 to 20 in a hidden layer. It is found that the best network structure for the LM and BR algorithms is 4-10-1 and for the SCG algorithm is 4-9-1. The LM, BR, and SCG algorithms are able to produce predictions that are very close to the experimental results. Based on network performance, the algorithms that produce the best mean square error and coefficient of determination are SCG, LM and BR, respectively.

INTRODUCTION

The final goal to be achieved from the machining process is a component that conforms to the dimensions, geometry, and surface integrity specifications. These characteristic specifications are obtained empirically from experience requiring a large number of materials, tools and certain machining conditions. Another method for achieving this is characteristic analysis or optimization of the relationship between input and output parameters. The conventional method commonly used by researchers is statistical analysis based on the mathematical equation of the input-output variable relationship. The second method is based on a non-conventional approach, namely computerized modeling and simulation. This method can reduce experimental work which is costly and time consuming. One of the advanced optimization techniques with good quality is an artificial neural network (ANN). The advantages of this ANN method offer the ability to create models and nonlinear interactions that are more complex than statistical methods [1, 2].

The ANN method has been widely applied in various fields of science. The Feed Forward Back Propagation (BP) network model has been chosen by many researchers for ANN analysis because of the combination of a flexible network structure with many layers. Analysis using pattern recognition-based ANN through the learning process

Proceeding of the 1st International Conference on Manufacturing Engineering Technology (IConMET 2021)
AIP Conf. Proc. 2544, 020013-1–020013-9; https://doi.org/10.1063/5.0117323
Published by AIP Publishing. 978-0-7354-4322-8/$30.00

(training and testing) of input and output data. The learning process in the BP model has many algorithms available. The BP algorithm that many researchers apply to train neural networks is Levenberg Marquardt (LM), Bayesian Regularization (BR), scale conjugation gradient (SCG). This is because the algorithm produces a better level of accuracy and determination coefficient than other algorithms [3, 4, 5]. The regularization method developed in the LM and BR algorithms can overcome the overfitting problem so that the mean square error (MSE) is better than other algorithms [6]. The results of research that have been conducted show that the LM, Br and SCG algorithms have competitive results for network performance [7].

In manufacturing, several studies were analyzed using ANN in predicting and optimizing the results of their investigations. Angelos et al. [8] conducted a comparison of the LM and BR Algorithms with the Radial Basis Function (RBF) network to predict surface hardness. The trained network with RBF network and BR algorithm was more stable than LM algorithm. Mia and Dhar [9] compared three algorithms namely LM, BR, and the SCG algorithm which are used to train the developed model. The best prediction accuracy is obtained based on the BR algorithm. Kilickap et al. [10] uses the LM algorithm and Response Surface Methodology (RSM) to predict cutting forces, tool wear, and surface roughness. The results from these two methods are very similar to experimental data. Tool wear analysis was applied by Thangarasu et al. [11] with the LM algorithm, quasi-Newton BFGS, and Gradient Descent with Momentum and Gradient descent with adaptive learning rate (GDA). It finds that the trainbfg algorithm provides the lowest MSE with minimum computation time and recommends five hidden layers for predicting tool wear.

Surface roughness is an important requirement for the functional behavior of a component and affects the mechanical and chemical properties of machine parts (wear, lubrication, corrosion, fatigue, and surface friction) [12, 13]. In the machining process, the use of cutting fluid is one way to improve surface quality. The application of environmentally friendly machining using the MQL method and vegetable oil as cutting fluid has been widely studied. The goal of environmentally friendly machining is to reduce waste and pollution, health impacts, and biodegradable. Gupta and Laubscer [14] in the review informed grinding of Ti6Al4V with commercial MQL vegetable oil. Elmunafi et al. [15] AISI 420 stainless turning uses MQL castor oil as the cutting fluid. Sultan et al. [16] used MQL sunflower oil containing 20% Tween 85 during AISI 304 drilling material. MQL resulted in lower surface roughness compared to wet and dry cutting.

