Prediction of surface roughness during hard turning of AISI 4340 steel (69 HRC)

Regression modelling for prediction of surface roughness during hard turning of AISI 4340 steel (69 HRC)

Anupam Agrawal[a], Saurav Goel[b*], Waleed Bin Rashid[c] and Mark Price[b]

[a]Department of Business Administration, University of Illinois at Urbana-Champaign, USA

[b]School of Mechanical and Aerospace Engineering, Queen's University, Belfast, BT95AH, UK

[c]Institute of Mechanical, Process and Energy Engineering, Heriot-Watt University, Edinburgh, UK

*Corresponding author Tel.: +44-028-90975625, Email address: s.goel@qub.ac.uk, Fax: +44-028-90974148

Abstract:

In this study, 39 sets of hard turning (HT) experimental trials were performed on a Mori-Seiki SL-25Y (4-axis) computer numerical controlled (CNC) lathe to study the effect of cutting parameters in influencing the machined surface roughness. In all the trials, AISI 4340 steel workpiece (hardened up to 69 HRC) was machined with a commercially available CBN insert (Warren Tooling Limited, UK) under dry conditions. The surface topography of the machined samples was examined by using a white light interferometer and a reconfirmation of measurement was done using a Form Talysurf. The machining outcome was used as an input to develop various regression models to predict the average machined surface roughness on this material. Three regression models - Multiple regression, Random Forest, and Quantile regression were applied to the experimental outcomes. To the best of the authors' knowledge, this paper is the first to apply Random Forest or Quantile regression techniques to the machining domain. The performance of these models was compared to each other to ascertain how feed, depth of cut, and spindle speed affect surface roughness and finally to obtain a mathematical equation correlating these variables.

Keywords: Hard turning; Random Forest regression; Quantile regression

Abbreviations:

AISI	*American Iron and steel institute*
ANOVA	*Analysis of variance*
HT	*Hard turning*
HRC	*Hardness on Rockwell 'C' Scale*
CBN	*Cubic boron nitride*
CNC	*Computer numerically controlled lathe*
DOE	*Design of experiments*
MSE	*Mean squared error*
OOB	*Out of bag*
GA	*Genetic algorithm*
NN	*Neural Networks*
RFR	*Random forest regression*
RPM	*Rotation of spindle per minute*
RSM	*Response surface methodology*
var	*Variation*

Nomenclatures:

α	Constant (intercept)
ε_i	Normally distributed error
f	Feed
a_p	Depth of cut
t	the *n*umber of trees in a Random Forest specification
m	number of variables to use at each tree split in Random Forest
β	Expected increment in the response
n	Spindle speed (RPM)
R	Tool nose radius

Ra	Average value of machined surface roughness
Ra_i	per unit change in surface roughness for i^{th} experiment

1. Introduction

Hard turning (HT) process has now become a viable method to machine automotive components made of ferrous alloys with hardness above 45 HRC. On account of reduced lead time and production cost, HT eliminates some of the processing steps and procedures involved during classical machining processes for hard ferrous alloy materials; indeed, 80% of the cycle time was saved when hard turning a pinion shaft (59-62 HRC) [1]. AISI 4340 medium carbon (0.4%C) high strength martensitic steel is one such desirable material used very frequently to manufacture critical components in aerospace engineering and automotive transmissions, including the manufacture of bearings, gears, shafts, and cams, which require tighter geometric tolerances, longer service life, and good surface finish [2]. In order to carry out a hard turning operation in a deterministic fashion, a machine tool with high rigidity, and a cutting tool with high toughness, hardness, and chemical inertness supplemented with appropriate machining conditions are necessary. In its current state, hard turning differs from conventional turning on account of a number of factors including the cutting tool, workpiece, or the process itself, all of which may influence the machining outcome. These variables are:

1. Cutting tool: Tool rake angle, tool clearance angle, nose radius, tool material
2. Workpiece: Hardness, microstructure, grain size, workpiece material, etc.
3. Machining parameters: feed, depth of cut, cutting speed

Because of the many complexities involved, the task to machine a component with a determinisitic level of precision becomes a challenging one. In an attempt to understand the contribution of these variables during the hard turning of 69 HRC steel with a CBN cutting tool, 39 trials were performed in this work.

2. Literature review

Hard turning owes its popularity primarily to the capability of generating complex geometric surfaces with better form accuracy and improved tolerances in one single machining pass [3]. Previous decades of manufacturing research on hard turning have focused on finding out the influence of tool geometry [4-5], tool wear [6-9], cutting temperature, and cutting forces [10]. Based on the outcome of these studies, the suggested cutting conditions for HT are cutting speeds between 100 and 250 m/min, a feed rate in the range 0.05 to 0.2 mm/rev, and a depth of cut of less than 0.25 mm [11]. A machining trial performed by Lima and co-workers [12] on AISI 4340 steel (42 and 48 HRC) between the feed range of 0.1-0.4 mm/rev using both carbide and a PCBN insert revealed high magnitude of cutting forces and poor machined surface. Chou *et al.* [13-14] found that an increase in the tool nose radius results in an increase in the amount of specific cutting energy and thereby an improved machined surface, but at the expense of tool wear.

Surface finish is the most common tangible outcome of any machining process that can be used to characterize the quality of the machining since it dictates the functional properties of a machined component. This is because surface roughness changes the contact tribology which is central to processes ranging from adhesion to friction, wear, lubrication, and coating systems [15-16]. This, in turn, influences the corrosion resistance, fatigue resistance, creep resistance, and service life of the component. Therefore, manipulating machined surface roughness to high level of precision is a key requirement of many industrial applications. In an attempt to accomplish this task, a wide variety of soft computing tools have been applied to the domain of hard turning. Chandrasekaran *et al.* [17] reviewed number of soft computing tools *viz.* neural networks, fuzzy sets, genetic algorithms, simulated annealing, ant colony optimization, and particle swarm optimization, all of which can conveniently be applied to the machining process depending on the complexity of the variable involved. Mital *et al.* [18] have reviewed a great deal of literature concerning the application of statistical methods on finish turning a variety of materials. The statistical data applied to the experimental data in their work suggest that surface finish is primarily dependent on the type of workpiece, feed rate used, and nose radius of the cutting tool.

The primary focus of this work is to investigate the influence of various machining parameters affecting the machined surface roughness. Some of the major studies found in the literature pertaining to the optimization of hard turning are tabulated in Table 1. It can be seen from this table that none of the studies has attempted to optimize the hard turning of 69 HRC hardened AISI 4340 steel with a CBN tool, whereas it is very clear from the literature that workpiece hardness could be an important variable in influencing the machined surface roughness.

