Scientia Iranica (2024) 31(3), 186–205

SCIENTIA IRANICA

Sharif University of Technology
Scientia Iranica
Transactions E: Industrial Engineering
https://scientiairanica.sharif.edu

Discriminant analysis-based parametric study of an electrical discharge machining process

B. Sarker and S. Chakraborty*

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

Received 25 June 2020; received in revised form 11 September 2020; accepted 15 November 2021

KEYWORDS
Discriminant analysis;
EDM process;
Process parameter;
Response;
Cut-off score.

Abstract. This paper presents the application of discriminant analysis in an Electrical Discharge Machining (EDM) process to determine the comparative contribution of each of its input parameters on the measured responses. It also identifies the most significant EDM process parameters influencing those responses. For this process, voltage, current, pulse-on time, and pulse-off time are treated as the input parameters, whereas, material removal rate, electrode wear rate, and surface roughness are the responses. Based on the past and simulated experimental data, both simultaneous and step-wise estimations are carried out for each of the three responses showing the relationships between the EDM process parameters and the considered responses. It is observed that in both these estimations, pulse-off time, current and pulse-on time respectively evolve out as the most significant parameters for material removal rate, electrode wear rate, and surface roughness. Step-wise estimation identifies voltage as the least significant input parameter for all these responses. The developed discriminant functions, which can also help in predicting the responses, are finally cross-validated.

1. Introduction

Electrical Discharge Machining (EDM) is a popular industrially-accepted non-traditional machining process, used for machining various advanced engineering materials which are difficult to cut by the conventional machining processes. It is particularly suitable for generating complex contours, patterns, and cavities, especially on electrically conductive materials, like tungsten and its alloys, bronze, copper, carbon and stainless steels, inconel, titanium and its alloys, carbon graphite, composites and other selected ceramic materials [1]. In this process, material removal takes place by a series of periodic and controlled electric discharges between two electrodes (the tool and the workpiece), with a very small gap maintained between them [2]. During EDM operation, the workpiece and the tool are submerged in a dielectric liquid (kerosene or deionized water), which acts as an insulator to control the spark discharges.

In the EDM process, on increasing the voltage between the electrodes, the electric field intensity in this region increases, until it exceeds the dielectric strength. At higher voltage, the dielectric breaks down and allows the current to flow between the electrodes in the form of sparks, which erode material from the workpiece. After stopping the current flow, another volume of dielectric liquid is used to flush into the inter-electrode gap, carrying away the solid debris, so that the dielectric property of the medium is restored.

*. *Corresponding author. Tel.: +91-9831568294*
E-mail address: s_chakraborty00@yahoo.co.in (S. Chakraborty)

doi: 10.24200/sci.2021.56262.4628

Thus, EDM is a thermal process where the material is removed by the application of heat. The electrical discharges between the tool and the workpiece act as the source of heat. The points between the tool and the workpiece, where the spark begins and ends respectively, are heated to such an extent that the work material melts and then vaporizes [2].

Due to its unique material removal characteristics, it can generate intricate shape geometries on materials irrespective of their hardness and brittleness properties. Since there is no direct contact between the tool and the workpiece, there is almost no force applied during the machining operation, enabling the use of soft and easy-to-machine electrodes even while machining extremely hard workpiece materials [3]. Another major advantage of the EDM process is that there is almost no mechanical vibration and residual stress generation during the machining operation. It can also machine components with higher dimensional accuracy while maintaining close tolerances. There are mainly three types of EDM processes, i.e. hole-drilling EDM, die-sinking EDM, and wire-EDM. Although the material removal mechanism is almost the same in these EDM processes, they have been employed to machine different components to satisfy the end requirements of varying manufacturing industries. The performance of an EDM process can often be characterized by various responses, like Material Removal Rate (MRR), Tool Wear Rate (TWR), Surface Roughness (SR), Electrode Wear Rate (EWR), etc.

As EDM is a complex and transient micro-physical process, its stochastic material removal mechanism is affected by multiple factors, making it difficult to establish an appropriate model to investigate the relations between the input parameters and responses [4]. Past researchers have already attempted to implement different techniques, like multiple regression analysis [5–8], Response Surface Methodology (RSM) [9–12], Support Vector Machine (SVM) [13,14], Artificial Neural Network (ANN) [15–18], Adaptive Neuro-Fuzzy Interference System (ANFIS) [19,20] etc. to ascertain these relationships between the input and output parameters of the EDM processes. A list of the input parameters, responses, and mathematical techniques adopted by past researchers for parametric analysis of EDM processes is provided in Table 1. It can clearly be noticed from this table that there exists no research work where any of the binary statistical classification techniques has been successfully applied for the modeling of conventional or non-traditional machining processes. Binary classification techniques, like logistic regression, probit model, decision tree, discriminant analysis, etc. are the methods of categorizing objects of a particular set into two groups on the basis of a classification rule, and thus predicting the group in which an object would be positioned. They can determine the relationships between the input and output variables, and identify the most significant input variable for each of the output variables.

As observed from Table 1, different techniques, like multiple regression analysis, RSM, SVM, ANN, ANFIS, etc. have already been deployed for machining performance prediction of EDM processes. But, they all have their own drawbacks. The SVMs cannot specify the score of the observations due to the lack of a linear combination of independent variables or functions, thereby indicating a paucity of transparency in the result. Thus, the contribution of each independent variable to the dependent variable cannot be definitely represented. Similarly, the ANN is a black-box type approach that cannot provide the causal relationship between the independent and dependent variables. Besides suffering from the inability to deal with high-dimensional data, ANFIS also faces a problem similar to SVM and ANN, i.e., lack of ability to establish a relationship between independent and dependent variables in the form of an equation [21]. Thus, the influence of each independent variable, absolute or comparative, on the dependent variable cannot be explained by these techniques due to the lack of an appropriate equation or function. Interpretation of the derived results and classification of further observations may also be difficult to achieve. Furthermore, all these three techniques suffer from overfitting of observations and are complex, requiring higher computational time.

On the other hand, discriminant analysis is a statistical approach that finds its application in comprehending the relation between a non-metric dependent variable and multiple metric independent variables while predicting the category into which an observation can be classified. It has evolved as an effective prediction tool in marketing, finance, social sciences, and other allied areas. It develops a discriminant function, similar to a multivariate regression equation to enable prediction and explanation of the contribution of each independent variable on the behavior of the dependent variables. The discriminant score calculated based on the discriminant function helps in the classification of the observations into relevant groups. While the independent variables can be compared based on their individual contributions to the changing values of the dependent variables during simultaneous estimation, the discriminant analysis also allows step-wise estimation, which can identify only the significant independent variables, while removing the insignificant ones from further analysis.

Although discriminant analysis is analogous to multiple regression analysis, it has an added advantage based on its ability to compute the difference between group means and the influence of the independent variables behind this difference. Since, it attempts to maximize the difference between the group means and

Table 1. List of input parameters, responses, and mathematical tools for parametric study of EDM processes.

Sl. no.	Author(s)	Input parameters	Responses	Tool(s)
1	Debnath et al. [5]	Pulse-on time, pulse-off time, current	MRR, SR, TWR	Multiple regression analysis
2	Singh et al. [6]	Pulse-on time, pulse-off time, induced current	MRR, SR	Multivariate regression analysis
3	Gudipudi et al. [7]	Pulse-on time, pulse-off time, current	MRR, average recast layer thickness	Multiple regression analysis
4	Kishan et al. [8]	Pulse-on time, pulse-off time, discharge current	MRR, TWR	Multivariate regression analysis
5	Kumar et al. [9]	Pulse-on time, pulse-off time, peak current, types of tool and powder material	MRR, TWR	RSM, desirability function
6	Sinha et al. [10]	Current, pulse-on time, voltage	MRR, TWR	RSM
7	Rajneesh et al. [11]	Pulse-on time, pulse-off time, discharge current	EWR, MRR, SR	RSM
8	Soundhar et al. [12]	Pulse-on time, pulse-off time, voltage, current	MRR, EWR, SR	RSM
9	Aich and Banerjee [13]	Pulse-on time, pulse-off time, current	MRR, SR	SVM
10	Jiang et al. [14]	Pulse-on time, pulse-off time, current	MRR, SR	SVM
11	Rajesh and Anand [15]	Discharge current, discharge voltage, pulse-on time, pulse-off time, oil pressure, gap width	MRR, SR	ANN
12	Bharti [16]	Pulse-on time, current, voltage	MRR, SR	ANN

Table 1. List of input parameters, responses, and mathematical tools for parametric study of EDM processes (continued).

