An Improved Mathematical Model to Predict Surface Roughness Using Hybrid Method

Mohd Fadzil Faisae Ab. Rashid

***Abstract*—Surface roughness is one of the most important requirements in machining process. In order to obtain needed surface roughness, the proper setting of cutting parameters is crucial before the process take place. Therefore, an accurate mathematical model to predict surface roughness is totally needed. This research presents a hybrid method which combine conventional multiple regression analysis and genetic algorithm to improve the accuracy of mathematical model to predict surface roughness. In experiment, three independent variables: spindle speed, feed rate and depth of cut were manipulated in collecting data. Full factorials cut were performed using FANUC CNC Milling α-T14ιE. The results show that the proposed hybrid method capable to improve accuracy of model with 23% and 28% of reduction in error.**

***Index Terms*—Surface roughness, linear regression, genetic algorithm.**

I. Introduction

To realize full automation in machining, computer numerically controlled (CNC) machine tools have been implemented during the past decades. CNC machine tools require less operator input, provide greater improvements in productivity, and increase the quality of the machined part. End milling is the most common metal removal operation encountered. It is widely used to mate with other part in die, aerospace, automotive, and machinery design as well as in manufacturing industries [1].

Surface roughness is an important measure of the technological quality of a product and a factor that greatly influences manufacturing cost. The quality of the surface plays a very important role in the performance of milling as a good-quality milled surface significantly improves fatigue strength, corrosion resistance, or creep life [2]. In addition, surface roughness also affects surface friction, light reflection, ability of holding a lubricant, electrical and thermal contact resistance. Consequently, the desired surface roughness value is usually specified for an individual part, and specific processes are selected in order to achieve the specified finish [3]. Surface specification can also be a good reference point in determining the stability of a production process, because the stability of the machine is contingent on the quality of the operating part [4].

In multiple regression analysis, the model is built by considering the statistical fact. By using this method, each of variable will be assigned a coefficient which being calculated using statistical method. The coefficient determines how strong the variable will influence the overall mathematical model. This method is widely used because of effectiveness and reliability of the technique [5].

However, the model that generated from multiple regression analysis sometimes does not fulfill the requirement especially on the accuracy. This paper presents a method to improve the accuracy of mathematical model to predict surface roughness which combines multiple regression and genetic algorithm technique. By using this method, multiple regression analysis is used to obtain an early coefficient value for each variable.

II. Multiple Regression Analysis

Multiple regression analysis is a statistical technique that allows us to predict score on one variable on the basis of their score on several other variables [6]. In this case the dependent variable is surface roughness, while the independent variables are spindle speed, depth of cut and feed rate.

In general, multiple regression equation takes the form;

$$y = b_0 + b_1x_1 + b_2x_2 + \cdots + b_nx_n \tag{1}$$

where n is number of independent variables y is dependent variable and x_1, x_2 ... x_n are independent variables [7]. Referring to this problem the general form of multiple regressions is as follows;

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 \tag{2}$$

y→ surface roughness (μmm)
x_1→ spindle speed (rpm)
x_2→ feed rate (mm/min)
x_3→ depth of cut (mm)
b_0 → error coefficient
b_1, b_2, b_3 → variable coefficient

In analyzing data, SPSS 16.0 software was used, where ‘Enter’ method is applied to develop mathematical model. ‘Enter’ method is one of method that regularly used when no theoretical model is in mind. By using this method, all variables are entered into the model in one single step [8].

A. Design of Experiment

The experiment is performs by using a FANUC CNC Milling α-T14ιE. The workpiece tested is 6061 Aluminum 400mm × 100mm × 50mm. The end-milling and four-flute high speed steel is chosen as the machining operation and cutting tool. The diameter of the tool is D=16mm.