This study aims to compare the predictive ability of surface roughness values using the LM, BR and SCG algorithms. Surface roughness is the result of side milling of AISI 1045 material using coconut oil as cutting fluid in the MQL method. The input parameters are cutting speed, feeding rate, radial and axial depth of cut. The comparison criterion is the network output with experimental value.

METHODOLOGY

Design of Experiment Setup

The input data used for analysis on ANN are cutting speed (V_c), feeding rate (f_z), radial (a_r) and axial (a_x) depth of cut. The data consisted of 30 training data and 4 test data. The target data (machining performance) is arithmetic surface roughness (R_a). The selected input data ranges are determined by the maximum and minimum levels listed in Table 1. A rotatable CCD type with a rotation radius of 2 was selected as the experimental design to predict the relationship between the data and the output data [17].

TABLE 1. Input data at each level at the main point of the CCD

Input Data		Levels				
		Lowest (-2)	Low (-1)	Center (0)	High (+1)	Highest (+2)
V_c- Cutting speed	(m/min)	8.9	16.3	23.7	31.1	38.5
f_z - Feed rate	(mm/tooth)	0.0365	0.053	0.0695	0.086	0.1025
a_x - Axial DOC	(mm)	4.0	6.0	8.0	10.0	12.0
a_r - Radial DOC	(mm)	0.2	0.2	0.3	0.5	0.6

The machining process is carried out using a conventional vertical milling machine (maximum power of 1.5 kW). The cutting tools used were uncoated carbide (K2 EMC 54100), 10 mm in diameter, 4 flutes, and 60° helical angles. To ensure rigidity when machining is running, the end mill is installed with an effective length of 30 mm on the tool holder. The machining is side milling on AISI 1045 material (dimensions 25x100x200 mm) with cutting conditions using coconut oil as the cutting fluid in the MQL method. It has been considered as one of the environmentally friendly

cutting fluids and machining methods. The coconut oil specification has a flashpoint of 286°C, a density of 925.8 kg/m3 at 15°C, and absolute viscosity of 1.84 cP at 40°C [7]. Specifications of AISI 1045 material and tool and are given in Table 2 and Table 3.

TABLE 2. AISI 1045 material specifications

Properties	Values		Properties	Values			
Mechanical Properties	Tensile strength	: 565 N/mm^2	Chemical Composition (% weight)	C	: 0.42-0.48	P	: ≤ 0.03
	Hardness (Brinell)	: 163		Si	: 0.15-0.35	S	: ≤ 0.035
	Elongation	: 16 %		Mn	: 0.60-0.90	Ni	: ≤ 0.2
	Density	: 7.87 gr/cm^2					

TABLE 3. Uncoated EMC 54100 K2 carbide (chemical composition)

W	Co	Ti
71.31±1.97	8.78±0.51	19.91±1.79

Artificial Neural Networks (ANN)

ANN is a computational method for prediction, optimization, identification, approach, or control inspired by human brain systems. This method is widely applied by researchers, especially for nonlinear systems approaches. ANN structure consists of layers of input, hidden and output where each layer has data information called neurons as given in Fig.1.