Sl. no.	Author(s)	Input parameters	Responses	Tool(s)
13	Ong et al. [17]	Pulse interval, pulse duration, peak current	MRR, EWR	Radial basis function neural network, moth search algorithm
14	Moghaddam and Kolahan [18]	Pulse-on time, pulse-off time, voltage, duty factor, peak current	MRR, TWR, SR	ANN, particle swarm optimization
15	Sethuramalingam and Sundararaj [19]	Pulse-on time, pulse-off time, pulse current	MRR, EWR, SR	ANFIS
16	Fazlollahtabara and Gholizadeh [20]	Pulse-on time, duty cycle, current	MRR, SR, EWR	ANFIS
17	This paper	Pulse-on time, pulse-off time, voltage, current	MRR, EWR, SR	Simultaneous and step-wise discriminant analysis

subsequently, the separation between the groups, it is a more efficient technique to identify the influence of the independent variables and their significance on the dependent variables. On the other hand, the RSM technique attempts to fit the data of a system to a polynomial even if it is not adequately explained by second-order polynomials [22]. When the equation developed using the RSM technique fails to explain the system behavior properly, it becomes necessary to reduce the range of values of the independent variables. Discriminant analysis is a causal model that maximizes the group difference by computing weights associated with the independent variables. Hence, the influence of each independent variable on the dependent variable is mirrored in the developed discriminant function. Apart from this, it is not influenced by the range of the independent variable. It also runs faster with less overfitting of observations.

The separation between the observations, classified into binary groups, is usually accounted by the Mahalanobis distance (D^2), providing an acceptable notion of distance with respect to standard deviation. Its higher value signifies more effectiveness of the discriminant function correlating the independent and dependent variables. On the contrary, multiple regression analysis is based on the calculation of the Euclidean distance using least square estimation. Similarly, while classifying new observations into binary groups, their corresponding D^2 values from the group means are calculated and it is allotted to the group having the minimum D^2 value.

It can be considered similar to principal component analysis based on the aspect of dimensionality reduction. It reduces the problem to a single dimension using the discriminant scores of the observations. These observations are segmented into corresponding groups formed by maximizing the separation between the group centroids while minimizing the scatter or variation across the scale. A two-group discriminant analysis can also be considered similar to multiple linear regression analysis. Both permit analysis of the causal relationship between the independent variables and dependent variables while envisaging the behavior of each of the dependent variables for different values of the independent variables. Both these techniques have the same set of assumptions and have provisions for step-wise estimation. However, discriminant analysis takes into account only non-metric dependent variables, while multiple linear regression deals with metric or continuous variables. Multiple regression is observed to be inadequate if the dependent variables are categorical. However, discriminant analysis can be used to search out the causal relationships even for continuous variables, if they are converted into categorical ones. All the measured values of the considered dependent variables are segmented into high and low categories for successive application of discriminant analysis. This paper thus presents the distinctive application of discriminant analysis in the parametric study of an EDM process to analyze the influences of its various input parameters on the responses and identify the relative importance of those parameters in enhancing the process performance.

The organization of this paper is structured as follows: After providing a brief overview of the EDM process and the need for discriminant analysis for its study in the Introduction section, the mathematical background of discriminant analysis is presented in Section 2. Section 3 exhibits the past experimental data considered for discriminant analysis. The discriminant functions for the EDM process based on simultaneous estimation and step-wise estimation methods are respectively provided in Sections 4 and 5. Discussions and conclusions are respectively furnished in Sections 6 and 7.

2. Discriminant analysis

Discriminant analysis is a multivariate statistical technique deployed for separating distinct sets of objects (observations) and assigning new objects to previously defined groups [23]. It evaluates the connection between dependent variables, which are categorical (non-metric or nominal), and independent variables, which are metric. A discriminant function, which is the linear combination of two or more independent variables, precisely discriminating the objects within a group with *a priori*, is developed in this analysis [24]. The discriminant function can be represented as follows:

$$Z_{qr} = \alpha + \beta_1 X_{1r} + \beta_2 X_{2r} + ... + \beta_n X_{nr}, \qquad (1)$$

where Z_{qr} denotes the score of discriminant function q for object r, α is the intercept, X_{nr} is the independent variable n for object r and β_n represents the discriminant coefficient for independent variable n.

Discriminant analysis fittingly validates the hypothesis of equality of group means of all the independent variables for two or more groups [24]. The group mean is determined by calculating the simple average of the discriminant scores for all the elements within a certain group. This group is also referred to as a centroid, with one group centroid for each group. The group centroid indicates the most representative position of an element in a particular group, while a comparison of group centroids shows the separation between the groups due to discriminant function. It can also predict the group in which a certain observation would fit depending on the proximity of the discriminant score of the observation to the group centroids. The test for statistical significance of the discriminant function is a hypothesized measure of the distance between the group centroids [24]. For this, the discriminant score distributions of the considered groups are contrasted and the function is tested on the basis of overlap between the groups. A small overlap indicates that the discriminant function significantly separates the groups, while a large overlap symbolizes that the groups cannot be properly segmented. The discriminant analysis generates more than one discriminant function if the dependent variables consist of more than two groups. In fact, this analysis produces $(g-1)$ functions, where g denotes the number of groups, with each function calculating a different discriminant score. This paper deals with dependent variables containing two groups, with a combination of independent variables and their relationships with the dependent variables through a single discriminant function, where the responses of the considered EDM process are treated as the dependent variables and input parameters as the independent variables.

The application methodology of discriminant analysis is depicted in the form of a flowchart in Figure 1. The first step involves in identification of the problem statement and objectives of the analysis. Discriminant analysis can act as a profile analysis, where it can provide an objective assessment of the differences between groups on a set of independent variables [24]. The aim of this paper is focused on adopting this binary classification technique in identifying the effects of various EDM process parameters on the responses and also finding out the most significant process parameter influencing each of the outputs. In the next step, the research framework is built. Determination of the input and output variables takes place, followed by classification of the output variables into the corresponding binary categories. If the output variable is metric, it is converted into non-metric data. The sample size is also required to be checked at this stage. Brown and Tinsley [25] suggested that the ratio between the sample size and the number of independent variables (input variables) should be a minimum of 10:1.

Then, the assumptions of discriminant analysis need to be validated. These assumptions are related to normality, multicollinearity, and equality of covariance matrices. The independent variables must be checked for univariate normality because it is the most effective measure of confirming multivariate normality [24]. Multicollinearity indicates that two or more independent variables are highly correlated, and one independent variable in the analysis can be predicted and described by the other independent variables, adding little to the explanatory power of the entire dataset [24]. Hence, the absence of multicollinearity among the independent variables is highly desired. Multicollinearity can be tested by the Variance Inflation Factor (VIF). It measures the degree to which variance for each variable would be higher for multi-

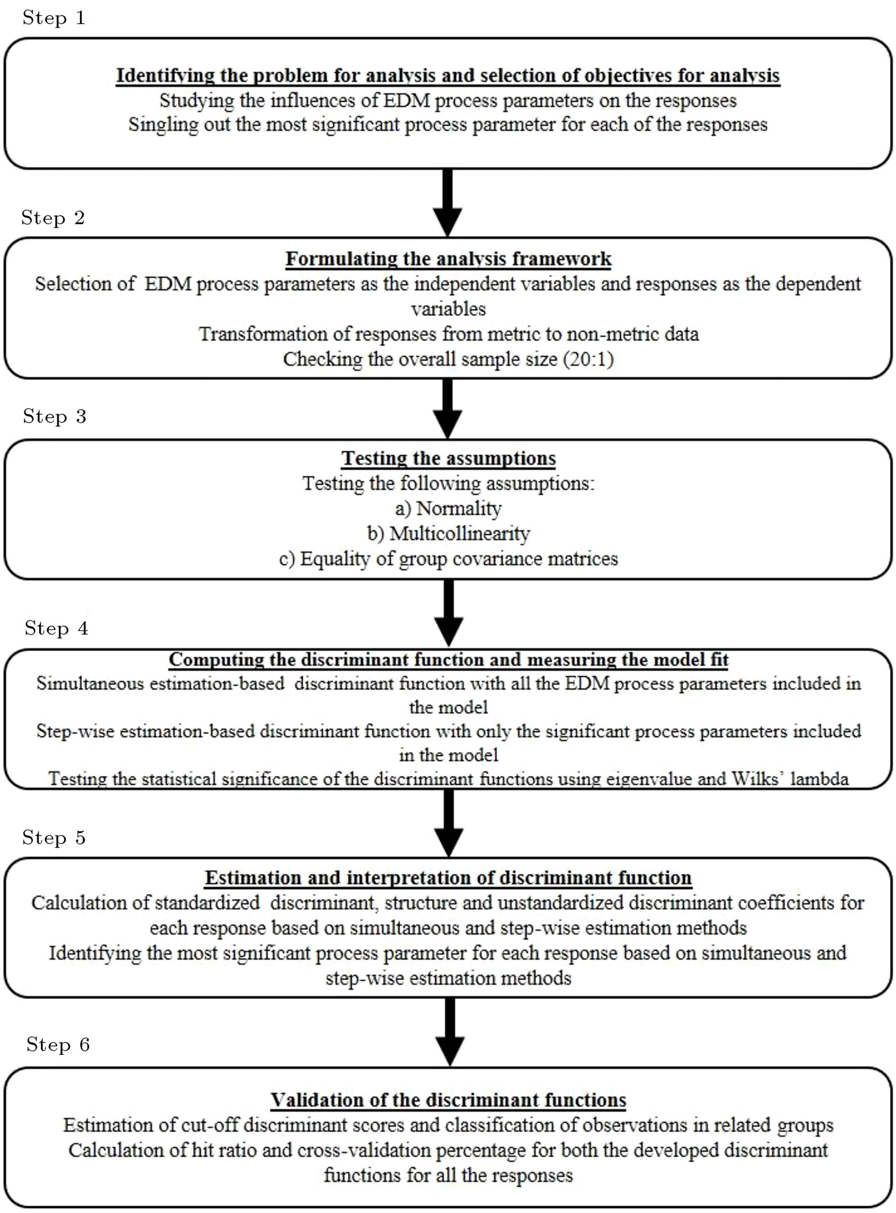

Figure 1. Flowchart showing steps of discriminant analysis.

