

## PARAMETRIC ANALYSIS OF A GRINDING PROCESS USING THE ROUGH SETS THEORY

Subham Agarwal<sup>1</sup>, Shruti Sudhakar Dandge<sup>2</sup>, Shankar Chakraborty<sup>1</sup>

<sup>1</sup>Department of Production Engineering, Jadavpur University, Kolkata, West Bengal, India

<sup>2</sup>Mechanical Engineering Department, Government Polytechnic, Murtizapur,  
Maharashtra, India

**Abstract.** *With continuous automation of the manufacturing industries and the development of advanced data acquisition systems, a huge volume of manufacturing-related data is now available which can be effectively mined to extract valuable knowledge and unfold the hidden patterns. In this paper, a data mining tool, in the form of the rough sets theory, is applied to a grinding process to investigate the effects of its various input parameters on the responses. Rotational speed of the grinding wheel, depth of cut and type of the cutting fluid are grinding parameters, and average surface roughness, amplitude of vibration and grinding ratio are the responses. The best parametric settings of the grinding parameters are also derived to control the quality characteristics of the ground components. The developed decision rules are quite easy to understand and can truly predict the response values at varying combinations of the considered grinding parameters.*

**Key Words:** *Data Mining, Rough Sets Theory, Grinding, Parameter, Rule*

### 1. INTRODUCTION

With the rapid advancements of various data analysis tools and network technology, data mining has now become an emerging area in computational intelligence which offers new concepts and methods to analyze voluminous data. Availability of a large volume of data in different forms has significantly accelerated the applications of data mining. Data mining, also known as ‘knowledge discovery from databases’, thus deals with the application of various competent tools and techniques to refine the extracted knowledge from a large database so as to envisage, categorize and characterize the mined data [1, 2]. It can identify interesting patterns in data to aid in valuable decision-making

---

Received November 18, 2019 / Accepted February 01, 2020

**Corresponding author:** Shankar Chakraborty

Department of Production Engineering, Jadavpur University, Kolkata, West Bengal, India

E-mail: s\_chakraborty00@yahoo.co.in

where the applications of the popular statistical and predictive models fail. Understanding the patterns inherent in the data sometimes becomes important when the data sources are heterogeneous and differently distributed. Data mining mainly consists of the applications of various mathematical tools for machine learning, cluster analysis, regression analysis and neural networks. Using a predetermined set of features and a training dataset, regression analysis and neural networks develop a single model. On the other hand, a machine learning algorithm develops a number of models in the form of decision rules while providing the interrelationships between various input features and the final decision. Cluster analysis can also create the same decision rules when the set of features included in each rule is independent of all other rules. The rules developed by the data mining tools are always expected to be explicit [3, 4].

Rough sets theory (RST), developed by Pawlak in 1982 [5], falls under the broad category of machine learning. Based on extraction of knowledge from the datasets, it can also provide valuable tools for data analysis and generation of independent decision rules for effective data classification. Having a strong mathematical foundation, it is well suited to efficiently solve various decision-making problems. Although it has some similarities with the fuzzy set theory, today it has evolved out as a separate discipline in data mining. Its main advantage is that it does not require any additional information about the dataset to be mined, like the probability theory in statistical approaches, membership functions in the fuzzy set theory, etc. As a non-statistical approach in data analysis, it thus classifies and analyses imprecise, uncertain or incomplete information and knowledge to generate minimal and non-redundant rule sets [6, 7].

In the contemporary automated manufacturing industries, a huge volume of data related to product design, bill of materials, production planning and control, production processes and systems, monitoring and diagnosis, etc. is being regularly captured and stored using various data acquisition tools. Valuable information in the form of rules, patterns, clusters, associations and dependencies are always expected to be hidden in the dataset collected from the manufacturing organizations. Thus, it becomes the responsibility of the production engineers to augment effective data mining tools to analyze this huge manufacturing-related dataset to identify potential patterns in various input parameters that control a manufacturing process or quality of the output products. It is observed that the RST has already been successfully applied in various domains of engineering and management decision-making, like manufacturing process control [8], quality assurance [9], supplier selection [10, 11], automotive warranty data analysis [12], operations of security forces [13], forecasting [14], etc.

In the present day manufacturing industries, grinding has been successfully applied as an efficient material removal process to almost all types of materials providing an extremely high material removal rate (MRR) (more than  $2000 \text{ mm}^3/\text{s}$ ) and ultra-precision surface finish (up to nanometer level). The precision and surface finish attained through grinding can be up to ten times better than the other machining processes, like turning and milling. Due to high hardness of the abrasives used as the cutting medium in grinding, it has now become the first choice for removing materials from the workpieces. Grinding process requires least pressure which makes it easy to hold the work material even during automated process using simple devices. It has been found out that application of grinding includes more than a quarter of total machining processes and is still showing an upward trend. Thus, in order to study the material removal mechanism in a grinding process,

while examining the possible interactions between different grinding parameters and responses, it has become critical to provide guidance for further improving the grinding quality and productivity through the identification of appropriate settings of the considered grinding parameters. On the other hand, the RST has several advantages over the other data mining tools, like it does not require any preliminary or additional information about the data to be analyzed, it provides efficient algorithms for searching out the hidden patterns in the data, it is able to find out minimal sets of data for effective pattern generation, it evaluates significance of the data, it generates sets of decision rules from the data automatically, it is easy to comprehend, it is capable to provide straightforward interpretation of the derived rules, it is suited for parallel processing, etc. Thus, in this paper, a maiden endeavor is put forward to apply the concepts of the RST to a grinding process so as study the effects of various grinding parameters on different measured responses and predict the optimal settings of those parameters.