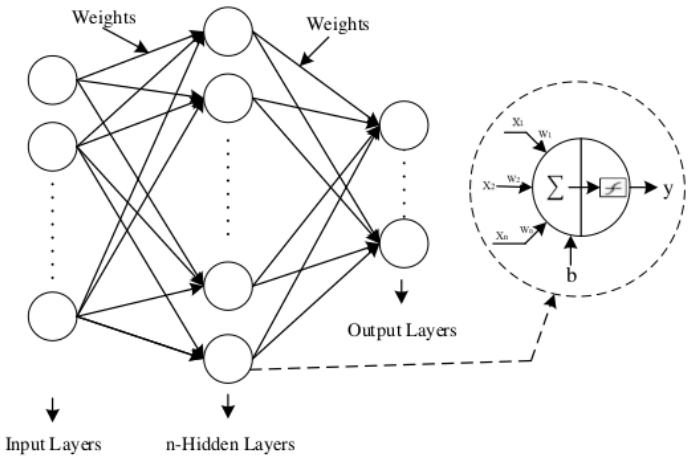

FIGURE 1. A network structure with n-hidden layers

Data processing from input to hidden layer and hidden to output layer is calculated using the weight (w) and bias (b) functions. This step is referred to as the learning process or the training and testing process. The neuron output value is determined by the sum of the weights called the net value (n_v), and is calculated using Equation (1). This net value is then transferred by the activation function, which in this study uses the tansig function (bipolar sigmoid-t_h) as shown in Equation (2). To evaluate network performance using error statistics with Mean Square Error (MSE) which is determined from Equation (3).

$$n_v = f\left(\sum_{i=0}^{n} w_i x_i + b\right) \quad (1)$$

$$t_h = \frac{1 - e^{-x}}{1 + e^{-x}} \tag{2}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(|x_i - y_i|\right)^2 \tag{3}$$

Where, n is the total experimental data, x_i and y_i are input and output data.

Levenberg-Marquardt Algorithm (LM)

Levenberg-Marquardt (LM) algorithm for numerical solutions to problems of minimizing nonlinear functions. It is great for small and medium-sized data. LM has stable convergence which works faster than other algorithms. The process of updating the weights and biases of the LM algorithm uses the Hesian (H) matrix approach with a gradient (g) which can be calculated by the following Equation (4) [18].

$$\begin{aligned} H &= J^T J \\ g &= J^T e \end{aligned} \tag{4}$$

From Equation (4), J is the Jacobian matrix, which contains the first order derivative of the network error (e) with respect to the network weights and bias. The improvement in weight change (x_{k+1}) is determined from Equation (5) as follows:

$$x_{k+1} = x_k - [J^T J + \mu I]^{-} J^T e \tag{5}$$

where, x = [v_{11}, v_{12}, v_{13}, …, v_{ij}; v_{01}, v_{02}, v_{03}, …, v_{0j}; w_{11}, w_{12}, w_{13}, …, v_{jk}; w_{01}, w_{02}, w_{03}, …, w_{0k}]. μ is the learning constant and I is the identity matrix.

Bayesian Regulation Algorithm (BR)

The regularization method improves the generalization process by minimizing the weights and squared errors of the network. BR changes the error performance by adding the weights and bias standard deviation. The regularization method improves the generalization process by minimizing the weights and squared errors of the network. BR changes the error performance by adding the weight and standard deviation bias which is represented as in Equation (6) [18].

$$F = \alpha E_w + \beta E_d \tag{6}$$

where, E_w and E_d are the sum of the network weights squared and the number of network errors. Both α and β are objective (regularization) function parameters. The E_w and E_d functions are determined from Equation (7) and Equation (8) as follows:

$$E_w = \frac{1}{n}\sum_{i=1}^{n}(W_i)^2 \tag{7}$$

$$E_d = \sum_{i=1}^{n}(t_i - a_i)^2 \tag{8}$$

where, n is the number of inputs to the training data, t_i is the target value of the i^{th} data and a_i of the i^{th} data output, with the weights or thresholds for the i^{th} data.

Scaled Conjugated Gradient Algorithm (SCG)

SCG is based on conjugation direction, but does not perform line searches in every iteration. It is designed to avoid time-consuming line searches. In Matlab, SCG is a network training function that updates the weights and bias values based on the scale conjugation gradient method. The network is trained as long as its inputs, weights and transfer functions are derivative functions. The SCG algorithm is based on a quadratic approximation for the error E around the point w (weight) determined from Equation (9) [18, 19].

$$E_{qw}(y) = E(\omega) + E^{/}(\omega)^{T} y + \frac{1}{2} y^{T} E^{//}(\omega) y \tag{9}$$

The critical point of Equation (9) must be found in order to determine the minimum for $E_{qw}(y)$. To determine the minimum value, the critical point of $E_{qw}(y)$ must be found using the Moller solution (linear system solution) as shown in Equation (10).

$$E^{/}_{qw}(y) = E^{//}(\omega) y + E^{/}(\omega) = 0 \tag{10}$$

SCG can produce very linear convergence in most cases. This algorithm is faster than other second-order algorithms because it avoids time-consuming row searches per iteration.