collinear data than for orthogonal data [26]. Equality of covariance matrices or homoscedasticity signifies whether all the variance-covariance matrices across the groups are equal or not [27]. It is verified by the Box's M test, which assumes the null hypothesis that the within-class covariance matrices are equal; hence, an insignificant result is desired, which cannot reject the null hypothesis. Suppose that a particular dependent variable is classified into a number of groups with n_j observations in each group and the estimated within-group covariance is S_j. The Box's M can then be calculated using the following equation:

$$M = (N-a)\ln|S| - \sum_{j=1}^{a}(n_j-1)\ln|S_j|, \tag{2}$$

where $N = \sum_{j=1}^{a} n_j$ and $S = \frac{\sum_{j=1}^{a}(n_j-1)S_j}{(N-a)}$.

In the next step, the corresponding discriminant functions are developed based on both the simultaneous estimation and step-wise estimation methods. In the simultaneous estimation method, all the variables are involved in the model, irrespective of their ability to discriminate objects between the groups. The discriminant function would thus be a function of all the independent variables considered in the analysis. Step-wise estimation is applied to determine the set of the most significant independent variables, i.e. those variables having the maximum discriminating powers. The discriminant model is constructed in steps, where after every step, the variables, not included in the model, are assessed to identify the one having the highest discriminating power. This results in the selection of the significant independent variables and the process would continue step-wise until all the significant variables are involved in the model. The inclusion and exclusion of the variables in the model are respectively controlled by the 'F to enter' and 'F to remove' values. The F-value for a particular independent variable denotes the statistical significance in discriminating between the groups of dependent variables. The result of both the simultaneous and step-wise estimation methods is the development of the discriminant function, as shown earlier in Eq. (1). The model fit of the developed function can be evaluated here with the help of eigenvalue and Wilks' lambda. The eigenvalue can be explained as the ratio of the between-groups sum of squares to the within-group sum of squares [28]. A higher eigenvalue is always desired for establishing the model fit. The Wilks' lambda is a likelihood ratio statistic for validating the hypothesis that the group means are equal in the population and approach zero if any two groups are well separated [29]. Thus, an insignificant Wilks' lambda value is always preferred. Suppose B is the 'between-groups' matrix and W is the 'within-group' matrix, the Wilks' lambda (Λ) can then be calculated as below [30]:

$$\Lambda = \frac{\det(W)}{\det(W+B)}. \tag{3}$$

The discriminant functions are usually interpreted with the help of standardized coefficients, unstandardized coefficients, and structure matrix. Each independent variable has a standardized coefficient in each of the discriminant functions. The size of the standardized coefficient is proportional to the influence of the respective independent variable on the discrimination power to segregate observations between the corresponding groups. These coefficients help comparison between the independent variables, which are measured in different scales. The standardized coefficients of the independent variables with larger absolute values represent their higher discriminating ability. Another way to understand how each independent variable explains a discriminant function is to analyze the related structure matrix. The structure coefficients, also known as discriminant loadings or structure correlations, are the correlations between the independent variables and discriminant functions. Thus, they can be considered similar to factor loadings in measuring the relative influences of all the independent variables in the discriminant function. One distinct feature of structure correlations is that they can be computed for all the variables, irrespective of their selection in the model during step-wise estimation of the discriminant function. The unstandardized coefficients, calculated for each of the independent variables, are used to form the discriminant function, which in turn, is employed to calculate the corresponding discriminant score. These scores are quite helpful during cross-validation and allocation of the objects, which are not classified into the relevant groups. The prediction of the group to which an object would relate is indicated by its discriminant score and its proximity to the group centroid. Group centroids are basically the group means of the discriminant scores. The group centroids are employed to estimate the cut-off score, which acts as the standard with respect to which the discriminant scores are compared to find out the group in which an object is to be categorized. The cut-off score (Z_C) between the two groups is calculated using the following expressions:

a) For unequal groups:

$$Z_C = \frac{N_A Z_B + N_B Z_A}{N_A + N_B}, \tag{4}$$

where N_A and N_B represent the group sizes, and Z_A and Z_B denote the group centroids respectively.

b) For equal groups:

$$Z_C = \frac{Z_A + Z_B}{2}. \tag{5}$$

Finally, the results derived from this analysis are validated, while measuring the degree of accuracy of the discriminant function in categorizing the objects. This validation can be performed in two ways, i.e. calculating the hit ratio and cross-validation. The discriminant analysis begins with a sample of objects classified into pre-defined groups. So, when the analysis is completed and the discriminant scores are calculated, the objects are re-classified into different groups according to the proximity of their discriminant scores with respect to those of the group centroids. Thus, the hit ratio is the percentage of objects classified correctly by the discriminant function. Besides calculating the hit ratio, cross-validation also needs to be performed. In discriminant analysis, the discriminant coefficients are computed so as to maximize the difference between the groups [31]. While doing this, discriminant analysis takes advantage of the differences among the groups that occur only because of the specific characteristics of the sample. Thus, there is a requirement to generalize the findings for all samples, in order to demonstrate the universality of the analysis and computed function based on cross-validation. In this paper, the leave-one-out approach of cross-validation is adopted [24], where one object from the sample is methodically left out and the analysis is performed on the remaining items. The observation, which was previously excluded, is then categorized into any of the binary groups based on the discriminant score calculated using the discriminant function derived by the analysis. This process is continued until all the objects in the sample are left out and categorized into groups. High values of hit ratio and cross-validation are recommended to validate the results obtained from the discriminant analysis.

3. Discriminant analysis of an EDM process

As mentioned earlier, this paper aims to determine the effects of different input parameters of an EDM process on each of its responses based on discriminant analysis. Using a central composite design plan, Soundhar et al. [12] performed 30 experiments on a die-sinking EDM set-up (Grace D-6030S make). A specially treated titanium alloy (TZN) (Ti-13Zr-13Nb) was chosen as the work material having a 20 mm diameter and 35 mm length. Four EDM process parameters, i.e. voltage, current, pulse-on time, and pulse-off time are considered here as the independent variables, whereas, MRR (in g/min), EWR (in g/min), and SR (in μm) are the dependent variables (responses) in this analysis. In order to achieve higher MRR and lower EWR, a graphite electrode having a 10 mm diameter with negative polarity was used. All the specimens were machined for 20 min and commercial-grade kerosene was utilized as the dielectric fluid. The MRR measures the amount of material removed from the workpiece during unit machining time, whereas, EWR denotes the amount of material removed from the tool electrode during unit machining time. On the other hand, SR characterizes the surface quality of a machined component. It is quantified by the deviations in the direction of normal of a machined surface from its ideal form. A machined surface would be rough if there are large deviations; otherwise, it would be smooth. Separate discriminant analyses are now performed here for all three dependent variables, using both simultaneous estimation and step-wise estimation methods. For the said purpose, IBM SPSS Statistics 25.0 software is employed.

To study the influences of the EDM process parameters on the responses, Soundhar et al. [12] varied the values of the four input variables at three different operating levels, as provided in Table 2. The experimental plan and the measured responses are displayed in Table 3. Soundhar et al. [12] identified the operating levels of the considered EDM process parameters based on several pilot runs and the availability of different settings of the control parameters in the EDM set-up. This experimental data is considered in this paper for the development of the subsequent discriminant functions for the responses.

As all the values of the measured responses in Table 3 are metric, they need to be categorized into two non-metric groups, i.e., high and low, based on their calculated median values. The response values higher than their corresponding medians are considered here as high and indicated by 2, whereas, the response values having less than the medians are treated as low and are denoted by 1. Among the three responses, MRR requires higher values as the machining efficiency/productivity of an EDM process is directly proportional to MRR. On the contrary, lower values for EWR and SR are always preferred. The EWR represents wearing out of the tool electrode during the EDM operation and higher EWR incurs additional machining costs due to frequent tool replacement. The quality of the machined components is appraised using SR values. Now, for carrying out the discriminant analysis, the number of experimental runs in Table 3 is supposed to be not enough. According to Brown and Tinsley [25], the ratio between the sample size and number of independent variables should be a minimum of 10:1. Pituch and Stevens [32] recommended a ratio of 20:1, with a minimum of 20 members in the group having the least number of objects. In order to fulfill this requirement, another 170 experimental runs are simulated to make the total sample size equal to 200. Table 4 shows the number of members in each group for all three responses in this discriminant analysis. These experimental runs are simulated in such a way that all the EDM process parameters and measured responses must lie within their respective minimum and maximum values. As evident from Table 4, groups for all the responses are almost similar in size. The discriminant analysis would be particularly robust if the difference in group sizes is preferably low. Besides this, group sizes also play an important role in calculating the cut-off discriminant score, which would identify the groups for each member according to this analysis.