## 2. LITERATURE REVIEW

Chadha and Lee [15] developed a new optimization tool in the form of a variable length evolution strategy for a cylindrical traverse grinding process. While considering dressing feed, grinding feed, dwell time and cycle time as the grinding parameters, Siddiquee et al. [16] presented an approach combining grey relational analysis (GRA) and principal component analysis to attain the most preferred values of surface roughness (SR), cylindricity error and diametric tolerance in a centerless cylindrical grinding process. Lee et al. [17] applied the Taguchi-sliding-based differential evolution algorithm for optimizing wheel speed, workpiece speed, depth of dressing and lead of dressing for a surface grinding process. Based on experimental studies on rough-grinding and finish-grinding processes, it was concluded that the proposed approach would provide better solutions as compared to the already adopted methods. Asiltürk et al. [18] proposed the application of an adaptive network-based fuzzy inference system for effectively predicting SR and vibration in cylindrical grinding, while taking into account workpiece speed, feed rate and depth of cut as the input parameters. Khan et al. [19] presented the application of GRA technique for optimizing an in-feed centerless cylindrical grinding process. The considered grinding parameters were dressing feed, grinding feed, dwell time and cycle time, whereas, SR and cylindricity error were the responses. Neşeli et al. [20] combined response surface methodology (RSM) and Taguchi method to find out the optimal settings of workpiece revolution, feed rate and depth of cut so as to minimize SR and vibration in an external cylindrical grinding process. Rudrapati et al. [21] applied the RSM technique to study the relationships between three grinding parameters (in-feed, longitudinal feed and work speed) and SR in a cylindrical grinding process. It was observed that the considered grinding parameters had no significant influence on SR. Using the grey-based Taguchi methodology, Köklü [22] investigated the influences of workpiece speed, depth of cut and number of slots on SR and roundness error in a grinding process. The optimal parametric mix and the most important grinding parameter were also identified. Using RSM technique, Pai et al. [23] developed regression models correlating three grinding parameters with two responses, i.e. MRR and SR, during grinding of Al6061-SiC composites. Elitist non-dominated sorting genetic algorithm (enhanced NSGA-II) was later employed to determine the optimal grinding

conditions. Pawar and Rai-Kalal [24] adopted NSGA-II technique for determining the optimal operating levels of wheel speed, workpiece speed, depth of dressing and lead of dressing to minimize production cost, production rate and SR in a grinding process. Winter et al. [25] applied geometric programming and a weighted max-min model for single as well as multi-objective optimization of an internal cylindrical grinding process. The corresponding Pareto optimal solutions were also identified to enhance the eco-efficiency of the considered grinding operation. Khare and Agarwal [26] developed an analytical model representing the relationship between SR and chip thickness for surface grinding of AISI 4340 steel material. Aleksandrova [27] applied generalized utility function for parametric optimization of different dressing parameters in a cylindrical grinding operation. Deng et al. [28] applied genetic algorithm (GA) technique to solve a multi-objective optimization model having minimum processing time and optimal carbon efficiency as the two objectives. The optimal parametric combination of different grinding parameters was also identified. Rudrapati et al. [29] considered three input parameters, i.e. in-feed, longitudinal feed and workpiece speed, and later applied multi-objective GA to minimize SR and vibration in traverse cut cylindrical grinding operation of stainless steel material. Çaydaş and Çelik [30] combined RSM and GA techniques to optimize speed of the workpiece, depth of cut and number of grooves in cylindrical surface grinding operation of AISI 1050 steel material. The effects of those grinding parameters on SR of the machined components were also investigated. Chang et al. [31] conducted an orthogonal test to investigate the influences of wheel speed, workpiece speed and grinding depth on surface integrity of a bearing raceway. Based on the experimental data, two support vector machine models were proposed in order to significantly reduce the optimization time and derive the global optimal solution. Kuo et al. [32] proposed a multi-criteria model to obtain the optimal parametric setting of a grinding process while taking into consideration SR and MRR as the responses. The effects of minimum quantity lubrication on MRR, surface integrity and temperature while grinding Ti-6Al-4V workpiece material were also studied. Considering the stochastic nature of a grinding process and based on orthogonal experiment method, Ming et al. [33] optimized different input parameters for a five-axis blade grinding setup. Liu et al. [34] integrated signal-to-noise analysis with GRA technique to optimize different grinding parameters for attaining the most desired values of MRR and grinding efficiency.

It can be observed from the exhaustive review of the past research that the investigation of the influences of different grinding parameters, like grinding wheel speed, workpiece rotational speed, depth of cut, cutting speed etc. on various responses, such as MRR, SR, vibration, cylindricity error, grinding efficiency etc., and determination of the optimal combinations of those grinding parameters have been the main topics of interest amongst the researchers. Several optimization tools in the form of GA, NSGA-II, GRA, utility theory, etc. have been implemented to fulfill the above-cited objectives. The applications of those tools and techniques often lead to sub-optimal or near-optimal solutions, and identify the optimal parametric settings of the considered grinding processes which are sometimes difficult to set and maintain in the existing machining systems. Thus, in this paper, the application of a data mining tool, in the form of the RST, is proposed to analyze the experimental dataset of a grinding process and generate the corresponding 'If-Then' decision rules to visualize the effects of different grinding parameters on three important responses and guide the concerned production engineers to identify the most appropriate parametric mix for the said grinding process for attaining the desired quality characteristics

of the ground components. These decision rules follow a general structure, i.e. if the machining conditions are met then certain response values can be attained or predicted. They are probably the most interpretable prediction models, semantically resembling the natural language and human thinking process. The developed rules aid in solving complex machining problems, providing explanations of how the final decisions have been arrived at and why a particular decision has been made. This rule generation process has been proved to have high speed and scalability. The developed rules for the considered grinding process can also act a repository and executable knowledge base to facilitate the decision-making process in the domain of grinding technology.

### 3. ROUGH SETS THEORY

In the present day automated manufacturing industries, it has now become quite essential to analyze the huge dataset to correctly estimate the real nature of knowledge inherent in it. For this purpose, 'If...Then' rules have emerged as a reliable tool for decision-making while effectively representing information or bits of knowledge. The expression of 'If...Then' rule attains a form, like 'If *condition* Then *conclusion*'. In order to demonstrate the importance of 'If...Then' rules in data mining, the data presented in Table 1 can be cited. In this data matrix, there are three input parameters, each having two different levels and three responses with three varying levels. For example, in the first experimental run, when all the three input parameters are set at '0' (minimum) level, all the responses would have 'low' observations.