RESULTS AND DISCUSSIONS

This study compares the surface roughness prediction ability with the Levenberg-Marquardt, Bayesian Regularization and Scaled Conjugate Gradient algorithms. These algorithms have advantages over one another, therefore testing is needed to compare their performance. Performance is evaluated based on the criteria of mean square error (MSE), determinant coefficient (R^2), best predictive value, and network structure. Levenberg-Marquardt (lm), trains faster to medium-sized networks weighing up to hundreds and supports training and testing with validations and test vectors. Bayesian regularization (br) can solve the problem of noise in wrong data input and over fitting as well as inappropriate problems in network training. The scaled conjugate gradient is a network training algorithm that updates the bias values and weights based on the scaled conjugation gradient method. The number of iterations required can be more, but the number of computations per iteration is very low [19].

The network structure developed for training and testing with one hidden layer and the number of neurons in the hidden layer is selected from 1 to 20. The toolbox on MATLAB R2015a has been used for training and network testing. The network uses learningdd as a training type and the sigmoid hyperbolic tangent (tansig) sigmoid in the hidden and output layers as a function of activation. Network performance using the mean square error (MSE). The experimental data used for learning were 30 training data and 4 testing data. The input layer consists of neurons for four machining variables (V_c, f_z, a_r and a_x), while the output layer is surface roughness (R_a). The best network structure is selected from a combination of the smallest MSE and the largest determinant coefficient. The training results for the network structure above are shown in Fig. 2 and Fig. 3.

These figures indicate the performance of a 4-n-1 network structure (n values 1 to 20) trained by three algorithms. For one training time, a maximum of 1000 epochs is used. Network performance is the result of repeated network learning until the lowest combination of MSE and R^2 values on testing approaches 100%. The trend of MSE points from the training results shows that performance increases with the number of neurons in the hidden layer from 4 to 20.

In LM modeling, the trend of the network structure with neurons in the hidden layer n = 4 to n = 20 results in MSE training values that are not much different from certain values. However, for testing the results fluctuate. The best network for the LM algorithm is the 4-10-1 structure. In modeling with BR, both training and testing the MSE value on neurons n = 4 to n = 20 showed more constant results than LM. Training with the LM resulted in a smaller MSE compared to the BR algorithm. The best network for the BR algorithm is the same as for the LM network, that is, with a 4-10-1 structure. In SCG modeling, the training value varies more than the other two algorithms, but at some point the test results are better than the training value. The best structure for the SCG algorithm is the 4-9-1 structure. Based on these results, increasing the number of neurons in the hidden layer did not improve the quality of network recognition. The MSE and R^2 values for the best network structure for the three algorithms are tabulated in Table 4. These results can be observed that the overall best prediction uses the SCG algorithm. The value of R^2 on the LM and SCG has a very small difference.