The assumptions required to be tested here are normality, non-multicollinearity, and homogeneity of covariance matrices. The results of normality and multicollinearity tests are equally applicable to both the simultaneous and step-wise estimation methods, while the results of the test for homogeneity of covariance matrices for simultaneous and step-wise estimations can be different from each other. Hence, the test for homogeneity of covariance matrices is separately carried

Table 2. Input variables and their levels in the EDM process [12].

Input variable	**Symbol**	**Unit**	**Level**		
			−1	**0**	**1**
Voltage	V_o	V	50	60	70
Current	I	A	8	12	16
Pulse-on time	T_{on}	μs	6	8	10
Pulse-off time	T_{off}	μs	7	9	11

V_o: Voltage; I: Current; T_{on}: Pulse-on time; T_{off}: Pulse-off time.

Table 3. Experimental details of the EDM process [12].

Exp. no.	V_o	I	T_{on}	T_{off}	MRR	MRR group	EWR	EWR group	SR	SR group
1	70	8	10	7	0.23	1	0.003	1	11.558	1
2	60	12	8	9	0.473	2	0.007	2	14.717	2
3	60	12	8	9	0.441	2	0.005	1	14.867	2
4	70	8	6	7	0.0789	1	0.004	1	7.647	1
5	50	8	6	11	0.5075	2	0.0004	1	6.245	1
6	50	16	10	7	0.2574	1	0.0115	2	14.514	2
7	60	12	10	9	0.6162	2	0.007	2	13.608	2
8	70	16	6	11	0.086	1	0.004	1	10.168	1
9	60	12	8	9	0.4731	2	0.006	1	10.325	1
10	60	12	8	9	0.4482	2	0.008	2	15.851	2
11	70	8	6	11	0.4272	1	0.0004	1	9.04	1
12	70	16	10	11	0.616	2	0.0117	2	14.514	2
13	60	12	8	9	0.4623	2	0.017	2	16.24	2
14	60	16	8	9	0.5707	2	0.0101	2	11.728	1
15	60	12	8	9	0.4572	2	0.007	2	12.485	2
16	50	16	6	7	0.2193	1	0.008	2	10.008	1
17	60	12	8	7	0.322	1	0.008	2	12.512	2
18	50	8	6	7	0.099	1	0.0042	1	6.301	1
19	70	16	6	7	0.3206	1	0.0105	2	9.577	1
20	50	12	8	9	0.205	1	0.0076	2	12.629	2
21	50	8	10	11	1.051	2	0.0043	1	10.389	1
22	60	12	8	11	0.6305	2	0.0057	1	12.196	1
23	60	12	6	9	0.34	1	0.0044	1	7.545	1
24	50	16	6	11	0.0448	1	0.003	1	6.753	1
25	50	8	10	7	0.2004	1	0.0039	1	13.289	2
26	60	8	8	9	0.7129	2	0.0041	1	9.149	1
27	70	12	8	9	0.2412	1	0.0085	2	18.214	2
28	50	16	10	11	0.525	2	0.0107	2	16.758	2
29	70	16	10	7	0.4086	1	0.0133	2	14.814	2
30	70	8	10	11	1.0208	2	0.0036	1	14.322	2
Median					0.4341		0.0065		12.3405	

V_o: Voltage; I: Current; T_{on}: Pulse-on time; T_{off}: Pulse-off time; MRR: Material Removal Rate; EWR: Electrode Wear Rate; SR: Surface Roughness.

Table 4. Number of members in each group for the discriminant analysis.

Group	Number of members in each group		
	MRR	EWR	SR
1	104	106	104
2	96	94	96

out in this paper for each of the estimation methods. In order to validate the normality assumption for the independent variables, the corresponding skewness and kurtosis values are estimated using Eqs. (6) and (7) respectively. Table 5 provides the results of normality and multicollinearity tests for the considered input (independent) variables.

$$\text{Moment measure of skewness} = \frac{m_3}{\sigma^3} = \frac{m_3}{(\sqrt{m_2})^3}, \quad (6)$$

$$\text{Kursotis} = \frac{m_4}{\sigma^4} - 3, \quad (7)$$

where σ is the standard deviation of the observations, and m_2 and m_3 are the second and third-order central moments of the observations respectively.

Table 5. Tests for normality and multicollinearity.

Input variable	Normality test		Multicollinearity test	
	Skewness	Kurtosis	Tolerance	VIF
V_o	0.105	–1.348	1.000	1.000
I	–0.009	–1.322	1.000	1.000
T_{on}	0.000	–1.365	1.000	1.000
T_{off}	–0.088	–1.353	1.000	1.000

V_o: Voltage; I: Current; T_{on}: Pulse-on time; T_{off}: Pulse-off time;
VIF: Variance Inflation Factor.

Table 6. Assessment of the model fit for simultaneous estimation.

Output variable	Eigenvalue	Canonical correlation	Wilk's lambda	Chi-square	Df	p-value
MRR	0.703	0.643	0.587	104.404	4	< 0.001
EWR	0.986	0.705	0.503	134.519	4	< 0.001
SR	0.741	0.652	0.574	108.661	4	< 0.001

MRR: Material Removal Rate; EWR: Electrode Wear Rate; SR: Surface Roughness; Df: Degree of Freedom.

According to Pituch and Stevens [32], the absolute skewness and kurtosis values must be smaller than 2 to validate the conclusion that the distribution is practically consistent and normal. It can be observed from Table 5, that the absolute values of skewness and kurtosis values are less than the corresponding threshold value of 2. Hence, it can be inferred that the input variables have normal distributions. Tolerance is the degree of variability in one independent variable that cannot be explained by the other independent variables. Its value ranges between 0 and 1, where 1 indicates that an independent variable cannot be explained at all by the other independent variables. The VIF is the reciprocal of tolerance. In Table 5, values of tolerance and VIF are both 1 for all the independent variables, which indicates that these variables are orthogonal in nature with no multicollinearity [33].

4. Simultaneous estimation of the discriminant function

As mentioned earlier, in this estimation method, all the independent variables are involved in the model and the corresponding discriminant function is then developed. At first, the equality of covariance matrices based on the Box's M test needs to be checked. The Box's M values for MRR, EWR, and SR responses are estimated as 102.14, 62.625, and 45.565 respectively. However, the p-values are smaller than 0.001, rendering them significant and thereby rejecting the null hypothesis that the within-group covariance matrices are equal for the discriminant analyses for the three responses. It thus violates one of the basic assumptions as mentioned in the earlier section. However, violation of the assumption of equality of covariance matrices bears less significance during a discriminant analysis and the discriminant analysis may also be robust in spite of this violation [31]. Table 6 depicts the assessment of the model fit based on Wilks' lambda. The Wilks' lambda is a degree of ability of the discriminant function in separating objects into a given number of groups. A smaller value of the Wilks' lambda indicates a higher discriminating power of the developed function. On the other hand, a lower p-value ($p < 0.05$) reiterates the same conclusion. As the discriminant analyses for the three responses show very low p-values, it can be inferred that the developed functions would perform well in separating objects into the corresponding binary groups. Tables 7–9 cumulatively help in investigating the effects of voltage, current, pulse-on time, and pulse-off time on the EDM responses, i.e., MRR, EWR, and SR.

Table 7. Group centroids for simultaneous estimation.

Group	Group centroids		
	MRR	EWR	SR
1 (Low)	–0.802	–0.931	–0.823
2 (High)	0.869	1.049	0.891

MRR: Material Removal Rate; EWR: Electrode Wear Rate; SR: Surface Roughness.

4.1. Discriminant analysis for MRR

Table 7 exhibits that for MRR, the group with higher values of MRR (more than 0.4341 g/min) has a positive centroid. As a result, it can be propounded that the input variables with positive standardized discriminant coefficients would attract the discriminant score of observation towards the group with higher values of MRR (group 2). Similarly, the negative coefficients would attempt to attract the discriminant score of observation towards the group with lower MRR values

Table 8. Standardized discriminant function and structure coefficients for simultaneous estimation.

Input variable	MRR		EWR		SR	
	Standardized discriminant function coefficient	Structure coefficient	Standardized discriminant mfunction coefficient	Structure coefficient	Standardized discriminant function coefficient	Structure coefficient
V_o	–0.082	–0.096	0.073	0.06	0.066	0.041
I	–0.260	–0.18	0.955	0.69	0.383	0.191
T_{on}	0.699	0.458	0.557	0.282	0.98	0.876
T_{off}	0.900	0.695	–0.557	–0.322	–0.309	–0.209

V_o: Voltage; I: Current; T_{on}: Pulse-on time; T_{off}: Pulse-off time; MRR: Material Removal Rate; EWR: Electrode Wear Rate; SR: Surface Roughness.

Table 9. Unstandardized discriminant function coefficients for simultaneous estimation.