**Table 1** Illustrative dataset

Experiment run	Input parameter			Response		
	$a_1$	$a_2$	$a_3$	$r_1$	$r_2$	$r_3$
1	0	0	0	Low	Low	Low
2	0	1	1	Low	Low	Low
3	0	0	0	Medium	High	High
4	1	1	1	High	High	High
5	1	0	0	High	Low	Low

In the initial dataset containing a large number of observations, it has been sometimes noticed that many attributes and responses are duplicative in nature which may be responsible for unwanted bulkiness of the dataset. Hence, it has become mandatory to reduce the numbers of attributes and responses in the original dataset to enable extraction of explicit knowledge with framing of simple rules. The RST technique helps in reduction of number of attributes or responses while estimating the dependency between two or more of them. The attributes/responses having higher dependency indexes as compared to the predefined threshold value are removed from the initial dataset. When the dependency index is greater than the threshold value, either the first or second attribute/response of a common pair is excluded from the dataset without losing any valuable information. The dependency index can be calculated using the following equation [8]:

$$K(a_i, a_j) = \sum_{L \in a_j^*} \frac{|a_i(L)|}{N} \quad (1)$$

$$a_i(L) = \cup \{Y \in a_i^* | Y \subseteq L\} \quad (2)$$

where  $a_i^*$  and  $a_j^*$  are the equivalence classes of attributes  $a_i$  and  $a_j$  respectively (the equivalence class is the set of objects having the same value for attributes  $a_i$  and  $a_j$ ),  $L$  is the equivalence class of  $a_j$ ,  $Y$  is the equivalence class of  $a_i$ ,  $N$  is the total number of objects in the dataset,  $|\bullet|$  is the cardinality of a set (number of elements in the set) and  $a_i(L)$  is the lower approximation of set  $L$  over attribute  $a_i$ .

In RST, a dependency index of  $K(a_i, a_j) = 0$  signifies that the attributes  $a_i$  and  $a_j$  are independent to each other, whereas,  $K(a_i, a_j) = 100$  implies that they are totally dependent to each other. Thus, elimination of any one of them would not affect extraction of the knowledge from the dataset. But, the attributes cannot be eliminated only by observing the value of  $K(a_i, a_j)$ , the value of  $K(a_j, a_i)$  also needs to be checked. Elimination of the attributes can only be possible if  $\min\{K(a_i, a_j), K(a_j, a_i)\}$  is greater than the predefined threshold value. Thus, determination of the corresponding threshold value plays an important role in generation of the decision rules. If the threshold value is estimated to be high, a greater number of incompetent attributes exists in the dataset, making formation of the rules highly complicated. On the other hand, if its value is low, there remains a high chance of many useful attributes getting eliminated with loss of valuable information. Hence, it is always recommended to set the threshold value based on the prevailing situation and experts' opinions. Traditionally, the threshold value is set as 85-90%.

Based on the dataset of Table 1, and using Eqs. (1) and (2), the corresponding dependency level matrix for the considered attributes is developed, as shown in Table 2. From this table, it can be noticed that  $K(a_2, a_3) = K(a_3, a_2) = 100$ , and  $K(r_2, r_3) = K(r_3, r_2) = 100$ , i.e. input parameters  $a_2$  and  $a_3$ , and responses  $r_2$  and  $r_3$  are dependent on each other (threshold value is taken as 90). Thus, either  $a_2$  or  $a_3$  and  $r_2$  or  $r_3$  can be eliminated from the initial dataset, while keeping other attributes remain intact as earlier. In this illustrative example,  $a_3$  and  $r_3$  are eliminated, and a new dataset is formed in Table 3.

**Table 2** Calculated values of dependency level

Attribute	Input parameter			Response		
	$a_1$	$a_2$	$a_3$	$r_1$	$r_2$	$r_3$
$a_1$	-	0	0	100	0	0
$a_2$	0	-	100	20	0	0
$a_3$	0	100	-	20	0	0
$r_1$	40	0	0	-	0	0
$r_2$	0	0	0	60	-	100
$r_3$	0	0	0	60	100	-

**Table 3** Reduced dataset

Experiment run	Input parameter		Response	
	$a_1$	$a_2$	$r_1$	$r_2$
1	0	0	Low	Low
2	0	1	Low	Low
3	0	0	Medium	High
4	1	1	High	High
5	1	0	High	Low

From the reduced dataset, with the help of  $k$ -means or any other clustering algorithm, the considered attributes are now discretized with continuous numeric values. It helps the attributes to be well organized themselves in different clusters/groups to provide the rules more proficiency. Now, a decision rule generation algorithm is applied to extract ‘If...Then’ rules from the reduced set of the attributes categorized into appropriate number of clusters. The decision rule generation algorithm is presented as follows [8]:

*Step 1:* Initialize:  $A = \{a_1, a_2, \dots, a_n\}$ ;  $R = \{r_1, r_2, \dots, r_m\}$

*Step 2:* Evaluate  $X_{ij} = A_i \cap R_j$  for  $i = 1, 2, \dots, p$ ;  $j = 1, 2, \dots, q$

*Step 3:* For each  $X_{ij} \neq \phi$ , a rule is assigned as

If  $a_1 = V(A_i, a_1)$  and...and  $a_n = V(A_i, a_n)$  Then  $r_1 = V(R_j, r_1)$  and...and  $r_m = V(R_j, r_m)$

[P, Q, C, QTY] [T]

where  $P = \frac{|X_{ij}|}{|A_i|}$ ;  $Q = \frac{|X_{ij}|}{|R_j|}$ ;  $C = \frac{|X_{ij}|}{N}$ ;  $QTY = |X_{ij}|$ ;  $T = P + Q + C$

In the above algorithm, the ‘‘If’’ statement contains the input or independent parameters, whereas, the ‘‘Then’’ statement consists of the dependent parameters or responses. Here, T signifies the total weight (relative importance) assigned to a rule for effective decision-making. The higher the value of T, the greater the weight of a particular rule is. The maximum value of T identifies a rule to be the optimal one among all the generated rules, and it has the highest chance of occurrence when the whole system is repeated again and again.

Using the reduced dataset of Table 3, two sets  $A = \{a_1, a_2\}$  and  $R = \{r_1, r_2\}$  are initially generated, and based on the steps as presented for the rule generation algorithm, the following rules are formulated.

a) Rules for single response:

For response ( $r_1$ ):

Rule 1: If  $a_1 = 0$  Then  $r_1$  is Low.

[P = 66.67%, Q = 100%, C = 40%, QTY = 2][T = 206.67%]

Rule 2: If  $a_1 = 1$  Then  $r_1$  is High.

[P = 100%, Q = 100%, C = 40%, QTY = 2][T = 240%]

Rule 3: If  $a_1 = 0$  and  $a_2 = 0$  Then  $r_1$  is Medium.