In contrast to the literature detailed above, this paper focuses on modeling the results of experiments *via* three regression models. Multiple regression modeling has been used in literature, however the prevalent analysis is focused on describing the mean of the response variable for each fixed value of the regressors, using the conditional mean of the response. This paper adds to this knowledge base by applying the Quantile Regression technique, which fits regression curves to other parts of the distribution of the response variable (and not merely the mean) and the Random Forest regression (RFR) which seeks to achive higher accuracy in predicting the outcomes. The Quantile Regression method helps to model the possibilities of different rates of change in different parts of the probability distribution of the response variable. RFR has been shown to be superior to other soft computing methods such as partial least squares, neural networks, and other techniques in the arena of species distribution prediction [19], biological activity prediction [20], and genetic applications [21], which was the motivation to apply RFR to the domain of hard turning in this work.

Table 1: Literature review of optimization studies on hard turning

Work material	Tool material	Optimization tools	Variables studied
AISI 52100	Ceramic inserts of aluminium oxide and titanium carbonitride [22]	ANOVA + RSM	Cutting velocity, feed, effective rake angle, and nose radius
	CBN cutting tool [6]	ANOVA + NN	Cutting speed, feed, workpiece hardness,

			cutting edge geometry
Aluminium alloy 390, Ductile case iron, Medium carbon steel, alloy steel, inconel	Carbide cutting tool [18]	Correlation analysis	Cutting speed, feed and nose radius (See reference stated therein)
AISI 4140 steel	TiC coated tungsten carbide [23-24]	Rotatable design + Multiple regression	Cutting speed, feed, depth of cut, time of cut
	Al_2O_3 + TiCN mixed ceramic [25]	ANOVA +Taguchi	Cutting speed, feed, and depth of cut
Mild steel	TiN-coated tungsten carbide (CNMG) [26]	RSM + GA	Speed, feed, depth of cut and nose radius
SCM alloy 440 steel	Al_2O_3 + TiC [27]	ANOVA +Taguchi	Cutting speed, feed, and depth of cut
SPK alloyed steel	Sintered carbide [28]	ANOVA + DOE	Cutting speed, feed, and depth of cut
AISI D2 Steel	Ceramic wiper inserts [29]	Multiple Regression + NN	Cutting speed, feed, and cutting time
AISI 4340 steel (below 60 HRC)	TiC/TiCN/Al_2O_3 coated carbide tipped [30]	Multiple Regression + Taguchi + RSM	Cutting speed, feed, and depth of cut
	Zirconia toughened alumina (ZTA) cutting [31]	RSM + ANOVA	Cutting speed, feed, and depth of cut
	CBN, ceramic and carbide tools [32]	Taguchi + ANOVA + Tukey- Kramer comparison,	Cutting speed, feed rate, depth of cut, workpiece hardness, and tool types

		correlation tests	
AISI H11 steel	CBN tool [33]	ANOVA + RSM	Cutting speed, feed rate, depth of cut, workpiece hardness

3. Experimental details and analysis

Longitudinal hard turning trials were performed on a Mori-Seiki SL-25Y (4-axis) CNC lathe. The workpiece specimen used was AISI 4340 steel that was hardened up to 69 HRC through heat treatment process. CBN cutting inserts (type CNMA 12 04 08 S-B) having a rake angle of 0°, clearance angle of 5°, and a nose radius of 0.8 mm were procured from Warren Tooling Limited, UK. Post-machining non-contact measurement of the surface roughness was done through a white light interferometer (Zygo NewView 5000) and the measurements were cross checked using Talysurf. In the subsequent section, the outcomes of the machining trials are discussed and analysed in terms of the statistical models. Machining by mechanical means has long been a conventional technique and unlike non-conventional machining processes it is applicable universally on almost all the real world materials [34]. Turning is one such basic machining process in which the workpiece is rotated at a particular speed (cutting speed) and the tool is fed against the workpiece (feed) at a certain level of engagement (depth of cut). Essentially, the combination matrix of these three parameters is of critical importance in determining the outcome of the process. Proper selection of these three parameters is an essential step to make the process more accurate in terms of the machined quality of the component and other favourable outcomes. Accordingly, the following experimental trials were done (Table 2) which became key input to the optimisation data. Since prior literature has shown feed (between 0.1 – 0.2 mm/rev) to be the dominant and limiting criteria for surface roughness [2], we accordingly chose closer values to cover a range of feeds (0.08, 0.09, 0.1 and 0.15) at several depths of cut and cutting speed combinations [11].

3.1. Experimental data

Table 2: Experimental data obtained from the hard turning trials

Experiment # i	Feed (f) (mm/rev)	Depth of cut (a_p) (mm)	Cutting speed (n) (RPM)	Experimental measurement of Ra (micron)
1	0.08	0.1	1608	0.502
2	0.08	0.105	1250	0.532
3	0.08	0.2	858	0.5902
4	0.08	0.2	965	0.539
5	0.08	0.452	1850	0.592
6	0.08	0.542	1072	0.5693
7	0.08	0.935	1072	0.5821
8	0.09	0.083	2145	0.667
9	0.09	0.125	1000	0.735
10	0.09	0.144	1072	0.683
11	0.09	0.2	858	0.6776
12	0.09	0.2	965	0.6179
13	0.09	0.2	1072	0.742
14	0.09	0.542	965	0.718
15	0.09	0.542	1072	0.65
16	0.09	0.753	2050	0.764
17	0.09	0.935	1072	0.625
18	0.1	0.045	2145	0.77
19	0.1	0.048	2681	0.781
20	0.1	0.133	1608	0.773
21	0.1	0.2	858	0.6687
22	0.1	0.2	965	0.7029
23	0.1	0.234	2145	0.772
24	0.1	0.352	2220	0.784
25	0.1	0.542	1072	0.6769
26	0.1	0.558	1400	0.812
27	0.1	0.754	858	0.809
28	0.1	0.935	1072	0.6966
29	0.15	0.019	2681	1.251
30	0.15	0.06	1287	1.361
31	0.15	0.1	2681	1.193
32	0.15	0.2	858	1.134
33	0.15	0.2	965	1.0854
34	0.15	0.2	1072	1.316
35	0.15	0.278	1608	1.312
36	0.15	0.542	1072	1.1083
37	0.15	0.657	1600	1.345
38	0.15	0.906	2600	1.523
39	0.15	0.935	1072	1.1337

Table 2 present the results of the average surface roughness for various combinations of tool feed (f), depth of cut (a_p), and cutting speed (n). It can be seen from Table 2 that the best value of the machined surface roughness obtained was 0.502 µm at a feed rate of 0.08 mm/rev, depth of cut of 0.1 mm, and cutting speed of 1608 RPM. A question may be asked as to why the feed rate was not lowered below this point. This is because the lowering the feed rate below a certain critical rate is governed by other factors involved in the machining operation. Below the critical feed rate, ploughing between the cutting tool with the workpiece worsens the machined surface and hence produces an undesirable outcome. From previous experience [35], 0.08 mm/rev was considered to be the critical feed rate and in order to avoid any loss to the useful life of the cutting tool, this feed was chosen as the minimum feed rate for the experiment detailed in this particular work.