Input variable	Unstandardized discriminant function coefficient		
	MRR	EWR	SR
V_o	–0.011	0.009	0.008
I	–0.085	0.373	0.125
T_{on}	0.477	0.368	0.782
T_{off}	0.665	–0.374	–0.2
Constant	–8.217	–4.593	–6.443

V_o: Voltage; I: Current, T_{on}: Pulse-on time; T_{off}: Pulse-off time; MRR: Material Removal Rate; EWR: Electrode Wear Rate; SR: Surface Roughness.

(group 1). As evident from Table 8, pulse-on time and pulse-off time have positive coefficients, and would thus increase the discriminant score towards the centroid of group 2. As a result, it can be concluded that with an increase in the values of pulse-on time and pulse-off time, MRR would increase. Conversely, with increasing values of voltage and current, MRR would tend to decrease. The absolute values of these coefficients also indicate the strengths of the effect of the independent variables on the discriminating power of the function, which can be deployed to infer their comparative effects on the output variable. Here, MRR depends mostly on pulse-off time, followed by pulse-on time, which is in close agreement with the findings of Soundhar et al. [12]. Based on Analysis of Variance (ANOVA) results, Soundhar et al. [12] determined the contributions of pulse-off time and pulse-on time on MRR as 25.01% and 24.49% respectively. The structure coefficients (structure correlations) which denote correlations between the input variables and discriminant variable (MRR) are respectively computed as −0.096, −0.18, 0.458, and 0.695 for voltage, current, pulse-on time, and pulse-off time.

In the EDM process, pulse-on time is the duration when the electrical discharges take place between the tool and the workpiece after the breakdown voltage of the dielectric is achieved. So, more sparking would result in more material removal in less machining time. An increase in pulse-off time also increases MRR. According to Singh et al. [34], if there is insufficient time available for cooling and removal of debris because of shorter pulse-off time, the dielectric liquid gets inadequate time to deionize at the beginning of the next cycle, causing the next electrical discharges to be unstable and thus slowing down the rate at which material is removed from the workpiece. During these experiments, current causes expansion of the plasma channel. A decrease in MRR at a higher current may be due to plasma column contamination resulting from fragmentation of the electrodes (tool and workpiece) [35]. Thus, an increase in current leads to decrement in MRR value. With increasing values of voltage, the gap distance for the initiation of a discharge increases [36]. This leads to an increase in the path of travel for the spark, causing reduced intensity of the spark, thereby decreasing the amount of material removed. Table 9 shows the unstandardized discriminant function coefficients, based on which the following discriminant function for MRR is developed.

$$Z_{MRR} = -8.217 - 0.011V_o - 0.085I + 0.477T_{on} + 0.665T_{off}. \tag{8}$$

Now, using Eq. (8), the corresponding cut-off score is computed as 0.0669. It symbolizes that the observations having discriminant scores (computed using Eq. (8)) higher than 0.0669, would be categorized into group 2 (MRR more than 0.4341 g/min). Similarly, observations with scores smaller than the cut-off value would be classified into group 1, having MRR values smaller than 0.4341 g/min.

4.2. Discriminant analysis for EWR

Table 7 already shows how different input EDM parameters affect EWR. The group with higher values of EWR has also a positive centroid. As observed from Table 8, voltage, current, and pulse-on time have positive coefficients, while pulse-off time has a negative coefficient for EWR. As a result, it can be realized that EWR would increase with increasing values of voltage, current, and pulse-on time, while it would decrease at higher values of pulse-off time. The EWR mostly depends on current, followed by pulse-on time and voltage, which is also in close congruence with the observations of Soundhar et al. [12]. The ANOVA results indicated that the contribution of current on EWR was 40.14%. The correlations between the EDM process parameters and discriminant variable (EWR) are respectively calculated as 0.06, 0.69, 0.282, and −0.322 for voltage, current, pulse-on time, and pulse-off time.

An increase in pulse-on time increases EWR due to higher spark intensity during the EDM operation. Thus, the discharge energy during longer pulse-on time dissipates more heat which would cause the tool to wear out more due to excess heat [37]. Voltage and current also show similar effects on EWR, as both of them are responsible for increased discharge energy, resulting in intense heating of the tool, wearing it out at a higher rate. Increased pulse-off time permits heat dissipation during sparking, while flushing away the debris from the machining zone, thus decreasing the amount of heat generated [37]. Hence, the tool wears out less, when pulse-off time increases. Based on Table 9, the following discriminant function for EWR is now developed.

$$Z_{EWR} = -4.593 + 0.009V_o + 0.373I + 0.368T_{on} -0.374T_{off}. \tag{9}$$

The corresponding cut-off score is estimated as 0.1184. It symbolizes that the observations with discriminant scores (calculated applying Eq. (9)) higher than 0.1184, would be classified into group 2 (EWR more than 0.0065 g/min). On the other hand, observations with discriminant scores less than 0.1184, would be allocated to group 1 (EWR less than 0.0065 g/min).

4.3. Discriminant analysis for SR

As evident from Table 7, the group with higher SR values has a positive centroid. From Table 8, it can be observed that voltage, current, and pulse-on time have positive coefficients, while pulse-off time has a negative coefficient. As a result, it can be anticipated that SR would increase with an increase in the values of voltage, current, and pulse-on time, while it would decrease at higher values of pulse-off time. The absolute values of these coefficients also indicate that SR mostly depends on pulse-on time, followed by current. It closely matches with the findings of Soundhar et al. [12]. The ANOVA-based results determined a contribution of 45.92% for pulse-on time on SR. The structure coefficients for SR response are respectively obtained as 0.041, 0.191, 0.876, and −0.209 for voltage, current, pulse-on time, and pulse-off time.

When pulse-on time increases, the surface machined by the EDM operation may have deep overlying craters, formed due to a series of sparks, extreme heat, melting, and vaporization of the workpiece material at discrete positions. The molten material, left over after flushing by the dielectric liquid, undergoes solidification to form lumps of debris, thus deteriorating the surface quality [38]. An increase in pulse-off time results in a decrease in SR. Longer pulse-off time enables the wearing away of the workpiece material, simultaneously providing a good cooling effect and time to rinse out the debris from the machining zone. On increasing current, discharge energy increases, leading to more erosion of the workpiece and subsequently, an increase in SR [38]. Higher voltage is also responsible for increasing discharge energy, resulting in poor machined surface quality due to increased erosion. Using the information from Table 9, the following discriminant function for SR is derived:

$$Z_{SR} = -6.443 + 0.008V_o + 0.125I + 0.782T_{on} -0.2T_{off}. \tag{10}$$

The related cut-off score is calculated as 0.0683. It indicates that the observations with discriminant scores more than 0.0683 would be assigned to group 2 (SR values more than 12.3405 μm). On the contrary, observations with discriminant scores smaller than 0.0683 would be allocated to group 1 (SR values less than 12.3405 μm).

4.4. Validation of the discriminant analysis

Finally, it is necessary to validate the discriminant analysis results in order to justify whether the developed discriminant functions can be employed as effective classification and prediction tools for the said EDM process. Table 10 provides the original and cross-validation results for the discriminant functions developed based on the simultaneous estimation method. It can be revealed from Table 4 that for MRR response, among 200 simulated experimental observations, 104 have low MRR values (less than 0.4341 g/min) and the remaining 96 have high MRR values (more than 0.4341 g/min). In Table 10, the discriminant function developed for MRR can correctly identify 80 group 1 observations (out of 104) and 58 group 2 observations (out of 96). So, the percentages of correct classification are 76.9% and 60.4% respectively. Thus, the hit ratio for the discriminant function for MRR is 69% (138 out of 200), with a misclassification error of 31%. It has already been mentioned that the prediction performance of the discriminant function is cross-validated using the leave-one-out approach based on

Table 10. Classification results for the simultaneous estimation method.

Output variable	Type of validation	Count	Group	Predicted group membership		Total
				1	2	
MRR	Original	Count	1	80	24	104
			2	38	58	96
		%	1	76.9	23.1	100
			2	39.6	60.4	100
	Cross-validated	Count	1	72	32	104
			2	38	58	96
		%	1	69.2	30.8	100
			2	39.6	60.4	100
EWR	Original	Count	1	95	11	106
			2	30	64	94
		%	1	89.6	10.4	100
			2	31.9	68.1	100
	Cross-validated	Count	1	95	11	106
			2	30	64	94
		%	1	89.6	10.4	100
			2	31.9	68.1	100
SR	Original	Count	1	85	19	104
			2	35	61	96
		%	1	81.7	18.3	100
			2	36.5	63.5	100
	Cross-validated	Count	1	85	19	104
			2	35	61	96
		%	1	81.7	18.3	100
			2	36.5	63.5	100

MRR: Material Removal Rate; EWR: Electrode Wear Rate; SR: Surface Roughness.

IBM SPSS Statistics 25.0 software. For MRR, the percentages of correct classification for group 1 and group 2 objects based on cross-validation are 69.2% and 60.4% respectively. Hence, the overall cross-validation percentage is 65% (130 out of 200). Similarly, for EWR, the hit ratio and cross-validation percentage are both 79.5%. For SR, these values are also the same as 73%. All these higher values indicate that the discriminant functions developed based on the simultaneous estimation method have the ability to classify the response values in appropriate lower and higher groups.