[P = 50%, Q = 100%, C = 20%, QTY = 1][T = 170%]

From the above-generated rules, it can be concluded that a value of P = 100% in Rule 2 indicates that all the elements present in the dataset with condition  $a_1 = 1$  satisfy this rule. On the other hand, a value of Q = 100% in Rules 1, 2 and 3 implies that all the elements present in the dataset, having response  $r_1$  as low, high and medium, satisfy Rules 1, 2 and 3,

respectively. It can also be observed that for Rules 1 and 2, 40% of the observations in the dataset ( $C = 40\%$ ) are covered by these rules. For Rule 1, a value of  $QTY = 2$  represents the total number of elements (cases) that follow this rule. Amongst the three generated rules, Rule 2 with the maximum T value of 240% is supposed to be the strongest rule. Although Rule 3 encompasses both the input parameters, it has poor strength ( $T = 170\%$ ).

For response ( $r_2$ ):

- Rule 1: If  $a_2 = 0$  Then  $r_2$  is Low.  
 $[P = 66.67\%, Q = 66.67\%, C = 40\%, QTY = 2][T = 173.34\%]$
- Rule 2: If  $a_1 = 0$  and  $a_2 = 1$  Then  $r_2$  is Low.  
 $[P = 100\%, Q = 33.33\%, C = 20\%, QTY = 1][T = 153.33\%]$
- Rule 3: If  $a_1 = 1$  and  $a_2 = 1$  Then  $r_2$  is High.  
 $[P = 100\%, Q = 50\%, C = 20\%, QTY = 1][T = 170\%]$
- Rule 4: If  $a_1 = 0$  and  $a_2 = 0$  Then  $r_2$  is High.  
 $[P = 50\%, Q = 50\%, C = 20\%, QTY = 1][T = 120\%]$

For response  $r_2$ , the value of T is maximum for Rule 1, identifying it as the strongest rule.

b) Rules for two responses:

- Rule 1: If  $a_1 = 0$  Then  $r_1$  is Low and  $r_2$  is Low.  
 $[P = 66.67\%, Q = 100\%, C = 40\%, QTY = 2][T = 206.67\%]$
- Rule 2: If  $a_1 = 0$  and  $a_2 = 0$  Then  $r_1$  is Medium and  $r_2$  is High.  
 $[P = 50\%, Q = 100\%, C = 20\%, QTY = 1][T = 170\%]$
- Rule 3: If  $a_1 = 1$  and  $a_2 = 1$  Then  $r_1$  is High and  $r_2$  is High.  
 $[P = 100\%, Q = 100\%, C = 20\%, QTY = 1][T = 220\%]$
- Rule 4: If  $a_1 = 1$  and  $a_2 = 0$  Then  $r_1$  is High and  $r_2$  is Low.  
 $[P = 100\%, Q = 100\%, C = 20\%, QTY = 2][T = 220\%]$

These rules as developed taking into consideration two responses simultaneously are supposed to be more reliable and useful as compared to those generated for only one response. For two responses, the value of T for Rules 3 and 4 is the maximum which implies that both these rules can collectively extract the optimal information from the considered dataset.

#### 4. ROUGH SETS THEORY IN A GRINDING PROCESS

Grinding is a machining process where a high volume of unwanted material is rapidly removed from the workpiece surface with the help of the abrasive grinding wheel or the grinder used as a cutting tool [35]. It is principally used as a fine finishing process which results in achievement of high surface quality and dimensional accuracy of the machined parts/components. In this process, each grain of abrasive on the grinding wheel removes material from the workpiece in the form of small chips through shear deformation. A grinding setup usually consists of a bed with a fixture to guide and hold the workpiece, and a power-driven grinding wheel with hard abrasives revolving with the required rotational speed. In order to cool the workpiece during the grinding operation, coolants (e.g. water, light duty oil, wax, heavy duty emulsifiable oil etc.) are also applied. The abrasives commonly used in the grinding wheels are aluminum oxide, silicon carbide, ceramics, diamond and cubic boron nitride. On the other hand, the workpiece materials include

aluminum, brass, plastics, cast iron, mild steel and stainless steel. In manufacturing industries, there are huge applications of grinding operation, e.g. surface finishing, slitting and parting, descaling, deburring, finishing of flat as well as cylindrical surface, and grinding and resharpening of tools and cutters.

Keeping in mind the large applicability of grinding operations, in this paper, a dataset is chosen where nine experiments have been conducted on low alloy steel workpiece samples ( $60 \times 40 \times 8$  mm size) using a vitrified bonded alumina grinding wheel. Spindle speed (SS) (in rpm), depth of cut (DOC) (in mm) and type of the cutting fluid (TCF) have been considered as the input grinding parameters. On the other hand, SR (Ra) (in  $\mu\text{m}$ ), amplitude of vibration (V) (in  $\mu\text{m}$ ) and grinding ratio (G-ratio) have been treated as the process outputs (responses). Spindle speed is the rotational speed of the grinding wheel and depth of cut is the thickness of material being removed during the grinding operation. Higher depth of cut provides more MRR while enhancing productivity of a grinding process. During the grinding operation, material removal takes place by abrasion, resulting in generation of substantial amount of heat. To cool the workpiece, the coolant is used to avoid overheating and meet the dimensional tolerances. The Ra (arithmetic mean roughness or centre line average roughness) symbolizes surface quality of the machined components. It is one of the most important parameters for measuring SR. If there are large form deviations in the machined surface, the corresponding Ra value would be high; otherwise for smooth surface, lower values of Ra are obtained. During the grinding operation, the maximum distance to which the grinding wheel goes from its central position is termed as the amplitude of vibration. The G-ratio indicates the efficiency of the grinding operation and can be defined as the ratio of MRR to wheel wear rate. For each of the grinding parameters, three different operating levels have been considered. The detailed experimental plan along with the measured values of the three responses is provided in Table 4. In this table, the numbers enclosed inside the parentheses show the respective operating levels of the considered grinding parameters. Now, this experimental dataset for the grinding operation is analyzed using the principle of the RST so as to identify those input parameters which are responsible for controlling the output characteristics of the ground parts/components. At first, data preprocessing in the form of attribute reduction and clustering of the considered attributes are performed. Table 5 exhibits the dependency indexes as computed for each pair of the attributes and smaller values of those indexes (all the values are less than the threshold limit of 90%) prove the independency of all the attributes as considered in this grinding process. It is worthwhile to mention that in Table 5, the values of two dependency indexes  $R(\text{SS}, \text{G-ratio})$  and  $R(\text{G-ratio}, \text{SS})$  are obtained as 33.33% and 100%, respectively. But, as the minimum of them, i.e. 33.33% is less than the predetermined threshold value of 90%, both of them can be treated as entirely independent attributes. Along with the data reduction, the measured responses are also grouped into appropriate number of clusters using  $k$ -means algorithm to convert their continuous values into separate distinguishable ranges.