3.2. *Multiple regression model*

First, multiple regression was applied to the data obtained from the experiment to predict the performance parameters of hard turning as well as for the optimization of the process. In the simplest formulation, average surface roughness (*Ra*) was considered to be the function of three linear predictors: feed (f), depth of cut (a_p), and RPM (n) which was modelled for the i^{th} experiment by assuming a linear function as follows:

$$Ra_i = \alpha + \beta_1 f_i + \beta_2 a_{p_i} + \beta_3 n_i + \varepsilon_i \qquad (1)$$

Equation (1) defines a straight line. The parameter α is the constant or intercept, and ε_i represents the error of this model estimation. The parameters β_1, β_2, and β_3 represent the expected increment in the response Ra_i per unit change in f_i, a_{pi}, n_i respectively. The linear model in equation (1) assumes that the three included variables are the most important determinants of surface roughness, and that the error ε_i is normally distributed and uncorrelated to the variables. Model A (shown later in Table 3) shows the results of the multiple regression model specified by equation (2). Standard errors that are robust to the assumptions outlined earlier are reported. These can be used to make valid statistical inferences about the coefficients, even though the data are not identically distributed. The regression results of Model A show that this model can explain 92.5% of variation in the data, and the model is therefore a very reasonable predictor of surface roughness. Model A is as follows:

$$Ra_i = -0.279 + 9.455 f_i + 0.0539 a_{p_i} + 5.61 \times 10^{-5} n_i \quad (2)$$

Among the three predictor variables, feed is the most significant predictor of surface roughness: the coefficient of feed β_1 is significant at a greater than 99.999 level (indicating that there is more than a 99.999% chance that feed has a strong dominance on the surface roughness). Similarly, cutting speed is also found to be a significant predictor of surface roughness: the coefficient β_3 is significant at a >99% level. The depth of cut is not found to be a significant predictor of surface roughness.

In figure 1, the relative importance of an individual regressor's contribution to the multiple regression model A is analysed by using four methods. Here, relative importance refers to each regressor's contribution (R^2) from univariate regression, and all univariate R^2 values add up to the full model R^2. The four methods used are as follows:

1. Averaging over orderings proposed by Lindeman, Merenda and Gold (LMG) [36]
2. Comparing what each regressor is able to explain in addition to all other regressors that are available by ascribing to each regressor the increase in R^2 when including this regressor as the last of the 3 regressors in our dataset (LAST)
3. Comparing what each regressor alone is able to explain by comparing the R^2 values from 3 regression models with one regressor only (FIRST)
4. Using the product of the standardized coeffcient and the marginal correlation, a measure proposed by Hoffman and detailed by Pratt (PRATT) [37].

In this work, 1000 bootstraps were used for replications for creating 95% confidence intervals (depicted as vertical lines within the bars in figure 1). The results show that irrespective of the method used, feed is by far the most important predictor of surface roughness, followed by cutting speed and depth of cut.

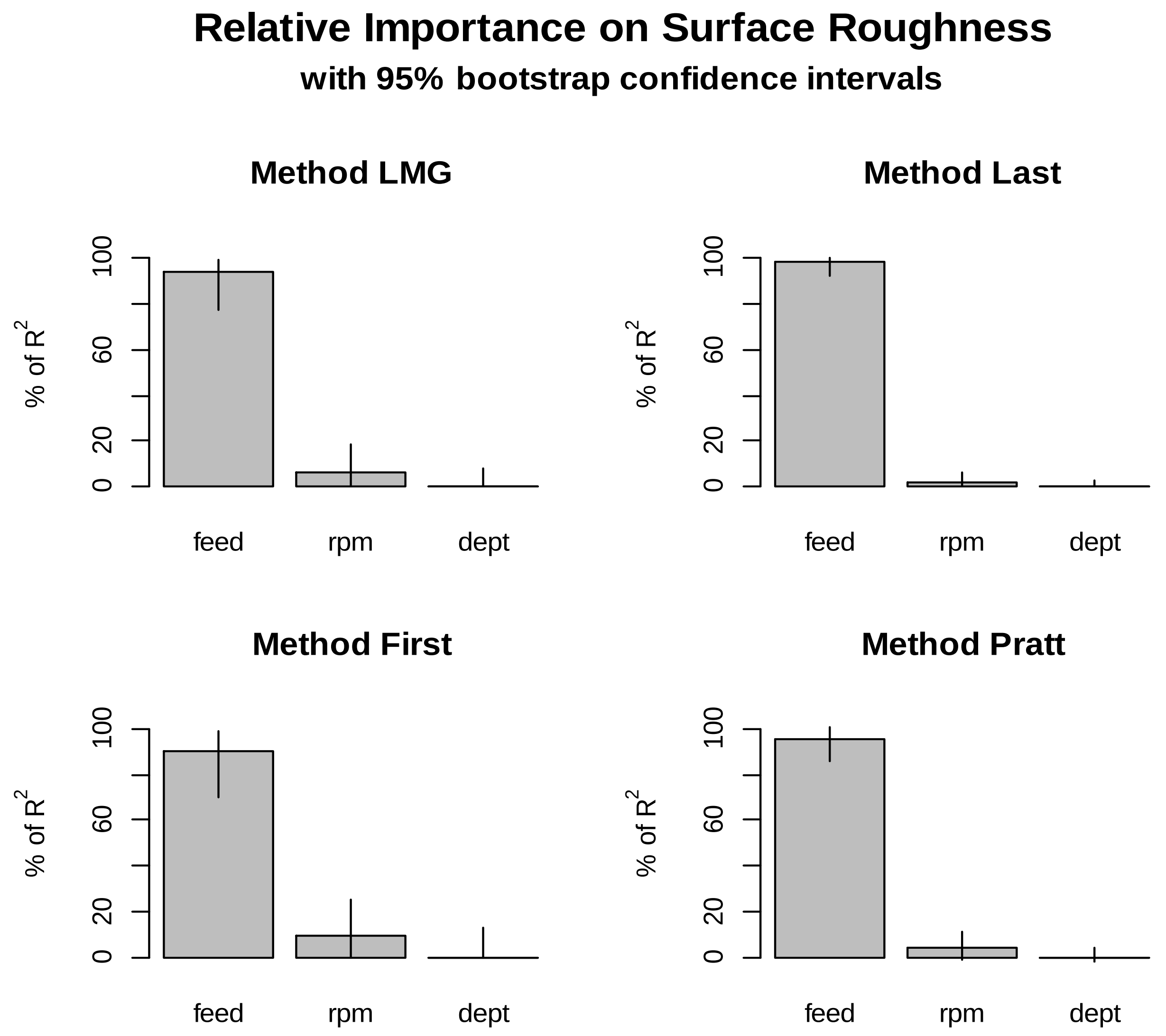

Figure 1: Relative importance of individual regressor's contribution tested by four methods