5. Step-wise estimation of the discriminant function

The step-wise estimation of the discriminant function is generally useful in selecting the most significant independent variables, which should be included in the developed model for further analysis. The Wilks' lambda values are considered to select the significant input variables. The variable with the lowest Wilks' lambda which reduces the overall Wilks' lambda maximally, has the first preference to be included in the model. This method starts with the model having no input variable. In every step, the input variable whose 'F to enter' value is the highest and exceeds the entry criterion, is added to the model. On the contrary, the 'F to remove' value is employed to eliminate a specific variable from the model. The 'F to enter' and 'F to remove' values, which are set at 3.84 and 2.71 as default in the software, correspond to p-values of 0.05 and 0.1 respectively. The process continues until all the variables, that meet the entry criterion, are included in the model. But before this analysis starts, the earlier assumptions need to be validated. Assumptions of normality and multicollinearity are already tested in Table 5, which also hold true for this analysis. The Box's M test to validate the assumption of equality of covariance matrices is again performed here for the

Table 11. Assessment of model fit for step-wise estimation.

Output variable	Eigenvalue	Canonical correlation	Wilk's lambda	Chi-square	Df	p-value
MRR	0.699	0.641	0.589	104.120	3	< 0.001
EWR	0.981	0.704	0.505	134.340	3	< 0.001
SR	0.738	0.652	0.575	108.574	3	< 0.001

MRR: Material Removal Rate; EWR: Electrode Wear Rate; SR: Surface Roughness; Df: Degree of Freedom.

Table 12. Variables included/not included in the model for MRR.

Variable included				Variable not included				
Input variable	Tolerance	F-value	Wilks' lambda	Input variable	Tolerance	Minimum tolerance	F-value	Wilks' lambda
T_{off}	0.922	88.353	0.854					
T_{on}	0.929	45.230	0.725	V_o	0.999	0.922	0.546	0.587
I	0.992	5.643	0.606					

V_o: Voltage; I: Current; T_{on}: Pulse-on time; T_{off}: Pulse-off time.

Table 13. Variables included/not included in the model for EWR.

Variable included				Variable not included				
Input variable	Tolerance	F-value	Wilks' lambda	Input variable	Tolerance	Minimum tolerance	F-value	Wilks' lambda
I	0.881	130.121	0.840					
T_{off}	0.939	33.732	0.592	V_o	0.997	0.881	0.519	0.503
T_{on}	0.923	32.212	0.588					

V_o: Voltage; I: Current; T_{on}: Pulse-on time; T_{off}: Pulse-off time.

Table 14. Variable included/not included in the model for SR.

Variable included				Variable not included				
Input variable	Tolerance	F-value	Wilks' lambda	Input variable	Tolerance	Minimum tolerance	F-value	Wilks' lambda
T_{on}	0.955	124.909	0.942					
I	0.963	12.601	0.612	V_o	0.995	0.953	0.361	0.574
T_{off}	0.989	8.406	0.600					

V_o: Voltage; I: Current; T_{on}: Pulse-on time; T_{off}: Pulse-off time.

three responses. The values of the Box's M for MRR, EWR, and SR are obtained as 72.170, 61.396, and 37.877 respectively, with p-values of less than 0.001 for all the analyses. Even though the assumption of the equality of covariance matrices is violated here, it can be assured that the discriminant analysis would be robust enough in spite of this violation [31]. The model fits need to be assessed using the Wilks' lambda values which test the significance of the discriminant functions for the three responses. Table 11 provides the eigenvalues and Wilks' lambda values for all the dependent variables, which test the significance of the discriminant functions. As all the Wilks' lambda and p-values are observed to be low, it can be propounded that the developed discriminant functions would perform satisfactorily as effective evaluation and prediction tools. Tables 12–14 show the variables entered into the models and removed from the models during the development of the step-wise discriminant functions for the three responses.

From Table 12, it becomes evident that the variables to be included in the analysis for MRR are pulse-off time, pulse-on time, and current. The input variables that influence EWR most are current, pulse-off, and pulse-on time. On the other hand, pulse-on time, current, and pulse-off time are the most important input variables for SR. Based on the calculated F-values, it can be unveiled that pulse-off time has

Table 15. Group centroids for step-wise estimation.

Group	Group centroid		
	MRR	EWR	SR
1 (Low)	–0.799	–0.928	–0.821
2 (High)	0.866	1.047	0.889

MRR: Material Removal Rate; EWR: Electrode Wear Rate; SR: Surface Roughness.

the most significant discriminating power in the case of MRR, followed by pulse-on time and current. For EWR, the most significant input variable is current, while SR is most significantly influenced by pulse-on time. All these findings are in close agreement with the observations of Soundhar et al. [12].

The voltage is identified as the least significant contributor in all three discriminant analyses. In discriminant analysis, an independent variable can separate objects only when there exists a significant difference between the group means of the independent variables. The difference in group means of voltage is not significant enough to create discrimination and hence, the changes in the values of MRR, EWR, and SR due to voltage seem to be insignificant. Like the simultaneous estimation method, as discussed earlier, Tables 15–17 highlight the effects of the input variables on the responses.

5.1. Discriminant analysis for MRR

Table 15 shows that for MRR, the centroid of group 2 with higher values is positive, while the centroid of group 1 with lower values is negative. Thus, pulse-on time and pulse-off time have positive effects on MRR, while MRR decreases with increased values of current. The structure coefficients, denoting correlations between the input variables and discriminant variable (MRR) are respectively estimated as −0.014, −0.180, 0.460, and 0.697 for voltage, current, pulse-on time, and pulse-off time. Both the standardized coefficient and structure coefficient indicate that pulse-off time has the maximum discriminating power on MRR, and it is the most significant input variable influencing MRR. Now, based on the unstandardized discriminant function coefficients of Table 17, the following discriminant function is developed for MRR.

Table 17. Unstandardized discriminant function coefficients for step-wise estimation.

Input variable	Unstandardized discriminant function coefficient		
	MRR	EWR	SR
I	–0.085	0.374	0.126
T_{on}	0.478	0.368	0.782
T_{off}	0.669	–0.376	–0.203
Constant	–8.881	–4.008	–5.918

V_o: Voltage; I: Current, T_{on}: Pulse-on time; T_{off}: Pulse-off time; MRR: Material Removal Rate; EWR: Electrode Wear Rate; SR: Surface Roughness.

$$Z_{MRRS} = -8.881 - 0.085I + 0.478T_{on} + 0.669T_{off}. \quad (11)$$

The cut-off score is estimated as 0.0668. It indicates that the observations with discriminant scores higher than 0.0668 would be categorized into group 2 (MRR more than 0.4341 g/min). Similarly, the observations with discriminant scores lower than 0.0668 would be categorized into group 1 (MRR less than 0.4341 g/min).

5.2. Discriminant analysis for EWR

Based on Table 16, it can be noticed that current and pulse-on time have positive effects on EWR, while longer pulse-off time decreases EWR. For this response, the structure coefficients are respectively determined as −0.013, 0.692, 0.282, and −0.323 for voltage, current, pulse-on time, and pulse-off time. Both the standardized coefficient and structure coefficient indicate that current is the most significant input parameter influencing EWR. The corresponding discriminant function

Table 16. Standardized discriminant function and structure coefficients for step-wise estimation.

Input variable	MRR		EWR		SR	
	Standardized discriminant function coefficient	Structure coefficient	Standardized discriminant function coefficient	Structure coefficient	Standardized discriminant function coefficient	Structure coefficient
V_o	–	–0.014	–	–0.013	–	–0.025
I	–0.262	–0.180	0.956	0.692	0.384	0.192
T_{on}	0.700	0.460	0.556	0.282	0.980	0.878
T_{off}	0.905	0.697	–0.562	–0.323	–0.313	–0.210

V_o: Voltage; I: Current, T_{on}: Pulse-on time; T_{off}: Pulse-off time; MRR: Material Removal Rate; EWR: Electrode Wear Rate; SR: Surface Roughness.

Table 18. Classification results for step-wise estimation.

Output variable	Type of validation	Count	Group	Predicted group membership		Total
				1	2	
MRR	Original	Count	1	98	6	104
			2	38	58	96
		%	1	94.2	5.8	100
			2	39.6	60.4	100
	Cross-validated	Count	1	98	6	104
			2	38	58	96
		%	1	94.2	5.8	100
			2	39.6	60.4	100
EWR	Original	Count	1	106	0	106
			2	38	56	94
		%	1	100	0	100
			2	40.4	59.6	100
	Cross-validated	Count	1	77	29	106
			2	38	56	94
		%	1	72.6	27.4	100
			2	40.4	59.6	100
SR	Original	Count	1	85	19	104
			2	43	53	96
		%	1	81.7	18.3	100
			2	44.8	55.2	100
	Cross-validated	Count	1	85	19	104
			2	43	53	96
		%	1	81.7	18.3	100
			2	44.8	55.2	100

MRR: Material Removal Rate; EWR: Electrode Wear Rate; SR: Surface Roughness.

for EWR is developed as below:

$$Z_{EWRS} = -4.008 + 0.374I + 0.368T_{on} - 0.376T_{off}. \quad (12)$$

The observations whose discriminant scores are higher than the cut-off score (0.1187) would be assigned to group 2 (EWR more than 0.0065 g/min) and those with scores less than the cut-off score would be allotted to group 1 (EWR less than 0.0065 g/min).