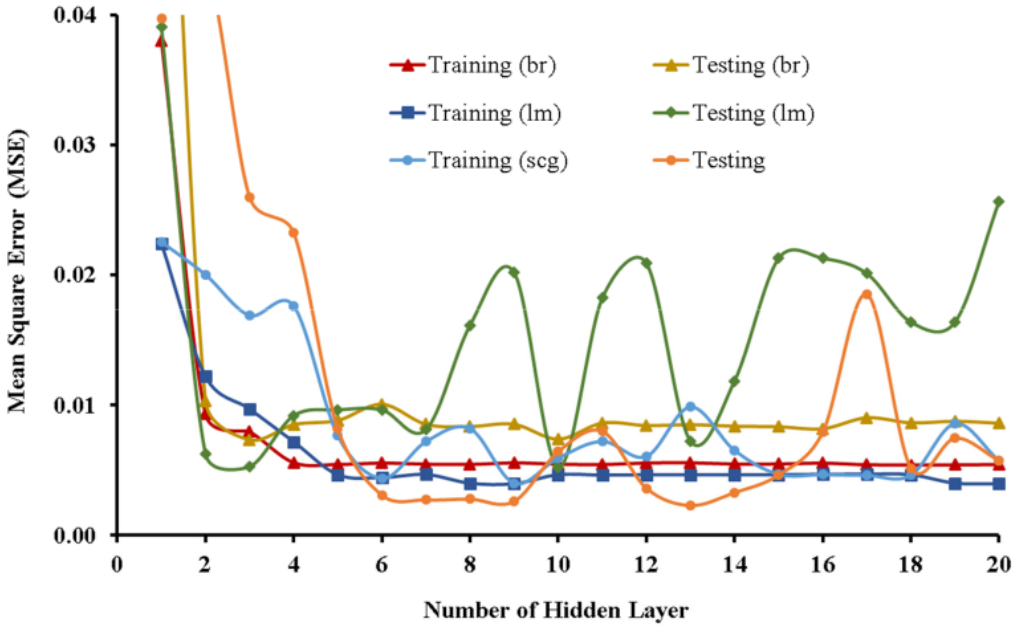

FIGURE 2. Mean square error value for each network structure

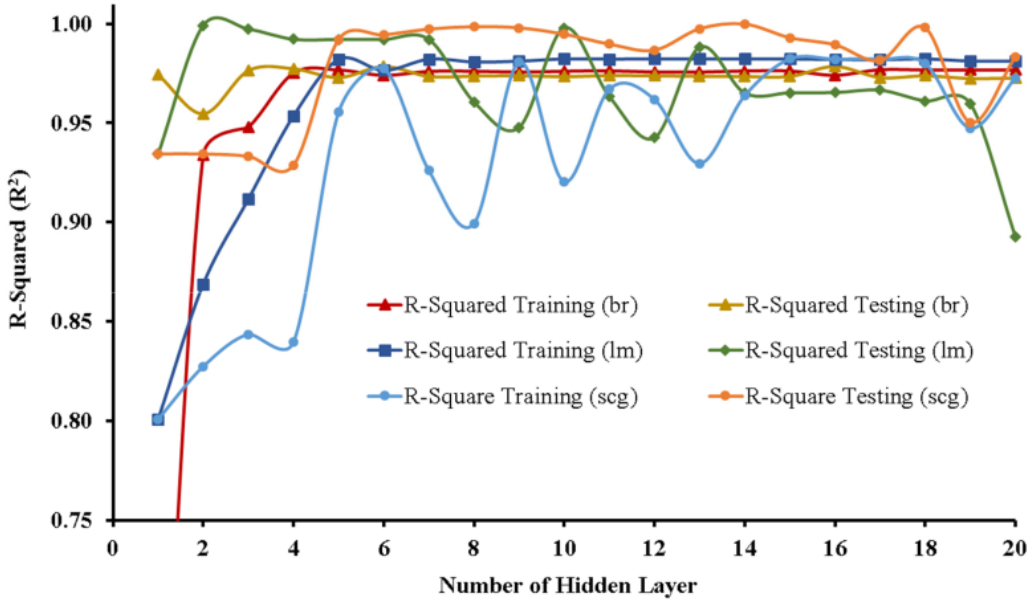

FIGURE 3. Determinant coefficient value for each network structure.

TABLE 4. Performance for LM, BR and SCG

No.	Algorithms	MSE		R^2	
		Training	Testing	Training	Testing
1	Levenberg Marquardt	0.0047	0.0050	**0.9820**	**0.9980**
2	Bayesian Regularization	0.0050	0.0074	0.9762	0.9734
3	Scale Conjugation Gradient	**0.0040**	**0.0026**	0.9810	**0.9980**

Based on the best network structure, Fig. 4 and Fig. 5 show the comparison of the surface roughness of the experimental and predicted results. The average error percentage and correlation coefficient of the three algorithms on the experimental results are shown in Table 5 below.