Equation (1) presupposes that the association between dependent variable Ra_i and the independent variables f_i, a_{pi}, *and* n_i is additive. However, the simultaneous influence of two independent variables (i.e. feed and depth of cut) on surface roughness may not be additive. For example, the impact of feed may depend on the depth of cut. Such an effect is known as an interaction effect, and these effects represent the combined effects of predictors on the dependent variable. In what follows, equation (1) is modified to include the interaction of each pair of independent variables, as well as the interaction of all three variables. The equation in (1) can be modified as follows:

$$Ra_i = \alpha + \beta_1 f_i + \beta_2 a_{pi} + \beta_3 r_i + \beta_4 f_i * a_{pi} + \beta_5 f_i \times n_i + \beta_6 n_i \times a_{pi} + \beta_7 f_i \times n_i \times a_{pi} + \varepsilon_i \tag{3}$$

Table 3: Multiple Regression models

Dependent Variable : Surface Roughness					
	Base Model	Interaction Models			
	A	B	C	D	E (better model)
Feed (β_1)	9.455	9.127	7.786	9.345	9.886
	(0.59)	(0.94)	(1.49)	(0.51)	(1.95)
Depth of Cut (β_2)	0.0539	-0.0452	0.0485	-0.271	0.414
	(0.06)	(0.21)	(0.05)	(0.08)	(0.31)
RPM (β_3)	5.61×10^{-5}	5.56×10^{-5}	-8.1×10^{-6}	-9.8×10^{-6}	-1.9×10^{-6}
	(2.6×10^{-5})	(2.5×10^{-5})	(9.6×10^{-5})	(2.2×10^{-5})	(2.2×10^{-5})
Feed × Depth of Cut (β_4)		0.892			-5.91
		(2.21)			(2.91)
Feed × RPM (β_5)			0.00116		-5.0×10^{-5}
			(0.00)		(0.00)
Depth of Cut × RPM (β_6)				0.000223	-0.00019
				(4.3×10^{-5})	(0.00)
Feed × Depth × RPM (β_7)					0.00335
					(0.00)
Constant	-0.279	-0.242	-0.0849	-0.164	-0.223
	(0.08)	(0.08)	(0.14)	(0.05)	(0.19)
Adjusted R^2	0.925	0.924	0.928	0.947	0.95
No. of trials	39	39	39	39	39

Values in parentheses indicate robust Standard Errors of the coefficients

Equation (3) represents an extended model where the objective is to explore whether or not the simultaneous effects of the three predictor variables (in pairs and all three together) are significant. In Table 3, Models B, C, and D show the interaction effect one pair at a time, and model E shows the interaction effect of all three variables. Adjusted R-squares have been reported for all models – these adjust for the number of explanatory terms in a model (the adjusted R-square value increases only if the new term improves the model more than would be expected by chance). Model B shows that the coefficient of β4 is not significant. Model C shows that the coefficient of β5 is not significant. Hence, models B and C are not significant improvements over model A. However, model D shows that the coefficient of β6 is significant, and therefore it can be asserted that model D is a better model to predict surface roughness than model A. Finally, model E shows that the coefficient of β7 is significant at 99.99%, and therefore model E is also a better model to predict surface roughness. Since Model E can explain a larger variation of data than model D (adjusted R^2 is higher), Model E can therefore be chosen as the preferred model.

Overall, multiple regression results, along with the interaction terms, suggest that the following model (E) is a better predictor of data than model A of equation (2).

$$Ra_i = -0.223 + 9.886 f_i + 0.414 a_{pi} - 1.93\times10^{-5} n_i - 5.91 f_i \times a_{pi} - 5.02\times10^{-5} f_i \times n_i - 0.00188 n_i \times a_{pi} + 0.00335 f_i \times n_i \times a_{pi}$$

(4)

Equation (4) explains 95% of the variation in the data, and therefore is a very good fit with the experimental data.

Overall, Multiple regression analysis helps in identifying two models that can be used for predicting surface roughness. Model A in equation (2) is a simpler model, which can be used for quicker prediction of the surface roughness, and can explain 92.5% of variation in the experimental data. Model E in equation (4) is a more complex model, but can explain 95% of variation in the experimental data.

3.3. Random Forest Regression Model

Random Forest [38] is an ensemble or divide-and-conquer approach that is similar to nearest neighbour predictor and is used to improve the performance of prediction while using regression. This decision tree methodology is based on machine learning technique [39] which asserts that it is possible to achieve higher prediction accuracy by using ensembles of trees, where each tree in the ensemble is grown in accordance with the realization of a random vector. Predictions are generated by aggregating over the ensemble. Aggregation over the ensemble results in a reduction of variance, and therefore the accuracy of the prediction is enhanced. Random Forests seek to reduce the correlation between the aggregated quantities by drawing a subset of the covariates at random. In a Random Forest, each node is split among a subset of predictors randomly chosen at that node. A Random Forest algorithm for regression is as follows:

1. Draw *t* bootstrap samples from the original data.
2. For each of the bootstrap samples, grow a regression tree by random sampling *m* of the predictors and choose the best split among those variables.
3. Predict new data by aggregating the average predictions of the *t* trees.

The Random Forest regression needs input data (the three predictors - feed, depth of cut, spindle speed, and the response variable of surface roughness), the number of trees (*t*), and the number of variables to use at each split (*m*). The random property arises out of two factors: (a) each of the *t* trees is based on a random subset of the observations, and (b) each split within each tree is created based on a random subset of *m* candidate variables.

Random Forests can be used to rank the importance of variables in a regression problem in a natural way. Essentially, a Random Forest Model tries to predict the outcome variable (surface roughness) from a group of potential predictor variables (feed, depth of cut, and cutting speed). If a predictor variable is "important" in making the prediction accurate, then by giving it random values, we must be able to obtain a larger impact on how well a prediction can be made, compared to a variable that contributes little. The variable importance score tries to capture this phenomenon. More formally, the importance of a given variable is increasing in mean square error for regression in the forest when the observed values of this variable are randomly permuted in the samples not considered for that tree (known as out of bag or OOB [38]). So, for each tree *t* of the forest, consider the associated OOB sample. Let error1 denote the mean squared error of a single tree *t* on this OOB (*t*) sample. Now, randomly permute the values of predictor x in the OOB (*t*) sample to get a perturbed sample and compute the error of predictor x on the perturbed sample. Denote this by error2. Then, the variable importance of predictor x can be denoted as $imp = \frac{1}{t}(\sum_t error2 - error1)$.

Random Forest Regression on the data was run for t = (500, 1000, 1500) and m = (1, 2, 3) to ascertain the sensitivity of the prediction to the number of trees and the number of splits. The number of trees (t) was increased until there was no increase in the variation explained by the model. Table 4 provides the importance scores for the three regressors for nine sets of regressions. A measure of the goodness-of-fit for Random Forest Regression Models is the pseudo-R^2 value, calculated from the OOB mean squared error (MSE) of the trees and the variation (*var*) of the response variable (surface roughness) explained by the model as follows: $pseudoR^2 = 1 - \frac{MSE(oob)}{\mathrm{var}}$. Table 4 also reports the pseudo-R^2 values, and the model with t=500 and m=3 provided the best fit.