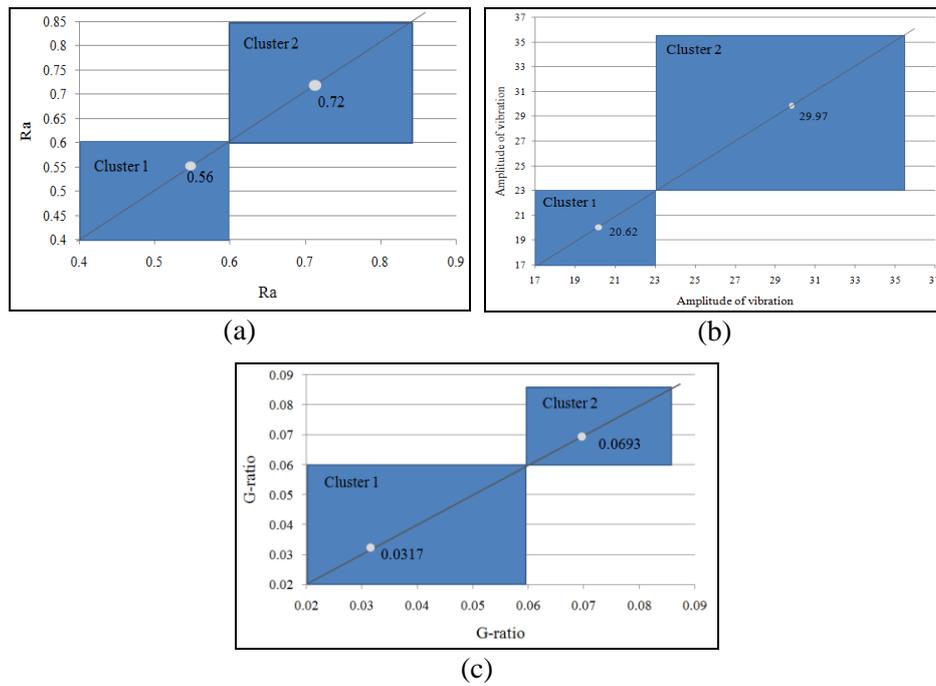
**Table 4** Experimental dataset for the grinding process

Exp. No.	Grinding parameter			Response		
	SS	DOC	TCF	Ra	V	G-ratio
1	2430 (1)	0.02 (1)	Coolant (1)	0.48	18.22	0.0253
2	2430 (1)	0.03 (2)	Water (2)	0.56	21.32	0.0262
3	2430 (1)	0.04 (3)	Coolant+Water (3)	0.57	26.23	0.0232
4	2560 (2)	0.02 (1)	Water (2)	0.61	22.32	0.0356
5	2560 (2)	0.03 (2)	Coolant+Water (3)	0.65	31.22	0.0323
6	2560 (2)	0.04 (3)	Coolant (1)	0.77	29.57	0.0476
7	2850 (3)	0.02 (1)	Coolant+Water (3)	0.72	26.45	0.0643
8	2850 (3)	0.03 (2)	Coolant (1)	0.8	31.56	0.0656
9	2850 (3)	0.04 (3)	Water (2)	0.65	34.78	0.0781

**Table 5** Dependency indexes for various grinding attributes

Attribute	SS	DOC	TCF	Ra	V	G-ratio
SS	-	0	0	0	0	33.33
DOC	0	-	0	0	0	0
TCF	0	0	-	0	0	0
Ra	66.67	0	0	-	33.33	33.33
V	33.33	33.33	33.33	55.56	-	33.33
G-ratio	100	0	0	44.44	33.33	-

In Fig. 1, the values of all the considered responses for this grinding process are clustered into two separate groups. For Ra and amplitude of vibration (both are non-beneficial characteristics requiring their lower values), the two formed clusters for them are respectively designated as 'Low' and 'High'. Here, lower values of Ra and amplitude of vibration are always preferred. On the other, for G-ratio (being a beneficial characteristic requiring only higher value), the corresponding clusters are also respectively termed as 'Low' and 'High'. But, for G-ratio, higher values are always desired. The number of classes in which the responses are to be segregated also plays an important role in subsequent generation of the decision rules. If the number of clusters is high, each generated rule would encompass a small number of elements. On the other hand, when the number of clusters is too small, interpretation of the rules would then become complicated. Thus, it is always recommended to choose the number of clusters in such a way so as to make a compromise between simplicity of the rules and the level of knowledge extraction. The details of the cluster analysis results for the three responses of the grinding process are provided in Table 6. In this table, the third and fourth columns respectively denote the mean and range values for each of the clusters formed for the considered responses. On the other hand, the column five represents the specific objects (experimental runs) and the last column denotes the total number of objects in each of the formed clusters.



**Fig. 1** Clustering of the considered responses

**Table 6** Details of the formed clusters for the responses

Response	Cluster number	Mean	Range of each cluster	Objects in each cluster	Total number of objects in each cluster
Ra	Cluster 1	0.56	0.40-0.60	1,2,3,4	4
	Cluster 2	0.72	0.60-0.85	5,6,7,8,9	5
Amplitude of vibration	Cluster 1	20.62	17.00-23.00	1,2,4	3
	Cluster 2	29.97	23.00-35.50	3,5,6,7,8,9	6
G-ratio	Cluster 1	0.0317	0.02-0.06	1,2,3,4,5,6	6
	Cluster 2	0.0693	0.06-0.085	7,8,9	3

Now, after performing all the required data preprocessing and clustering tasks, the decision rule generation algorithm is adopted to explore valuable information from the experimental dataset in the form of developed rules. These rules simply depict the relationships between various grinding parameters and responses to effectively control the said grinding operation. The first three sets of rules relate one or more grinding parameters to a single response. In contrast, the last set of rules relates multiple grinding parameters to all the three responses.