TABLE 5. The Average error percentage and correlation coefficient for LM, BR and SCG

No.	Algorithms	% - Error		Correlation Coefficient Exp. vs Prediction	
		Training	Testing	Training	Testing
1	Levenberg Marquardt	0.82	11.05	0.9924	0.9960
2	Bayesian Regularization	3.91	8.07	0.9807	0.9475
3	Scale Conjugation Gradient	0.64	7.08	0.9900	0.9960

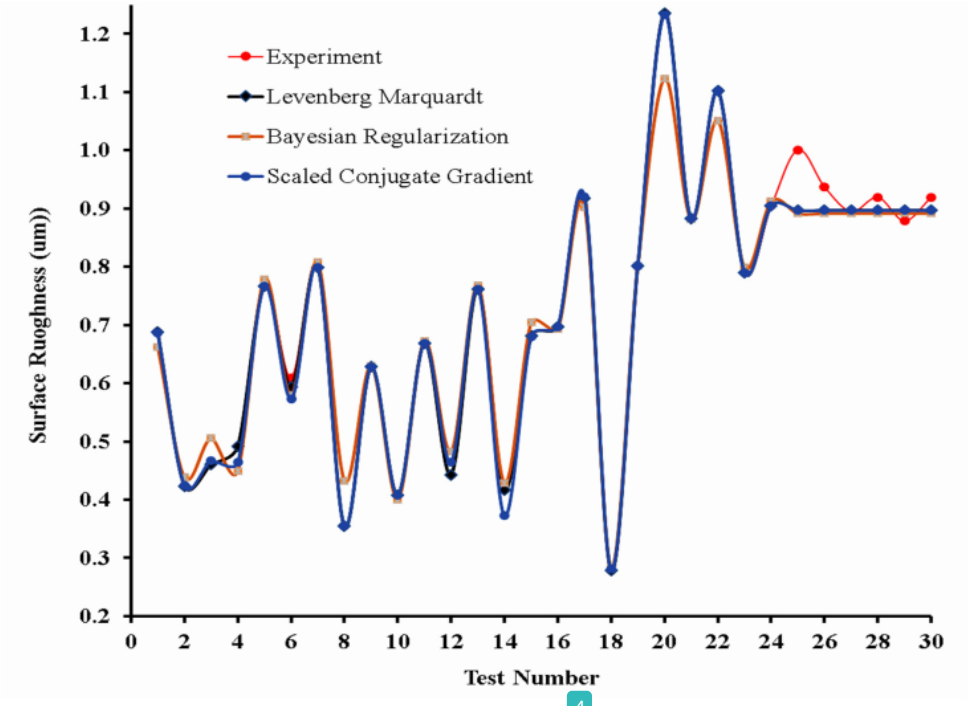

FIGURE 4. Comparison of experimental results and predictions on the LM, BR, and SCG training algorithms.