5.3. Discriminant analysis for SR

It can be revealed from Table 16 that on increasing current and pulse-on time, SR increases; while it decreases with longer pulse-off time. The values of the standardized coefficient and structure coefficient highlight that pulse-on time has the maximum discriminating power on SR, followed by current. The computed discriminant function for SR is shown below:

$$Z_{SRS} = -5.918 + 0.126I + 0.782T_{on} - 0.203T_{off}. \quad (13)$$

The corresponding cut-off score is estimated as 0.0683. The observations with discriminant scores higher than 0.0683 would be assigned to group 2 (SR more than 12.3405 μm). Similarly, the observations with scores less than 0.0683 would be allotted to group 1 (SR less than 12.3405 μm).

5.4. Validation of the discriminant analysis

In Table 18, numbers of the correctly classified items in each group and hit ratios for each group for all three output variables are provided along with the cross-validation results. For MRR, the hit ratio and cross-validation percentage are both 78%. In the case of EWR, the hit ratio is 81%, while the cross-validation percentage is 66.5%. The hit ratio and cross-validation percentage for SR are both 69%.

6. Discussions

Based on the developed discriminant functions, it is observed that voltage is an insignificant input parameter for MRR, TWR, and SR responses of the considered EDM process. The MRR is increased with

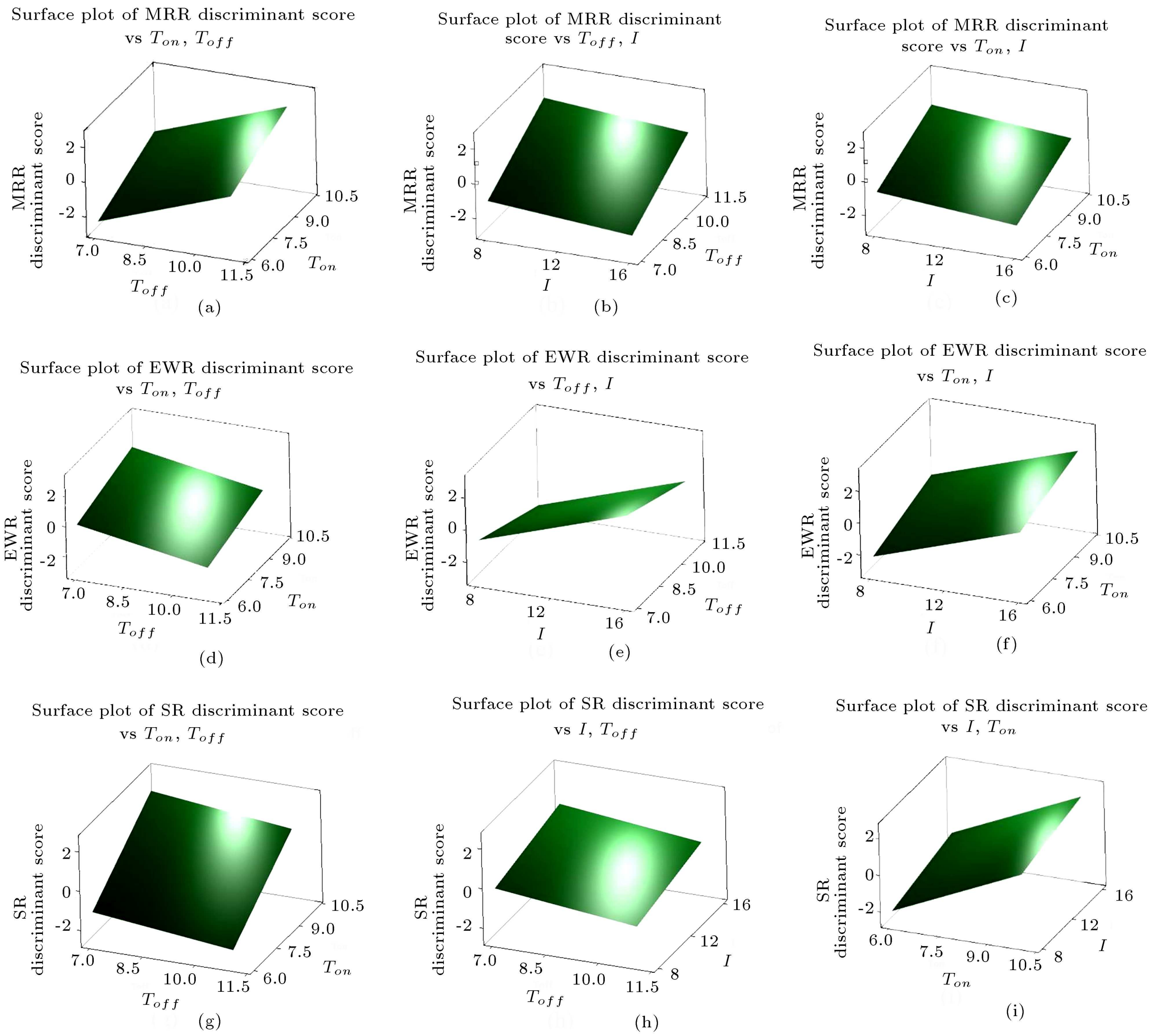

Figure 2. Effects of EDM process parameters on discriminant scores of the responses.

the increasing values of both pulse-on time and pulse-off time, and it tends to decrease with a higher current. Similarly, higher values of pulse-off time and lower values of pulse-on time and current would contribute to lower TWR. Excellent quality of the machined surface with lower SR can only be attained at higher pulse-off time, and lower pulse-on time and current. These observations would tremendously help the concerned process engineer to fix the settings of various input (control) parameters of the EDM set-up. Voltage can be set at any value. In order to simultaneously optimize all the responses, pulse-off time is to be maintained at its higher setting, while pulse-on time and current should be set at moderate and lower settings respectively. Thus, it would lead to multi-objective optimization of the said EDM process. The effects of various input parameters of the considered EDM process on the calculated discriminant scores of the three responses are also exhibited through the surface plots in Figure 2(a)–(i). They show the same trends of influences of the EDM process parameters on the responses as observed from the unstandardized discriminant function coefficients in the developed discriminant functions.

7. Conclusions

This paper presents the application of discriminant analysis in an Electrical Discharge Machining (EDM) process to envisage the effects of its four input parameters on three responses and identify the most significant input parameter for each of the responses. After verifying the corresponding assumptions of normality, multicollinearity, and homogeneity of covariance matrices, the respective simultaneous and step-wise estimation-based discriminant functions are developed. In both these methods, pulse-off time, current, and pulse-on time are respectively observed to be the most significant input variables for Material Removal Rate (MRR), Electrode Wear Rate (EWR), and Surface Roughness (SR). It can thus be realized that lower values of voltage and current, and higher values of pulse-on time and pulse-off time would lead to an increase in MRR. For a better surface finish, higher pulse-off time, and lower values of voltage, current and pulse-on time are essential. Similarly, EWR can be reduced by increasing pulse-off time, and decreasing voltage, current and pulse-on time. The discriminant analysis for all the responses reveals that the absolute values of the standardized discriminant function coefficient and structure coefficient for voltage are always less than 0.10. It can also be validated from the step-wise estimation method that voltage does not significantly affect MRR, EWR, and SR and hence, it is not included in the discriminant functions during step-wise analysis. Higher hit ratio and cross-validation percentages, computed for both simultaneous and step-wise estimation methods, indicate the capability of the developed functions in discriminating experimental observations of the EDM process into well-defined groups.

The step-wise estimation method finds its application when a relatively large number of independent variables are considered while constructing the function. As mentioned earlier, the input variables, which do not significantly discriminate the observations between the groups, are not included in the final discriminant function. The reduced function, without the insignificant input variables, is proved to be as good as the one with all the variables included in the model. The Wilks' lambda values for simultaneous discriminant analysis for the dependent variables are marginally less than those computed during step-wise estimation. A smaller value of the Wilks' lambda is an indication of the higher ability of the function to discriminate observations between the groups. Hence, although simultaneous estimation is a more effective method, the discriminant functions developed based on step-wise estimation are also capable of discriminating experimental observations between the groups. Due to several added advantages of discriminant analysis over the other existing statistical tools, it can be successfully applied as an effective statistical tool for multivariate analysis of different machining processes to identify the most significant input parameters affecting the outputs.