## Rules for Ra:

- Rule 1: If SS = 2430 Then Ra is 0.56 [0.40-0.60].  
[P = 100%, Q = 75%, C = 33.33%, QTY = 3] [T = 208.33]
- Rule 2: If SS = 2560 and DOC = 0.2 Then Ra is 0.56 [0.40-0.60].  
[P = 100%, Q = 25.00%, C = 11.11%, QTY = 1] [T = 136.11]
- Rule 3: If SS = 2850 Then Ra is 0.72 [0.60-0.85].  
[P = 100%, Q = 60.00%, C = 33.33%, QTY = 3] [T = 193.33]
- Rule 4: If SS = 2560 and DOC = 0.3 Then Ra is 0.72 [0.60-0.85].  
[P = 100%, Q = 20.00%, C = 11.11%, QTY = 1] [T = 131.11]
- Rule 5: If SS = 2560 and DOC = 0.4 Then Ra is 0.72 [0.60-0.85].  
[P = 100%, Q = 20.00%, C = 11.11%, QTY = 1] [T = 131.11]

## Rules for amplitude of vibration (V):

- Rule 1: If SS = 2430 and DOC = 0.2 Then V is 20.62 [17.00-23.00].  
[P = 100%, Q = 33.33%, C = 11.11%, QTY = 1] [T = 144.44]
- Rule 2: If SS = 2430 and TCF = Water Then V is 20.62 [17.00-23.00].  
[P = 100%, Q = 33.33%, C = 11.11%, QTY = 1] [T = 144.44]
- Rule 3: If SS = 2560 and DOC = 0.2 Then V is 20.62 [17.00-23.00].  
[P = 100%, Q = 33.33%, C = 11.11%, QTY = 1] [T = 144.44]
- Rule 4: If SS = 2850 Then V is 29.97 [23.00-35.50].  
[P = 100%, Q = 50.00%, C = 33.33%, QTY = 3] [T = 183.33]
- Rule 5: If DOC = 0.4 Then V is 29.97 [23.00-35.50].  
[P = 100%, Q = 50.00%, C = 33.33%, QTY = 3] [T = 183.33]
- Rule 6: If SS = 2560 and DOC = 0.3 Then V is 29.97 [23.00-35.50].  
[P = 100%, Q = 16.67%, C = 11.11%, QTY = 1] [T = 127.78]

## Rule for G-ratio:

- Rule 1: If SS = 2430 Then G-ratio is 0.0317 [0.02-0.06].  
[P = 100%, Q = 50.00%, C = 33.33%, QTY = 3] [T = 183.33]
- Rule 2: If SS = 2560 Then G-ratio is 0.0317 [0.02-0.06].  
[P = 100%, Q = 50.00%, C = 33.33%, QTY = 3] [T = 183.33].
- Rule 3: If SS = 2850 Then G-ratio is 0.0691 [0.06-0.085].  
[P = 100%, Q = 100.00%, C = 33.33%, QTY = 3] [T = 233.33]

## Rules for all the three responses:

- Rule 1: If SS = 2850 Then Ra is 0.72 [0.60-0.85] and V is 29.97 [23.00-35.50] and G-ratio is 0.0693 [0.06-0.085].  
[P = 100.00%, Q = 100.00%, C = 33.33%, QTY = 3] [T = 233.33]
- Rule 2: If SS = 2430 Then Ra is 0.56 [0.40-0.60] and V is 20.62 [17.00-23.00] and G-ratio is 0.0317 [0.02-0.06].  
[P = 66.67%, Q = 66.67%, C = 22.22%, QTY = 2] [T = 155.56]
- Rule 3: If SS = 2560 Then Ra is 0.72 [0.60-0.85] and V is 29.97 [23.00-35.50] and G-ratio is 0.0317 [0.02-0.06].  
[P = 66.67%, Q = 100.00%, C = 22.22%, QTY = 2] [T = 188.89]
- Rule 4: If SS = 2430 and DOC = 0.04 and TCF = Coolant + Water Then Ra is 0.56 [0.40-0.60] and V is 29.97 [23.00-35.50] and G-ratio is 0.0317 [0.02-0.06].  
[P = 100.00%, Q = 100.00%, C = 11.11%, QTY = 1] [T = 211.11]

Rule 5: If SS = 2560 and DOC = 0.02 and TCF = Water Then Ra is 0.56 [0.40-0.60] and V is 20.62 [17.00-23.00] and G-ratio is 0.0317 [0.02-0.06].  
[P = 100.00%, Q = 33.33%, C = 11.11%, QTY = 1] [T = 144.44]

From the developed rules, it can be propounded that for response Ra (a smaller-the-better type of quality characteristic), Rule 1 emerges out as the strongest rule with a T value of 208.33%. Based on this rule, it can be concluded that when the spindle speed is 2430 rpm, all the measured Ra values are expected to be 'Low' ranging between 0.40  $\mu\text{m}$  and 0.60  $\mu\text{m}$  with a rule confidence of P = 100%. Similarly, 75% of all the trials (Q = 75%) having Ra values between 0.40  $\mu\text{m}$  and 0.60  $\mu\text{m}$  have been experimented while setting the corresponding spindle speed at 2430 rpm, and 33.33% of the experimental trials (C = 33.33%) are covered by this rule (i.e. three trials have Ra values between 0.40  $\mu\text{m}$  and 0.60  $\mu\text{m}$ ). Amongst all the nine experimental trials, there are three runs that satisfy this rule (QTY = 3). Similarly, for Rule 3, when the spindle speed is 2850 rpm, the measured Ra values are expected to be 'High' falling within the range of 0.60-0.85  $\mu\text{m}$ . For Ra response, all the remaining rules have less strength with not so much importance in controlling this grinding operation. Rules 4 and 5 showing the combined influences of two separate grinding parameters on Ra appear to be interesting to the production engineers, but they have also low total strength. These rules state that moderate value of spindle speed and moderate/high value of depth of cut lead to higher Ra values causing generation of poor surface finish of the machined components. Spindle speed appears in all the developed rules signifying its maximum importance in this grinding operation, followed by depth of cut. It is quite interesting to notice that type of the cutting fluid does not appear in any of the generated rules, revealing the fact that it has no role in controlling the surface characteristics of the ground workpiece samples.