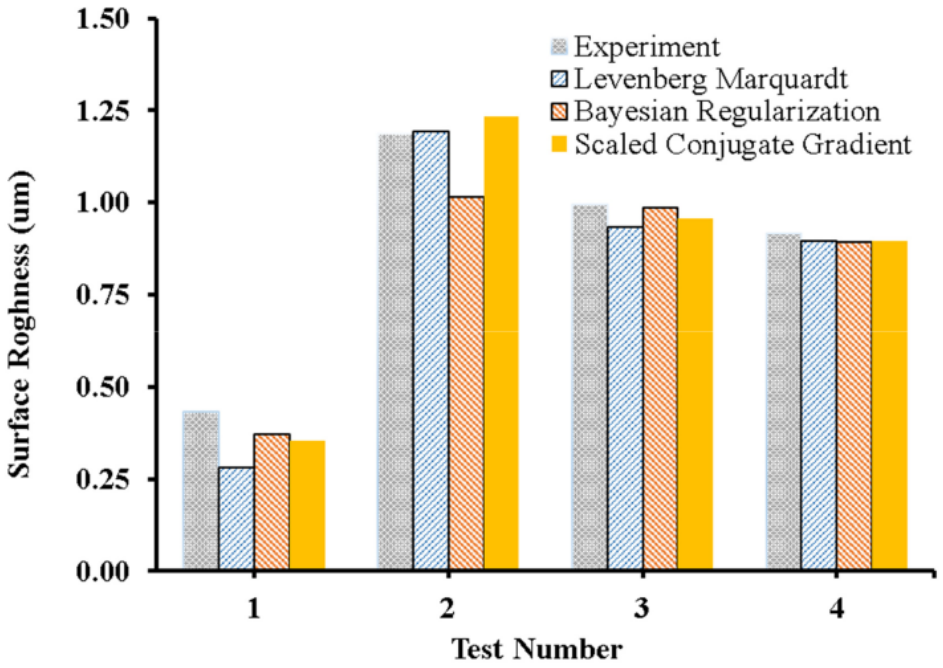

FIGURE 5. Comparison of experimental results and predictions on the LM, BR, and SCG testing algorithms.

The results of the prediction performance using LM, BR and SCG are very close to the experimental results. In several investigations conducted by researchers, the three algorithms can produce the best predictions. Ganovska et al. [20] conducted research on AWJ technology for surface roughness prediction using neural networks. He concluded that prediction by BR algorithm is better than LM. Similar results were obtained in the study conducted by Jazayeri et al. [21]. Different studies carried out by Garoosiha et al. [22] that predictions using LM are better than BR. Research by Mia and Dhar [9] obtained better surface roughness prediction based on BR algorithm than LM and SCG algorithms. Based on the study of several researchers above, the ability to predict accuracy is influenced by the size of the data to be extracted. For small and medium data sizes, LM is superior to BR, but for large data sizes, BR is better than LM [5]. Another influencing factor is the linear or non-linear level of the data being modeled.

CONCLUSIONS

This study is a comparison of the ability of the Levenberg Marquardt (LM) algorithm, Bayesian Regularization (BR), Scale Conjugation Gradient (SCG) in predicting surface roughness accuracy on AISI 1045 side milling. The results of this study can be summarized as follows:

- The best network structure for the LM and BR algorithms is 4-10-1 and the SCG algorithms are 4-9-1.
- LM, BR, and SCG produce predictive abilities that are very close to the experimental results. Based on network performance, the algorithms that produce the best MSE and R^2 are SCG, LM, and BR, respectively.
- The correlation percentage of the predictive ability of surface roughness values with experimental results was obtained for LM 99.24% (training) and 99.60% (testing). The correlation percentages for SCG were 99.00% (Training) and 99.60% (Testing). The correlation percentages for BR were 98.07% (Training) and 94.75% (Testing).

ACKNOWLEDGMENTS

This research was funded by DIPA Sriwijaya University in 2020. Based on the Rector's Decree Number: 0685/UN9/SK.BUK.KP/2020 (July-15, 2020).