84 specimens are run in this experiment. 60 randomly selected specimens are used to build a prediction model (training set) and the remaining 24 specimens are for testing set. Spindle speed, feed rate and depth of cut are selected as consider parameters. Four levels of spindle speed: 750, 1000,

Manuscript received January 10, 2014; revised May 14, 2014.
Mohd Fadzil Faisae Ab. Rashid is with the Faculty of Mechanical Engineering, Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia (e-mail: ffaisae@ump.edu.my).

1250, and 1500 revolutions per minute (rpm), seven levels of feed rate: 152, 229, 305, 380, 457, 515, 588 millimeter per minute (mm/m), and three levels of depth of cut: 0.25, 0.76, 1.27 millimeter (mm) are determined. The data that collected from experiment is shown in Table I.

TABLE I: DATA COLLECTED FROM EXPERIMEN

Spindle speed	Feed rate	Depth of cut	Actual surface
750	152	0.25	1.351
750	152	0.76	1.3
750	152	1.27	1.629
1000	152	0.25	1.173
1000	152	0.76	1.681
1000	152	1.27	1.275
1250	152	0.25	1.276
1250	152	0.76	1.301
1250	152	1.27	1.603
1500	152	0.25	0.64
750	229	0.25	2.469
750	229	0.76	2.212
750	229	1.27	2.113
1000	229	1.27	2.291
1250	229	0.25	1.707
1250	229	0.76	1.757
1250	229	1.27	2.037
1500	229	0.25	1.503
1500	229	0.76	1.554
750	305	0.76	2.291
750	305	1.27	2.088
1000	305	0.25	3.002
1000	305	0.76	1.833
1000	305	1.27	2.037
1250	305	0.25	2.265
1250	305	1.27	1.859
1500	305	0.25	1.935
750	380	0.76	2.799
750	380	1.27	2.342
1000	380	0.25	2.265
1000	380	0.76	2.443
1000	380	1.27	2.367
1250	380	1.27	2.137
1500	380	0.25	2.392
1500	380	1.27	2.215
750	457	0.25	4.399
750	457	1.27	2.773
1000	457	0.25	3.205
1250	457	0.25	2.621
1250	457	0.76	2.037
1250	457	1.27	2.113
1500	457	0.25	2.723
1500	457	1.27	2.342
750	515	0.76	3.84
750	515	1.27	3.61
1000	515	0.25	3.485
1000	515	0.76	3.383
1250	515	0.25	2.875
1250	515	0.76	2.24
1250	515	1.27	2.367
1500	515	0.25	2.697
1500	515	1.27	2.57
750	588	0.76	4.018
1000	588	0.25	3.84
1000	588	0.76	3.586
1250	588	0.25	3.637
1250	588	0.76	2.469
1500	588	0.25	2.723
1500	588	0.76	2.316
1500	588	1.27	2.469

The data above then being analyze using SPSS software to obtain the model by using multiple regression analysis. From the analysis, the following model summary was acquired.

TABLE II: REGRESSION MODEL SUMMARY

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0.8893	0.7908	0.7796	0.3649

R is a measure of correlation between the observed value and predicted value. While ‘R Square’ indicates the proportion of variation in the criteria value which is accounted by the model. In essence, this is the measure of how good a prediction of criterion can make by knowing the predictor variables. However, ‘R Square’ tends to somewhat over-estimate the success of the model when applied to the real world, so an ‘Adjusted R Square’ is calculated which takes into account the number of variables in the model and the number of observation. This ‘Adjusted R Square’ give the most useful measure of success of the model. The analysis of the data also came out with the following coefficient value;

TABLE III: COEFFICIENT VALUE FOR VARIABLES

Variables	**Coefficient**	**Std. Error**
(Constant)	2.3025	0.2443
Spindle Speed	-0.001201	0.0001777
Feed rate	0.004402	0.0003226
Depth of Cut	-0.2982	0.1117

Therefore, the mathematical model that produced by using multiple regression analysis was as follows;

$$y = 2.3025 - 0.001201x_1 + 0.004402x_2 - 0.2982x_3 \quad (3)$$

By having this model, the percentage of average error was calculated following the formula;

$$Error\ \% = \frac{\sum_{i=1}^{n} \left| \left(\frac{y_{act_i} - y_{est_i}}{y_{act_i}} \right) \times 100 \right|}{n} \quad (4)$$

where

y_{act}: actual surface roughness from experiment

y_{est}: estimated surface roughness using mathematical model

n: number of data in testing set (n = 24)

Using the model in (4), the percentage of average error for testing set was 14.24%.