For amplitude of vibration, six rules are similarly generated. Among them, Rules 4 and 5 are observed to be the most decisive ones with the total strength of 183.33% each. They signify that when spindle speed is 2850 rpm or depth of cut is 0.04 mm, amplitude of vibration is high, falling within the range of 23.00-35.50  $\mu\text{m}$ . Experiment trial number 9 follows both these rules/conditions (i.e. Rules 4 and 5), requiring attention of the concerned production engineers. Some of the developed rules also exhibit the conjoint influences of two grinding parameters on amplitude of vibration, but they have low strength with smaller T values. Among these rules, Rule 2 is supposed to be the interesting one, i.e. it reveals that low spindle speed and water as the cutting fluid lead to reduced amplitude of vibration during the grinding operation. Similarly, low/moderate spindle speed and low depth of cut also cause reduced vibration. For G-ratio, three decision rules are also formulated. Spindle speed only appears in all these rules. It can be thus stated that when the spindle speed is equal to 2560 rpm or less than it, the corresponding values of G-ratio are low, falling between 0.02 and 0.06. In Rule 3, having strength of 233.33%, a spindle speed value of 2850 rpm leads to higher G-ratio, in the range of 0.06-0.085. It is also interestingly observed that depth of cut and cutting fluid type do not affect G-ratio.

When all the three responses are taken into consideration while formulating the corresponding decision rules, they become more complicated. Amongst the five generated rules, Rule 1 has the maximum strength of 233.33%, followed by Rule 4 (211.11%). It states that when the rotational speed of the grinding wheel (spindle speed) is set at its highest operating level of 3 (i.e. 2850 rpm), higher values for all the considered responses

are simultaneously achieved. Higher grinding wheel speed thus leads to poor machined surface with higher Ra values, higher amplitudes of vibration and higher G-ratios. But, Rule 4 with the second maximum strength is supposed to be the most interesting one for the concerned production engineers, because it encompasses all the grinding parameters and responses. Based on this rule, it can be concluded that when the spindle speed is 2430 rpm, depth of cut is 0.04 mm, and a mixture of coolant and water is applied as the cutting fluid, low values of Ra and G-ratio along with high value of amplitude of vibration are observed. Thus, higher rotational speed of the grinding wheel always leads to higher G-ratio (grinding efficiency) with poor surface qualities of the machined components. Similarly, higher depth of cut causes higher vibration during the grinding operation. The application of coolant and water or ordinary water causes enhanced performance of the grinding operation. But, keeping in mind the additional cost of special purpose coolant, it may be advised to apply simple water as the cutting fluid while grinding low alloy steel work materials. Spindle speed plays the most significant role in controlling all the quality characteristics of the considered grinding process, followed by depth of cut and type of the cutting fluid.

## 5. CONCLUSIONS

In this paper, the RST, a machining learning algorithm of data mining, is employed to analyze the experimental data of a grinding process. Based on the generated rules, the effects of three grinding parameters, i.e. spindle speed, depth of cut and type of the cutting fluid on three different responses, i.e. average surface roughness value, amplitude of vibration and grinding ratio are studied. Using the calculated dependency indexes, the possibility of reduction of the initial experimental dataset is also explored. Depending on the type of the responses, they are subsequently grouped into two different clusters, i.e. 'Low' and 'High'. It is observed from the decision rules developed for average surface roughness that low spindle speed leads to better surface roughness of the ground workpiece samples. On the contrary, higher spindle speed or depth of cut causes increased amplitude of vibration. Similarly, higher spindle speed leads to higher grinding ratio (grinding efficiency). The rules formulated while taking into consideration all the three responses demonstrate that at higher rotational speed of the grinding wheel, higher values for all the considered responses are achieved. Type of the cutting fluid does not influence attainment of low surface roughness and higher grinding efficiency; it only affects the vibration generated during the grinding operation. These rules developed based on the application of the RST are easy to comprehend and would guide the concerned production engineers in setting the input parameters of a grinding process so as achieve the desired quality characteristics of the ground components.

It is observed that the classical RST approach can only process discrete data. However, in real time machining applications, most of the measured data are continuous. Hence, for its successful application, there is always an additional task to discretize the continuous response values with the help of a suitable clustering technique. On the other hand, the generalization ability of rough sets needs to be improved and the probability distribution of sample data requires to be further considered for its effective deployment. Moreover, in the RST approach, during data pre-processing, attribute reduction may often lead to over-fitting of a problem.

With automation of manufacturing industries and availability of high speed data acquisition systems, the manufacturing domain is now flooded with huge volumes of experimental data which if mined, can lead to effective and efficient control of different machining processes. The 'If-Then' decision rules generated using the RST approach can be applied to any of the conventional and non-conventional machining processes to visualize the influences of their input parameters on the responses under consideration. As these rules are quite easy to apprehend, even by a non-technical end user, they can lead to manufacturing process control and optimization with the fulfillment of the primary objective of enhanced productivity with better quality of final products.