REFERENCES

1. A. M. Zain, H. Haron, S. N. Qasem, S. Sharif, Appl. Math. Modell. 36, 1477–1492 (2012).
2. R. V. Rao, V. D. Kalyankar, Int. J. Adv. Manuf. Technol. 73, 1159–1188 (2014).
3. A. Suliman,B. S. Omarov, Int. J. Interact. Multimed. Artif. Intell. 5 (1), 68-72 (2018).
4. S. Bharati, M. A. Rahman, P. Podder, Md. R. A. Robel, N. Gandhi, "Advances in Intelligent Systems and Computing" (19th International Conference on Intelligent Systems Design and Applications, 2019), pp 69-79.
5. A. Priya, S. Garg, "Smart Intelligent Computing and Applications" (Proceedings of the Third-International Conference on Smart Computing and Informatics, 2020), pp. 657-664.
6. M. Kayri, Math. Comput. Appl. 21, 20, 1-11 (2016).
7. M. Yanis, A. S. Mohruni, S. Sharif, I. Yani, J. Technol. 81:6, 51-60 (2019).
8. A. P. Markopoulos, S. Georgiopoulos, D. E. Manolakos, J. Ind. Eng. Int. 12, 389-400 (2016).
9. M. Mia and N. R. Dhar, Measur 4168, 1-40 (2016).
10. E. Kilickap, A. Yardimeden, Y. H. ÇelikS, Appl. Sci. 7-1064, 1-12 (2017).
11. K. Thangarasu, S. Shankar, T. Mohanraj, K. Devendran, J. Mech. Eng. Sci. 234, 329–342 (2019).
12. Y. Chen, R. Sun, Y. Gao, J. Leopold, Measur. 4448, 1-19 (2016).
13. F. Abdulateef, J. Mech. Eng. Res. Dev. 43, 360-366 (2020).
14. K. Gupta and R. F. Laubscher, J. Manuf. Eng. 1-18 (2016).
15. M. H. S. Elmunafi, M.Y. Noordin, D. Kurniawan, Procedia Manuf. 2, 563 – 567(2015).
16. A. Z. Sultan, S. Sharif, D. Kurniawan, J. Adv. Mater. Res. 84, 809-813 (2014).
17. R. H. Myers, D. C. Montgomery, C. M. Anderson-Cook, "Response Surface Methodology", (4th Ed, John Wiley and Sons, New York, 2016).
18. O. Baghirli, "Comparison of Lavenberg-Marquardt, Scaled Conjugate Gradient And Bayesian Regularization Backpropagation Algorithms for Multistep Ahead Wind Speed Forecasting Using Multilayer Perceptron Feedforward Neural Network, Dissertation" Uppsala University, 2015).
19. L. Babani, S. Jadhav, B. Chaudhari, "Hal-ouvertes" (12th IFIP International Conference on Artificial Intelligence Applications and Innovations, 2017), pp384-395.
20. B. Ganovska, M. Molitoris, A. Hosovky, J. Pitel, J. B. Krolczyk, A. Ruggierio, G.m Krolczyk, S. Hloch, Indian J. Eng. Mater. Sci. 23, 279-287 (2016).
21. K. Jazayeri, M. Jazayeri, S. Uysal, *Advances in Data Mining-Applications and Theoretical Aspects* (16th Industrial Conference, ICDM 2016 Proceedings, 2016), pp. 80-95.
22. H. Garoosiha1, J. Ahmadi1, H. Bayat, Cogent Eng. 6, 1-14 (2019).

Levenberg-Marquardt, Bayesian-regularization, and scaled conjugate gradient algorithms for predicting surface roughness accuracy on side milling AISI 1045

ORIGINALITY REPORT

14%
SIMILARITY INDEX

PRIMARY SOURCES

1	jmerd.net Internet	191 words — 5%
2	aip.scitation.org Internet	117 words — 3%
3	Mia, Mozammel, and Nikhil Ranjan Dhar. "Prediction of surface roughness in hard turning under high pressure coolant using Artificial Neural Network", Measurement, 2016. Crossref	71 words — 2%
4	etd.aau.edu.et Internet	61 words — 2%
5	mdpi-res.com Internet	29 words — 1%
6	Azam Beg, Falah Awwad, Walid Ibrahim, Faheem Ahmed. "On the reliability estimation of nano-circuits using neural networks", Microprocessors and Microsystems, 2015 Crossref	23 words — 1%
7	copilosk.fbk.eu Internet	18 words — 1%

EXCLUDE QUOTES	OFF	EXCLUDE SOURCES	< 1%
EXCLUDE BIBLIOGRAPHY	ON	EXCLUDE MATCHES	OFF