Table 4: Importance scores of the three regressors for RFR (seed =99)

	t= 300			*t*= 500			*t*= 1000		
m	3	2	1	3	2	1	3	2	1
Feed	2.665	2.347	1.749	2.672	2.344	1.757	2.678	2.347	1.754
Depth of Cut	0.067	0.149	0.0361	0.066	0.138	0.346	0.064	0.139	0.351
RPM	0.116	0.298	0.488	0.118	0.302	0.468	0.116	0.311	0.463
Variation (*var*)	89.12%	89.04%	77.85%	**89.36%**	88.99%	78.81%	89.23%	89.03%	80.14%
MSE (oob)	0.0083	0.0084	0.0169	**0.0081**	0.0084	0.0162	0.0082	0.0084	0.0151
Pseudo R^2	0.991	0.991	0.978	0.991	0.991	0.979	0.991	0.991	0.981

The importance scores measure how much more helpful than random a particular predictor variable is in successfully predicting the outcome variable (surface roughness). The best fit estimation (t=500 and m=3) shows that feed is the best predictor of surface roughness, followed by spindle speed (rpm) and depth of cut.

3.4. Quantile Regression Model

Quantile Regression [40] is a method for estimating relationship between variables for all portions of a probability distribution. While multiple regressions provides a summary for the means of the distributions corresponding to the set of regressors, Quantile regression helps to compute several different regression curves corresponding to the various percentage points of the distributions and thus provides a complete picture of the data. The τ^{th} quantile could be thought of as splitting the area under the probability density into two parts: one with area below the τ^{th} quantile and the other with area 1-τ above it [40]. For example, 10% of the population lies below the 10th quantile. Thus, equation (1) for the τ^{th} quantile will reduce to the following equation (5):

$$Ra_i = \alpha^{\tau} + \beta_1 f_i^{\tau} + \beta_2 d_i^{\tau} + \beta_3 n_i^{\tau} + \varepsilon_i^{\tau} \tag{5}$$

While the Multiple Regression Model specifies the change in the conditional mean of the dependent variable (surface roughness) associated with a change in the regressors (feed, depth of cut, and spindle speed), the Quantile Regression Model specifies changes in the conditional quantile. Thus, the Quantile Regression model can be considered a natural extension of the Multiple Regression model. This model can help in inspecting the rate of change of surface roughness by quantiles. Thus, while equation (1) addresses the question “how does feed, depth of cut, and spindle speed affect surface roughness?”, it does not and cannot answer a more nuanced question: “does feed, depth of cut, and spindle speed influence surface roughness differently for samples with low surface roughness than for samples with average surface roughness?” The latter question can be answered by (for example) comparing the regression for the 50^{th} quantile with that for the 10^{th} quantile of surface roughness.

Table 5 and figure 2 show the estimated effect of feed, depth of cut, and spindle speed on surface roughness for the 10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} quantiles. The estimates shown here used bootstrapped standard errors [41] with 1000 replications.

Table 5: Quantile Regression

	Dependent variable : surface roughness				
Quantile	10^{th}	25^{th}	50^{th}	75^{th}	90^{th}
Feed	8.218 (0.42)	8.21 (0.57)	9.201 (1.10)	10.53 (0.91)	10.47 (1.04)
Depth of cut	0.0207 (0.04)	0.0281 (0.04)	0.052 (0.04)	0.0626 (0.04)	0.0362 (0.04)
RPM	6.39×10^{-5} (1.9×10^{-5})	6.04×10^{-5} (1.73×10^{-5})	4.51×10^{-5} (3.1×10^{-5})	7.29×10^{-7} (4.12×10^{-5})	0.0001 (6.7×10^{-5})
Constant	-0.213 (0.06)	-0.203 (0.06)	-0.251 (0.10)	-0.277 (0.11)	-0.34 (0.10)
Observations	39	39	39	39	39
Bootstrapped Standard errors in parentheses (1000 replications)					

According to the Multiple Regression model A (shown earlier in Table 3), for each change of one unit in feed rate, the average change in the mean of surface roughness is about 9.455 units. The quantile regression results indicate that the effect of feed on surface roughness has a lower impact for lower quantiles of surface roughness. For the 10^{th} quantile of surface roughness, for each change of one unit in feed rate, the average change in the mean of surface roughness is about 8.218 units. The Multiple Regression model overestimates this effect at the 10^{th} quantile. Similarly, for the 75^{th} quantile of surface roughness, for each change of one unit in feed rate, the average change in the mean of surface roughness is about 10.53 units. The Multiple Regression model underestimates this effect at the 75^{th} quantile.

Overall, quantile regression estimates suggest that the effect of feed on surface roughness is lower at lower levels of surface roughness and higher as surface roughness increases. The effect of spindle speed is in the opposite direction, i.e. the effect of spindle speed on surface roughness is higher at lower levels of surface roughness and reduces as surface roughness increases. However, it again becomes important as a variable at very high levels of surface roughness.

4. Comparison of Multiple Regression with Random Forest Regression

In this section, Multiple Regression and Random Forest Regression results are compared with each other to evaluate their effectiveness in predicting the value of surface roughness (the Quantile Regression methodology is not compared since that technique is used to understand how the effect of predictor variables is different at different quantiles of surface roughness, and therefore one-on-one comparison with other techniques is not possible). The values of the surface roughness obtained from the 39 experimental trials, and the predicted values of the three models presented in the work i.e. Model A (simplified multiple regression model), Model E (complex multiple regression model) and Random Forest Regression Model are correspondingly plotted in Figure 2, Figure 3, and Figure 4 to highlight the differences of each model with respect to experimental values.