#### REFERENCES

1. Han, J., Kamber, M., Pei, J., 2001, *Data Mining: Concepts and Techniques*, Morgan Kaufmann Publishers, USA.
2. Dunham, M.H., 2006, *Data Mining: Introductory and Advanced Topics*, Pearson Education India.
3. Kusiak, A., 2006, *Data mining: Manufacturing and service applications*, International Journal of Production Research, 44(18-19), pp. 4175-4191.
4. Wang, K., 2007, *Applying data mining to manufacturing: the nature and implications*, Journal of Intelligent Manufacturing, 18(4), pp. 487-495.
5. Pawlak, Z., 1982, *Rough sets*, International Journal of Computer and Information Sciences, 11(5), pp. 341-356.
6. Guo, J.Y., Chankong, V., 2002, *Rough set-based approach to rule generation and rule induction*, International Journal of General Systems, 31(6), pp. 601-617.
7. Nzaramba, A., Yang, W.J., Langat, G.K., 2018, *Decision rules making based on rough set approach*, International Journal of Scientific & Engineering Research, 9(2), pp. 752-757.
8. Sadoyan, H., Zakarian, A., Mohanty, P., *Data mining algorithm for manufacturing process control*, International Journal of Advanced Manufacturing Technology, 28(3-4), pp. 342-350.
9. Tseng, T.L.B., Kwon, Y., Ertekin, Y.M., 2005, *Feature-based rule induction in machining operation using rough set theory for quality assurance*, Robotics and Computer-Integrated Manufacturing, 21(6), pp. 559-567.
10. Tseng, T.L., Huang, C.C., Jiang, F., Ho, J.C., 2006, *Applying a hybrid data-mining approach to prediction problems: A case of preferred suppliers prediction*, International Journal of Production Research, 44(14), pp. 2935-2954.
11. Vasiljević, M., Fazlollahtabar, H., Stević, Ž., Vesković, S., 2018, *A rough multicriteria approach for evaluation of the supplier criteria in automotive industry*, Decision Making: Applications in Management and Engineering, 1(1), pp. 82-96.
12. Buddhakulsomsiri, J., Siradeghyan, Y., Zakarian, A., Li, X., 2006, *Association rule-generation algorithm for mining automotive warranty data*, International Journal of Production Research, 44(14), pp. 2749-2770.
13. Karavidić, Z., Projović, D., 2018, *A multi-criteria decision-making (MCDM) model in the security forces operations based on rough sets*, Decision Making: Applications in Management and Engineering, 1(1), pp. 97-120.
14. Sharma, H.K., Kumari, K., Kar, S., 2020, *A rough set theory application in forecasting models*, Decision Making: Applications in Management and Engineering, 3(2), pp. 1-21.
15. Chadha, M., Lee, C.W., 2010, *Optimisation of the multi-pass grinding operation using evolution strategy with variable length representation*, International Journal of Manufacturing Research, 5(3), pp. 286-304.
16. Siddiquee, A.N., Khan, Z.A., Mallick, Z., 2010, *Grey relational analysis coupled with principal component analysis for optimisation design of the process parameters in in-feed centreless cylindrical grinding*, International Journal of Advanced Manufacturing Technology, 46(9-12), pp. 983-992.
17. Lee, K.M., Hsu, M.R., Chou, J.H., Guo, C.Y., 2011, *Improved differential evolution approach for optimization of surface grinding process*, Expert Systems with Applications, 38(5), pp. 5680-5686.
18. Asiltürk, İ., Tinki, M., El Monuayr, H., Çelik, L., 2012, *An intelligent system approach for surface roughness and vibrations prediction in cylindrical grinding*, International Journal of Computer Integrated Manufacturing, 25(8), pp. 750-759.
19. Khan, Z.A., Siddiquee, A.N., Kamaruddin, S., 2012, *Optimization of in-feed centreless cylindrical grinding process parameters using grey relational analysis*, Pertanika Journal of Science & Technology, 20(2), pp. 257-268.

20. Neşeli, S., Asiltürk, İ., Çelik, L., 2012, *Determining the optimum process parameter for grinding operations using robust process*, Journal of Mechanical Science and Technology, 26(11), pp. 3587-3595.
21. Rudrapati, R., Pal, P.K., Bandyopadhyay, A., 2012, *Modelling for surface roughness in cylindrical grinding*, International Journal of Machining and Machinability of Materials, 12(1-2), pp. 28-36.
22. Köklü, U., 2013, *Optimisation of machining parameters in interrupted cylindrical grinding using the grey-based Taguchi method*, International Journal of Computer Integrated Manufacturing, 26(8), pp. 696-702.
23. Pai, D., Rao, S., D'Souza, R., 2013, *Application of response surface methodology and enhanced non-dominated sorting genetic algorithm for optimisation of grinding process*, Procedia Engineering, 64, pp. 1199-1208.
24. Pawar, P.J., Rai-Kalal, D.P., 2013, *Multi-objective optimisation of grinding process parameters using NSGA-II*, International Journal of Metaheuristics, 2(2), pp. 123-140.
25. Winter, M., Li, W., Kara, S., Herrmann, C., 2014, *Determining optimal process parameters to increase the eco-efficiency of grinding processes*, Journal of Cleaner Production, 66, pp. 644-654.
26. Khare, S.K., Agarwal, S., 2015, *Predictive modeling of surface roughness in grinding*, Procedia CIRP, 31, pp. 375-380.
27. Aleksandrova, I., 2016, *Optimization of the dressing parameters in cylindrical grinding based on a generalized utility function*, Chinese Journal of Mechanical Engineering, 29(1), pp. 63-73.
28. Deng, Z., Lv, L., Li, S., Wan, L., Liu, W., Yan, C., Zhang, H., 2016, *Study on the model of high efficiency and low carbon for grinding parameters optimization and its application*, Journal of Cleaner Production, 137, pp. 1672-1681.
29. Rudrapati, R., Pal, P.K., Bandyopadhyay, A., 2016, *Modeling and optimization of machining parameters in cylindrical grinding process*, International Journal of Advanced Manufacturing Technology, 82(9-12), pp. 2167-2182.
30. Çaydaş, U., Çelik, M., 2017, *Genetic algorithm-based optimization for surface roughness in cylindrically grinding process using helically grooved wheels*, Surface Review and Letters, 25(2), pp. 1-8.
31. Chang, Z., Jia Q., Yuan, X., Chen, Y., 2017, *Optimization of the grinding process to improve the surface integrity of bearing raceways*, International Journal of Advanced Manufacturing Technology, 91(9-12), pp. 4243-4252.
32. Kuo, C., Hsu, Y., Chung, C., Chen, C.C.A., 2017, *Multiple criteria optimisation in coated abrasive grinding of titanium alloy using minimum quantity lubrication*, International Journal of Machine Tools and Manufacture, 115, pp. 47-59.
33. Ming, X., Gao, Q., Yan, H., Liu, J., Liao, C., 2017, *Mathematical modeling and machining parameter optimization for the surface roughness of face gear grinding*, International Journal of Advanced Manufacturing Technology, 90(9-12), pp. 2453-2460.
34. Liu, G., Li, C., Zhang, Y., Yang, M., Jia, D., Zhang, X., Guo, S., Li, R., Zhai, H., 2018, *Process parameter optimization and experimental evaluation for nanofluid MQL in grinding Ti-6Al-4V based on grey relational analysis*, Materials and Manufacturing Processes, 33(9), pp. 950-963.
35. Kalpakjian, S., Schmid, S., 2014, *Manufacturing Engineering & Technology*, Prentice Hall, NJ, USA.