References

1. Rajesha, S., Sharma, A.K., and Kumar, P. "On electro discharge machining of Inconel 718 with hollow tool", *Journal of Materials Engineering and Performance*, **21**(6), pp. 882–889 (2012).
2. Jameson, E.C., *Electrical Discharge Machining*, Society of Manufacturing Engineers, Dearborn, Michigan (2001).
3. Schubert, A., Zeidler, H., Hackert-Oschätzchen, M., et al. "Enhancing micro-EDM using ultrasonic vibrations and approaches for machining of nonconducting ceramics", *Strojniški Vestnik - Journal of Mechanical Engineering*, **53**(3), pp. 156–164 (2012).
4. Lin, Z., Liu, Y., and Zhang, L. "Research of EDM (electrical discharge machining) process simulation based on grey neural network", In: D. Jin and S. Lin, (Eds.) Advances in Mechanical and Electronic Engineering, Springer, Heidelberg, **177**, pp. 373–379 (2012).
5. Debnath, S., Rai, R.N., and Sastry, G.R.K. "A study of multiple regression analysis on die sinking EDM machining of ex-situ developed Al-4.5Cu-Sic composite", *Materials Today: Proceedings*, **5**(2), pp. 5195–5201 (2018).
6. Singh, R., Debnath, R., and Rai, R.N. "Study the effect of machining process parameters on die-sinking EDM for Al5083/B4C composite", *Journal of Scientific and Industrial Research*, **78**, pp. 615–619 (2019).
7. Gudipudi, S., Nagamuthu, S., Subbian, K.S., et al. "Fabrication and experimental study to optimize the recast layer and the material removal in electric discharge machining (EDM) of AA6061-B4C composite", *Materials Today: Proceedings*, **19**(2), pp. 448–454 (2019).
8. Kishan, B., Sudheer Premkumar, B., Gajanana, S., et al. "Development of regression model for Al6061 + SiC tool and optimization of process parameters on electro discharge machining", *IOP Conference Series: Journal of Physics*, **1455**, pp. 1–10 (2020).
9. Kumar, S., Dhingra A.K., and Kumar, S. "Parametric optimization of powder mixed electrical discharge machining for nickel based superalloy inconel-800 using response surface methodology", *Mechanics of Advanced Materials and Modern Processes*, **3**(7), pp. 1–17 (2017).
10. Sinha, S., Ballav, R., and Kumar, A. "Investigation of material removal rate and tool wear rate on electrical discharge machining of Incoloy 800HT by using response surface methodology", *Materials Today: Proceedings*, **4**, pp. 10603–10606 (2017).
11. Rajneesh, R., Subhash, S., Mulik, R.S., et al. "Study of machining performance in EDM through response surface methodology", In: K. Shanker, R. Shankar and R. Sindhwani, (Eds.) Advances in Industrial and Production Engineering, Springer, pp. 207–216 (2019).
12. Soundhar, A., Zubar, H.A., Sultan, M.T.B.H.H., et al. "Dataset on optimization of EDM machining parameters by using central composite design", *Data in Brief*, **23**, Article 103671, pp. 1–11 (2019).

13. Aich, U. and Banerjee, S. "Modeling of EDM responses by support vector machine regression with parameters selected by particle swarm optimization", *Applied Mathematical Modelling*, **38**(11–12), pp. 2800–2818 (2014).

14. Jiang, H., Lin, Y., Wang, S., et al. "Research on optimum processing parameter of EDM based on support vector machine", *Applied Mechanics and Materials*, **775**, pp. 229–233 (2015).

15. Rajesh, R. and Anand, M.D. "Prediction of EDM process parameters for AISI 1020 steel using RSM, GRA and ANN", *International Journal of Recent Technology and Engineering*, **8**, pp. 51–63 (2019).

16. Bharti, P.S. "Process modelling of electric discharge machining by back propagation and radial basis function neural network", *Journal of Information and Optimization Sciences*, **40**(2), pp. 263–278 (2019).

17. Ong, P., Chong, C.H., Rahim, M.Z.B., et al. "Intelligent approach for process modelling and optimization on electrical discharge machining of polycrystalline diamond", *Journal of Intelligent Manufacturing*, **31**, pp. 227–247 (2020).

18. Moghaddam, M.A. and Kolahan, F. "Modeling and optimization of the electrical discharge machining process based on a combined artificial neural network and particle swarm optimization algorithm", *Scientia Iranica*, B, **27**(3), pp. 1206–1217 (2020).

19. Sethuramalingam, P. and Sundararaj, O.N.R. "Adaptive neuro-fuzzy interference system modelling of EDM process using CNT infused copper electrode", *International Journal of Intelligent Engineering and Systems*, **10**(3), pp. 102–116 (2017).

20. Fazlollahtabara, H. and Gholizadeh, H. "Fuzzy possibility regression integrated with fuzzy adaptive neural network for predicting and optimizing electrical discharge machining parameters", *Computers & Industrial Engineering*, **140**, pp. 1–9 (2020).

21. Chau, N.L., Dao, T-P., and Nguyen, V.T.T. "Optimal design of a dragonfly-inspired compliant joint for camera positioning system of nanoindentation tester based on a hybrid integration of Jaya-ANFIS", *Mathematical Problems in Engineering*, Article ID 8546095, pp. 1–16 (2018).

22. Chau, N.L., Le, H.G., Dao, T-P., et al. "Efficient hybrid method of FEA-based RSM and PSO algorithm for multi-objective optimization design for a compliant rotary joint for upper limb assistive device", *Mathematical Problems in Engineering*, Article ID 2587373, pp. 1–14 (2019).

23. Johnson, R.A. and Wichern, D.W., *Applied Multivariate Statistical Analysis*, Pearson Prentice Hall, Upper Saddle River, New Jersey (2007).

24. Hair, J.F., Black, W.C., Babin, B.J., et al., *Multivariate Data Analysis*, Prentice-Hall, New Jersey (2010).

25. Brown, M.T. and Tinsley, H.E.A. "Discriminant analysis", *Journal of Leisure Research*, **15**(4), pp. 290–310 (1983).

26. Mansfield, E.R. and Helms, B.P. "Detecting multicollinearity", *The American Statistician*, **36**(3a), pp. 158–160 (1982).

27. Eisenbeis, R.A. "Pitfalls in the application of discriminant analysis in business, finance, and economics", *The Journal of Finance*, **32**(3), pp. 875–900 (1977).

28. Verma, J.P. "Application of discriminant analysis: For developing a classification model", In: *Data Analysis in Management with SPSS Software*, Springer, India, pp. 389–412 (2013).

29. McGarigal, K., Stafford, S., and Cushman, S. "Discriminant analysis", In: *Multivariate Statistics for Wildlife and Ecology Research*, Springer, New York, pp. 129–187 (2000).

30. Nath, R. and Pavur, R. "A new statistic in the one-way multivariate analysis of variance", *Computational Statistics & Data Analysis*, **2**(4), pp. 297–315 (1985).

31. Brown, M.T. and Wicker, L.R. "Discriminant analysis", In: *H.E.A. Tinsley and S.D. Brown, (Eds.)*, Handbook of Applied Multivariate Statistics and Mathematical Modeling, Academic Press, Elsevier, pp. 209–235 (2000).

32. Pituch, K.A. and Stevens, J.P., *Applied Multivariate Statistics for the Social Sciences*, Routledge, New York (2016).

33. Sarker, B. and Chakraborty S. "Discriminant analysis-based modeling of cotton fiber and yarn properties", *Research Journal of Textile and Apparel*, **26**(1), pp. 18–40 (2021). https://doi.org/10.1108/RJTA-08-2020-0092

34. Singh, B., Kumar, J., and Kumar, S. "Investigating the influence of process parameters of ZNC EDM on machinability of A6061/10% SiC composite", *Advances in Materials Science and Engineering*, Article 173427, pp. 1–8 (2013).

35. Sahani, O.P., Kumar, R., and Vashista, M. "Effect of electro discharge machining process parameters on material removal rate", *Journal of Basic and Applied Engineering Research*, **1**(2), pp. 17–20 (2014).

36. Puthumana, G. and Joshi, S.S. "Investigations into performance of dry EDM using slotted electrodes", *International Journal of Precision Engineering and Manufacturing*, **12**(6), pp. 957–963 (2011).

37. Gaitonde, V.N., Manjaiah, M., Maradi, S., et al. "Multiresponse optimization in wire electric discharge machining (WEDM) of HCHCr steel by integrating response surface methodology (RSM) with differential evolution (DE)", In: *J.P. Davim (Ed.), Computational Methods and Production Engineering*, pp. 199–223 (2017).

38. Sultan, T., Kumar, A., and Gupta, R.D. "Material removal rate, electrode wear rate, and surface roughness evaluation in die sinking EDM with hollow tool through response surface methodology", *International Journal of Manufacturing Engineering*, Article 259129, pp. 1–16 (2014).

Biographies

Baneswar Sarker obtained his B. Tech. and M.E. degrees in 2014 and 2020, respectively. He is now pursuing his doctoral research at Indian Institute of Technology, Kharagpur. His areas of interest are modeling and simulation of engineering systems and data analysis.

Shankar Chakraborty is a Professor of the Production Engineering Department at Jadavpur University. He obtained his BSc (Tech.) and M. Prod. E. degrees in 1986 and 1989, respectively. He was awarded a PhD (Engg.) degree from Jadavpur University in 1994. His areas of interest are optimization techniques, data mining, multi-criteria decision-making techniques, operations management, etc. He has published more than 200 research papers in many reputed Journals and is also aregular reviewer of several International Journals.