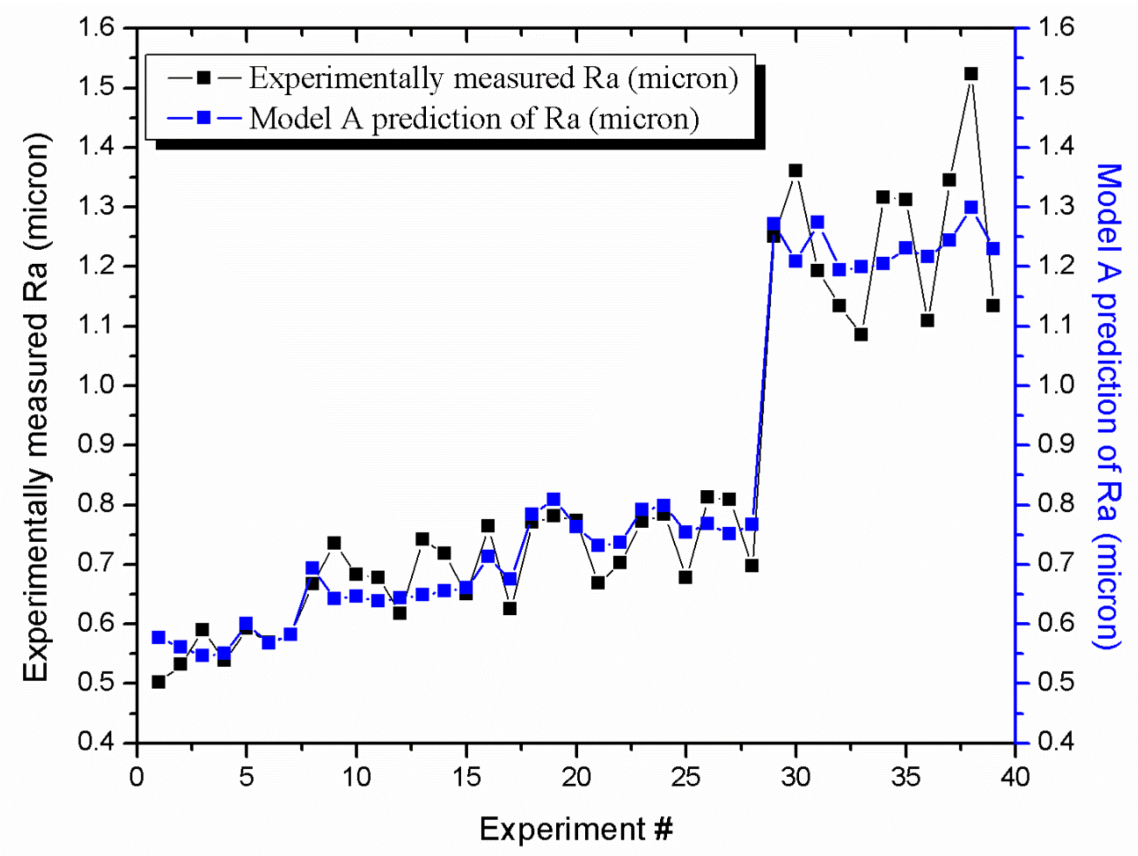

Figure 2: Comparison of experimental surface roughness with Multiple Regression Model A

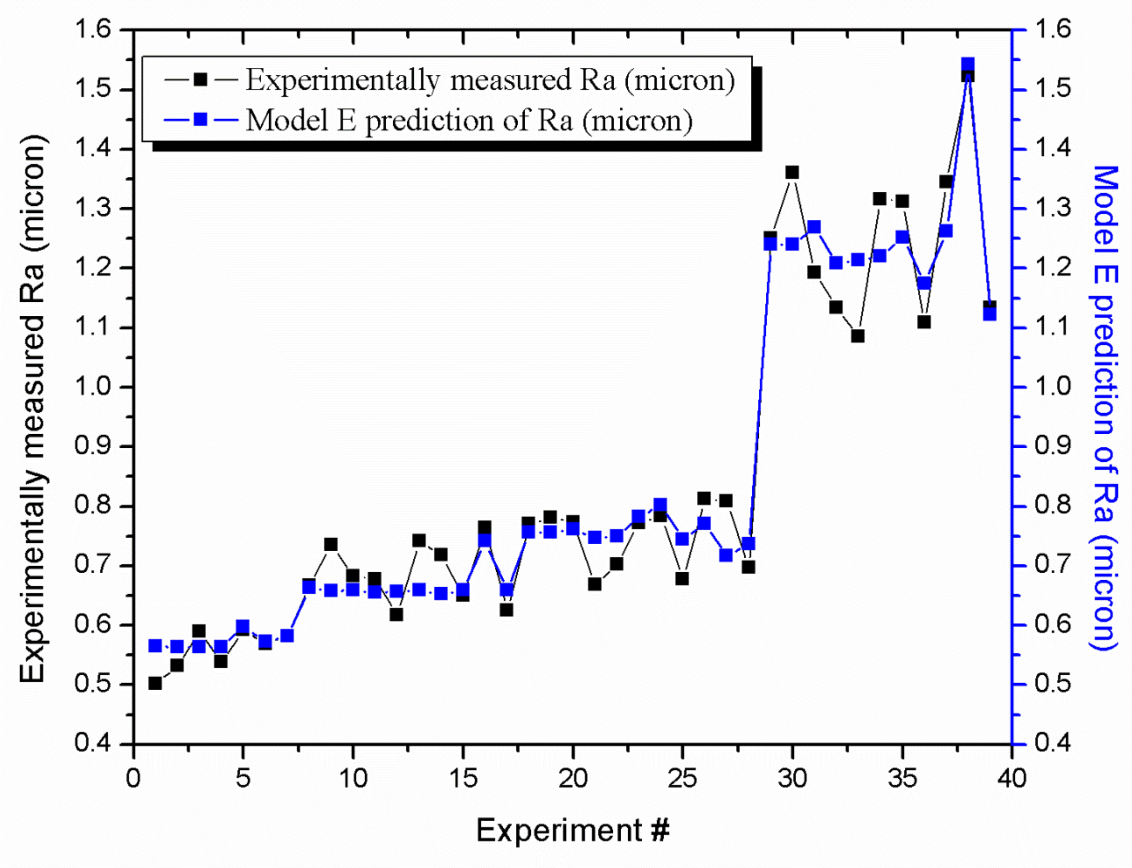

Figure 3: Comparison of experimental surface roughness with Multiple Regression Model E