III. GENETIC ALGORITHM

Genetic algorithm (GA) is a programming technique that mimics biological evolution as a problem-solving strategy [9]. The input of the GA is a set of potential solutions to that problem, the aim of the GA being to improve them with generated initialize randomly [9].

The purpose of using GA in this study is to determine the optimum value for b_0, b_1, b_2andb_3, so that the percentage of average error will be minimized. According to initial model that produced by multiple regression analysis, the early coefficients value were determined. Thus, the upper and lower limits for each coefficient are set as follow:

Upper limit = $|2b|$

Lower limit = $-|2b|$; b: initial coefficient value

Hence, the upper and lower limits for each coefficient are presented in Table IV.

TABLE IV: UPPER AND LOWER LIMIT

Coefficient	Initial Coefficient Value	Upper Limit	Lower Limit
b_0	2.3025	4.605	-4.605
b_1	-0.001201	0.002402	-0.002402
b_2	0.004402	0.008804	-0.008804
b_3	-0.2982	0.5964	-0.5964

A. Initialization

The purpose of this step is to generate initial chromosome of solution. Since the population is set to 100, thus 100 chromosomes for initial solutions within the limit range were generated.

B. Evaluation

Each of chromosomes from initial population is being evaluated in this step. For this purpose, the data in training set is used to calculate objective function. According to the principle of GA to minimize average error, the objective function for this problem is as follow;

$$f = minimize\ \frac{\sum_{i=1}^{n}\left|\left(\frac{y_{act_i}-y_{est_i}}{y_{act_i}}\right)\times 100\right|}{n} \tag{5}$$

The definition of the equation is similar with (4) except for n. Here, n is defined as number of data in training set.

C. Selection

The purpose of the selection is to emphasize the fitter individuals in the population. It's also must be balanced with variation of crossover and mutation. When strong selection means that suboptimal, highly fit individual will take over the population, meanwhile too weak selection will result in too slow evolution. Roulette wheel selection is used to select chromosomes to be reproduced in the next step.

D. Reproduction

In 'Reproduction', a new set of chromosome will be produced by using 'Crossover' and 'Mutation' method. The selected parents from previous step will undergo the Crossover which uses 'Two Point Crossover' technique. After that 'Uniform Mutation' taken place to avoid trapping in local optimum. For this problem, probability of crossover, P_c and probability of mutation, P_m were set to 0.6 and 0.2 respectively.

E. Termination

Termination step is to stop the simulation, when certain criterion was met. In this study, the termination was set when the number of generation achieve 10,000 generations.

IV. RESULTS

After 10,000 generations, five fittest points were selected from GA. The results are presented in Table V below.

TABLE V: FIVE FITTEST POINTS FROM GA

Model	Fitness value	b_0	b_1	b_2	b_3
1	11.9209	1.5819	-0.00068	0.00417	-0.0716
2	11.9154	1.6009	-0.00069	0.00416	-0.08464
3	11.6701	1.1682	-0.00038	0.00386	0.09931
4	11.1218	1.3643	-0.00052	0.00413	-0.00508
5	11.05	1.5916	-0.00066	0.00409	-0.07481

Referring to Table V, the fitness values represent the percentage of error for each model when being applied to data in training set. To ensure that the model can be accepted in other data group, the average error for each model when used in testing set were calculated and the results were as follow.