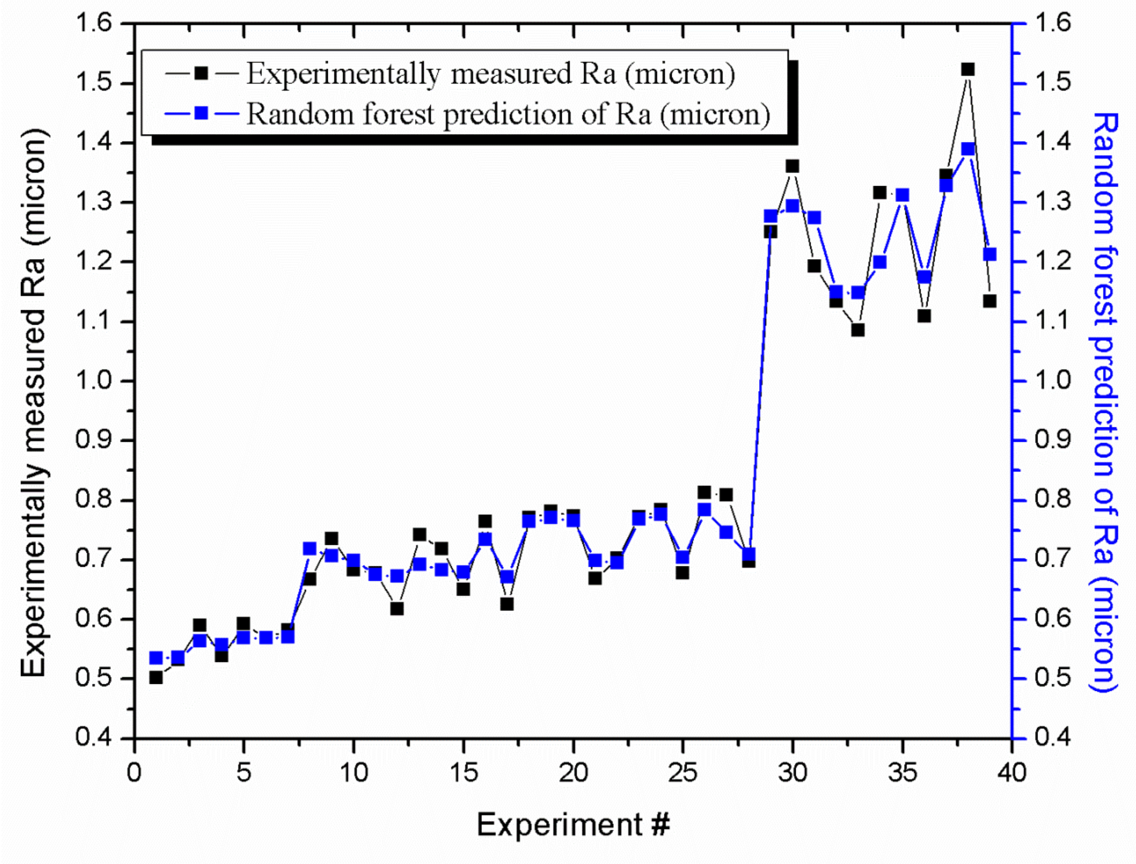

Figure 4: Comparison of experimental surface roughness with Random Forest Regression Model

From Figure 2, Figure 3, and Figure 4, it appears that while all three proposed models were good at predicting the surface roughness, however they were more accurate only when the surface roughness was below an average value of 1 micron. As the surface roughness tends to worsen beyond 1 micron, Model E becomes more accurate than Model A because it takes into consideration the pairing of the input variables. In general, the trend of the plot predicted by the Random Forest Regression Model shows a lot more consistency in the values in contrast to Model E and Model A.

Finally, the standard deviations of the differences of the predicted values from the three models versus the actual values from experiments are shown in Table 6.

Table 6: Standard deviation of the model with respect to experiments

	Model A	Model E	RFR
Standard deviation of experimental values vs. predicted values for the whole experiment	0.0740	0.0565	0.0465
Standard deviation of experimental values vs. predicted values for Ra below 1 micron	0.0479	0.0447	0.0298

It can be seen that both for the surface roughness measurement below 1 micron and for the whole set of experiments, the Random Forest Regression Model exhibits the least standard deviation compared to the Multiple Regression Models (Model A and Model E). Also, Model E shows lower standard deviation than Model A for the whole experiment, but for lower measure of the surface roughness either Model A or Model E can reliably be used.

5. Conclusions

This study presents an approach of modelling comprehensive experimental trials (39 trials) to predict the average value of machined surface roughness during hard turning of AISI 4340 steel (69 HRC) with a CBN cutting tool. For the first time, a novel approach, namely the Random Forest Regression Model has been applied to the machining domain and an excellent correlation has been found between the model and the experimental results, as the standard deviation of the predicted values from the 39 experimental result sets was only 0.0465. Among the other trials, the best value of the machined surface roughness obtained was 0.502 µm at a feed rate of 0.08 mm/rev, 0.1 mm depth of cut, and cutting speed of 1608 RPM. Based on the comprehensive models developed and proposed in this work, the following conclusions could be made:

1. Quite similar to other precision machining processes, the experimental outcome of 39 sets of trials of hard turning of AISI 4340 steel (69 HRC) showed that the value of machined surface roughness is most significantly impacted by the feed rate followed by the cutting speed and depth of cut. Although the feed rate was found to play a dominant role compared to the other two parameters, it cannot be lowered beyond a certain critical extent due to ploughing phenomena.
2. Multiple Regression Models applied to the 39 experimental datasets obtained from in-house trials revealed the following mathematical equations which could provide 92.5% and 95% accurate predictions of machined surface roughness compared with the experimental results:

   $$Ra_i = -0.279 + 9.455 f_i + 0.0539 a_{pi} + 5.61 \times 10^{-5} n_i$$

   $$Ra_i = -0.223 + 9.886 f_i + 0.414 a_{pi} - 1.93 \times 10^{-5} n_i - 5.91 f_i \times a_{pi} - 5.02 \times 10^{-5} f_i \times n_i - 0.00188 n_i \times a_{pi} + 0.00335 f_i \times n_i \times a_{pi}$$

3. While Multiple Regression Models were found suited to addresses the question “how does feed, depth of cut, and spindle speed affect surface roughness?”, further robustness check was performed using the Quantile Regression Model proposed in this work which answers the question “does feed, depth of cut, and spindle speed influence surface roughness differently for samples with low surface roughness than for those samples with average surface roughness?” It was found that the effect of feed on surface roughness is lower at lower levels of surface roughness and higher as surface roughness increases. The effect of spindle speed is in the opposite direction.
4. A novel modelling approach, i.e. Random Forest Regression, has been presented and applied to the machining process for the first time and is found to be more accurate than Multiple regression models in predicting surface roughness.

5. Multiple Regression Models were found more accurate for prediction only when the expected surface roughness is below 1 micron. Beyond this value the results showed higher deviation.

Acknowledgments:

Authors acknowledge the funding support of Ministry of Higher Education, Kingdom of Saudi Arabia for funding the PhD of WBR and an additional funding from the International Research Fellowship account of Queen's University, Belfast.

References:

1. Rashid, W.B., S. Goel, and X. Luo. *Enabling ultra high precision on hard steels using surface defect machining*. in *P 7.16- Proceedings of the 14th EUSPEN International Conference from 2nd-6th June*, 2014. Dubrovnik, Croatia: EUSPEN.
2. Rashid, W.B., S. Goel, X. Luo, and J.M. Ritchie, *The development of a surface defect machining method for hard turning processes.* Wear, 2013. **302**(1–2): p. 1124-1135.
3. Kundrák, J., B. Karpuschewski, K. Gyani, and V. Bana, *Accuracy of hard turning.* Journal of Materials Processing Technology, 2008. **202**(1–3): p. 328-338.
4. Özel, T., T.-K. Hsu, and E. Zeren, *Effects of cutting edge geometry, workpiece hardness, feed rate and cutting speed on surface roughness and forces in finish turning of hardened AISI H13 steel.* The International Journal of Advanced Manufacturing Technology, 2005. **25**(3-4): p. 262-269.
5. Thiele, J.D., S.N. Melkote, R.A. Peascoe, and T.R. Watkins, *Effect of Cutting-Edge Geometry and Workpiece Hardness on Surface Residual Stresses in Finish Hard Turning of AISI 52100 Steel.* Journal of Manufacturing Science and Engineering, 2000. **122**(4): p. 642-649.
6. Özel, T. and Y. Karpat, *Predictive modeling of surface roughness and tool wear in hard turning using regression and neural networks.* International Journal of Machine Tools and Manufacture, 2005. **45**(4–5): p. 467-479.
7. Kishawy, H. and M. Elbestawi, *Tool wear and surface integrity during high-speed turning of hardened steel with polycrystalline cubic boron nitride tools.* Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 2001. **215**(6): p. 755-767.
8. Poulachon, G., A. Moisan, and I.S. Jawahir, *Tool-wear mechanisms in hard turning with polycrystalline cubic boron nitride tools.* Wear, 2001. **250**(1–12): p. 576-586.
9. Grzesik, W., *Influence of tool wear on surface roughness in hard turning using differently shaped ceramic tools.* Wear, 2008. **265**(3–4): p. 327-335.
10. Huang, Y. and S.Y. Liang, *Modeling of Cutting Forces Under Hard Turning Conditions Considering Tool Wear Effect.* Journal of Manufacturing Science and Engineering, 2005. **127**(2): p. 262-270.
11. Bartarya, G. and S.K. Choudhury, *State of the art in hard turning.* International Journal of Machine Tools and Manufacture, 2012. **53**(1): p. 1-14.
12. Lima, J.G., R.F. Ávila, A.M. Abrão, M. Faustino, and J.P. Davim, *Hard turning: AISI 4340 high strength low alloy steel and AISI D2 cold work tool steel.* Journal of Materials Processing Technology, 2005. **169**(3): p. 388-395.
13. Chou, Y.K., C.J. Evans, and M.M. Barash, *Experimental investigation on CBN turning of hardened AISI 52100 steel.* Journal of Materials Processing Technology, 2002. **124**(3): p.