TABLE VI: PERCENTAGE OF DIFFERENCE BETWEEN TRAINING AND TESTING SET

Model	Average error (%)		Difference
	Training set	Testing set	
1	11.9209	11.6326	2.40%
2	11.9159	11.8009	0.96%
3	11.6701	11.2155	3.90%
4	11.1218	11.7789	0.90%
5	11.05	10.7577	2.60%

The results in Table VI show that the percentage of average error in both training and testing set for all models were in accepted range. The differences for both errors were less than 4%. It shows that the model is accepted to be applied in other groups of data within the same process and specifications.

According to Table VI, the fittest model is the fifth model with 11.05 % error in training set and 10.76% error in testing set. Therefore the new mathematical model that produced by hybrid technique is;

$$y = 1.5916 - 0.00066x_1 + 0.00409x_2 - 0.07481x_3 \tag{6}$$

V. CONCLUSIONS

This research proposed a hybrid method to improve accuracy of mathematical model by combining multiple regression analysis and genetic algorithm technique. In training set, 28% improvement was achieved when average error reduced from 14.24% to 11.05%. While in testing set, the average error was also reduced from 13.94% to 10.76% with 23% improvement.

The results also conclude that the models that were developed from genetic algorithm were accepted in other groups of data with less than 4% differences. Therefore this hybrid method was proven to improve accuracy of mathematical model compared to multiple regression analysis alone.

REFERENCES

[1] J. Z. Zhang, C. C. Joseph, and E. D. Kirby, "Surface roughness optimization in an end-milling operation using the Taguchi design method," *Journal of Materials Processing Technology*, vol. 184, iss. 1-3, pp. 233-239, 2007.

[2] M. S. Lou, C. C. Joseph, and C. M. Li, "Surface roughness prediction technique for CNC end-milling," *Journal of Industrial Technology*, vol. 15, no. 1, pp. 1-6, 1999.

[3] P. G. Benardos and G. Vosniakos, "Predicting surface roughness in machining: A review," *International Journal of Machine Tools and Manufacture*, vol. 43, iss. 8, pp. 833-844, 2003.
[4] P. Armitage, G. Berry, and J. Mathew, *Statistical Methods in Medical Research*, 4th ed., Blackwell Science, 2002.
[5] C. Oğuz, C. Kurbanoğlu, and M. C. Kayacan, "Milling surface roghness prediction using evolutionary programming methods," *Materials & Design*, vol. 28, iss. 2, pp. 657-666, 2007.
[6] G. Rao, G. Rangajanardhaa, D. Rao, and M. Rao, "Development of hybrid model and optimization of surface roughness in electric discharge machining using artificial neural networks and genetic algorithm," *Journal of Materials Processing Technology*, vol. 209, iss. 3, pp. 1512-1520, 2009.
[7] M. C. Cakir, C. Ensarioglu, and I. Demirayak, "Mathematical modeling of surface roughness for evaluating the effects of cutting parameters and coating material," *Journal of Materials Processing Technology*, vol. 209, iss. 1, pp. 102-109, 2009.
[8] B. Ozcelik and M. Bayramoglu, "The statistical modeling of surface roughness in high-speed flat end milling," *International Journal of Machine Tools and Manufacture*, vol. 46, pp. 1395-1402, 2006.
[9] M. Negnevitsky, *Artificial Intelligence: A Guide to Intellilgent Systems*, Pearson Education Limited: England, 2002.

Mohd Fadzil Faisae Ab. Rashid received the bachelor degree in mechanical (industry) from Universiti Teknologi Malaysia in 2003, M.Eng (manufacturing) from Universiti Malaysia Pahang in 2007 and Ph.D in applied sciences from Cranfield University, United Kingdom in 2013.

During the early of his career, he worked with a Japanese company as a production engineer until 2004. Later, he was appointed as a tutor and lecturer in Universiti Malaysia Pahang. During that period, he was awarded a scholarship to pursue master and doctoral degrees. Currently, he is a senior lecturer in Faculty of Mechanical Engineering, Universiti Malaysia Pahang.

His research interests are in engineering optimization, particularly focused on manufacturing process and system optimization, artificial intelligent and discrete event simulation techniques.