274-283.
14. Chou, Y.K. and H. Song, *Tool nose radius effects on finish hard turning.* Journal of Materials Processing Technology, 2004. **148**(2): p. 259-268.
15. Buzio, R., C. Boragno, F. Biscarini, F.B. De Mongeot, and U. Valbusa, *The contact mechanics of fractal surfaces.* Nature Materials, 2003. **2**(4): p. 233-236.
16. Fineberg, J., *Diamonds are forever - or are they?* Nature Materials, 2011. **10**.
17. Chandrasekaran, M., M. Muralidhar, C.M. Krishna, and U. Dixit, *Application of soft computing techniques in machining performance prediction and optimization: a literature review.* The International Journal of Advanced Manufacturing Technology, 2010. **46**(5-8): p. 445-464.
18. Mital, A. and M. Mehta, *Surface finish prediction models for fine turning.* The International Journal Of Production Research, 1988. **26**(12): p. 1861-1876.
19. Prasad, A.M., L.R. Iverson, and A. Liaw, *Newer classification and regression tree techniques: bagging and random forests for ecological prediction.* Ecosystems, 2006. **9**(2): p. 181-199.
20. Svetnik, V., A. Liaw, C. Tong, J.C. Culberson, R.P. Sheridan, and B.P. Feuston, *Random forest: a classification and regression tool for compound classification and QSAR modeling.* Journal of chemical information and computer sciences, 2003. **43**(6): p. 1947-1958.
21. Bureau, A., J. Dupuis, B. Hayward, K. Falls, and P. Van Eerdewegh, *Mapping complex traits using Random Forests.* BMC genetics, 2003. **4**(Suppl 1): p. S64.
22. Singh, D. and P.V. Rao, *A surface roughness prediction model for hard turning process.* The International Journal of Advanced Manufacturing Technology, 2007. **32**(11-12): p. 1115-1124.
23. Sundaram, R. and B. Lambert, *Mathematical models to predict surface finish in fine turning of steel. Part I.* The International Journal Of Production Research, 1981. **19**(5): p. 547-556.
24. Sundaram, R. and B. K LAMBERT, *Mathematical models to predict surface finish in fine turning of steel. Part II.* The International Journal Of Production Research, 1981. **19**(5): p. 557-564.
25. Aslan, E., N. Camuşcu, and B. Birgören, *Design optimization of cutting parameters when turning hardened AISI 4140 steel (63 HRC) with Al2O3+ TiCN mixed ceramic tool.* Materials & Design, 2007. **28**(5): p. 1618-1622.
26. Suresh, P., P. Venkateswara Rao, and S. Deshmukh, *A genetic algorithmic approach for optimization of surface roughness prediction model.* International Journal of Machine Tools and Manufacture, 2002. **42**(6): p. 675-680.
27. Thamizhmanii, S., S. Saparudin, and S. Hasan, *Analyses of surface roughness by turning process using Taguchi method.* Journal of Achievements in Materials and Manufacturing Engineering, 2007. **20**(1-2): p. 503-506.
28. Mehrban, M., D. Naderi, V. Panahizadeh, and H.M. Naeini, *Modelling of Tool Life in Turning Process Using Experimental Method.* International journal of material forming, 2008. **1**(1): p. 559-562.
29. Özel, T., Y. Karpat, L. Figueira, and J.P. Davim, *Modelling of surface finish and tool flank wear in turning of AISI D2 steel with ceramic wiper inserts.* Journal of Materials Processing Technology, 2007. **189**(1): p. 192-198.
30. Suresh, R., S. Basavarajappa, and G.L. Samuel, *Some studies on hard turning of AISI 4340 steel using multilayer coated carbide tool.* Measurement, 2012. **45**(7): p. 1872-1884.
31. Mandal, N., B. Doloi, and B. Mondal, *Force prediction model of Zirconia Toughened Alumina (ZTA) inserts in hard turning of AISI 4340 steel using response surface methodology.* International Journal of Precision Engineering and Manufacturing, 2012. **13**(9): p. 1589-1599.
32. Çydaş, U., *Machinability evaluation in hard turning of AISI 4340 steel with different cutting tools using statistical techniques.* Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 2010. **224**(7): p. 1043-1055.

33. Aouici, H., M.A. Yallese, K. Chaoui, T. Mabrouki, and J.-F. Rigal, *Analysis of surface roughness and cutting force components in hard turning with CBN tool: Prediction model and cutting conditions optimization.* Measurement, 2012. **45**(3): p. 344-353.
34. Brinksmeier, E. and W. Preuss, *Micro-machining.* Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2012. **370**(1973): p. 3973-3992.
35. Rashid, W.B., S. Goel, X. Luo, and J.M. Ritchie, *An experimental investigation for the improvement of attainable surface roughness during hard turning process.* Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 2013. **227**(2): p. 338-342.
36. Lindeman, R.H., P.F. Merenda, and R.Z. Gold, *Introduction to bivariate and multivariate analysis*. 1980: Scott, Foresman Glenview, IL.
37. Pratt, J.W. and T. Pukkila. *Dividing the indivisible: Using simple symmetry to partition variance explained*. in *Proceedings of the second international conference in statistics*. 1987: Tampere, Finland: University of Tampere.
38. Breiman, L., *Random forests.* Machine Learning, 2001. **45**(1): p. 5-32.
39. Breiman, L., *Bagging Predictors.* Machine Learning, 1996. **24**(2): p. 123-140.
40. Koenker, R. and G. Bassett Jr, *Regression quantiles.* Econometrica: journal of the Econometric Society, 1978: p. 33-50.
41. Hao, L. and D.Q. Naiman, *Quantile regression*. Vol. 149. 2007: SAGE Publications, Incorporated.