

14th CIRP Conference on Intelligent Computation in Manufacturing Engineering, CIRP ICME '20

Support vector machine regression for predicting dimensional features of die-sinking electrical discharge machined components

Kanka Goswami^a, GL Samuel^{a,*}

^aIndian Institute of Technology Madras, Chennai, 600036, India

* Corresponding author. Tel.: +91-44-2257-4699; fax: +91-44-2257-6699. E-mail address: samuelgl@iitm.ac.in

Abstract

Die-sinking electrical discharge machining produces components with low repeatability as the process is inherently stochastic. Effects of its inputs and process parameters on the components' dimensions are difficult to predict. This paper investigates the influence of input parameters like gap voltage, current, and pulse characteristics like percentage of “open”, “normal”, “arc” and “short” pulses on the dimensional features of the machined components. It discusses the methodology for extraction and estimation of amount of area machined, undercut and dimension by image processing. Support vector machine regression is applied to predict the dimension features based on the input and condition parameters.

© 2021 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

Peer-review under responsibility of the scientific committee of the 14th CIRP Conference on Intelligent Computation in Manufacturing Engineering, 15-17 July 2020.

Keywords: EDM; Support Vector Machine regression; Machine learning; Image processing

1. Introduction

Electrical discharge machining (EDM) finds its application for manufacturing difficult-to-machine with high accuracy like aerospace components, dies, diesel injectors, nozzles, and many more. The stochastic nature of EDM makes it challenging to predict the deviation of the machined profile causing low precision in machining. With die-sinking EDM, geometric inaccuracies like tapered holes, overcut, undercut are inevitable and difficult to predict. Attempts were made to address the precision of a similar process (wire-EDM) by prediction models developed through various machine learning techniques[1–3]. In these studies, prediction of the features with support vector machine regression (SVR) has been proven.

In the present work, the prediction of geometric features of the components machined with Die-sinking EDM machining is studied. The geometric features like “machined area”, “profile radius” and “undercut” were extracted from the images of the machined profiles. These features are predicted by developing respective SVR models from the input and the conditional (signal) features of the experiments. The discharge signals

acquired during the experiments were processed to extract the signal features. The SVR models were validated and tested by simulating unknown prediction conditions.

2. Experiment and Methodology

2.1. Experiments

The Experiments were performed in a Die-sinking EDM machine for 200 μ m depth. The machine inputs and the experimental conditions are given in table 1. A total of forty experiments were performed by varying voltage and the peak current as inputs. The tool was positive, and the workpiece was negative throughout the experiments. All the experiments were performed in a submerged condition and a medium flushing pressure of the dielectric fluid.

The overall experimental setup and the methodology is shown in Fig. 1. The current and the voltage data of the discharge pulses were captured by a Keysight DSO 2024A oscilloscope during the experiments. The captured pulse data were post-processed to extract the features for the predictors of the SVR models.

Table 1. Experimental conditions and input parameters

Input parameters	Values
Voltage (V)	5,10,15,20,25
Peak Current (A)	0.5,1,1.5,5,10,15,20,25
Pulse on time (μs)	100
Pulse off time (μs)	10
Depth of cut (μm)	200
Material (<i>workpiece</i>)	Cr-Steel
Material (<i>tool</i>)	Cu
Tool diameter (mm)	10
Dielectric	EDM oil
Machine Type	Die sinking submerged
Machine model	Electronica ZNC

After machining is performed, the images of the workpiece were taken with 18 megapixel digital camera. The images were post-processed to extract the geometrical and shape features for every experiment. The captured images were imported using ImageJ software and processed to identify the boundary of the machined profile. The shape and the geometric features were extracted from the processed images by using MATLAB software. The extracted features were used as the predictors and the targets for building the SVR models.

2.2. Data Acquisition and Signal Features Extraction

The pulse train data were captured at random for each experiment. The gap voltage data during the discharge were captured directly with the 300 MHz passive voltage probe. Gap voltage was measured directly between the tool and workpiece. The current data were acquired using the shunt method. Voltage-drop was measured to estimate the current across a resistor of 0.1Ω that was connected in parallel to the tool-gap. The signals were imported in MATLAB software for post-processing.

A typical discharge pulse in EDM has three phases. First, the charge accumulates between the electrodes and the voltage rises for some time. This phase is called “rise time” of the discharge pulse. Second, as the charge saturates and leaks through the electrodes, the voltage stabilizes in this region. The phase is called the “ignition delay time”. The dielectric is about to breakdown in the region. Finally, as the dielectric breaks down, the charge gushes towards the opposite electrodes and the electrode gap voltage falls. The phase is called the “breakdown time”. Thus, a gradient based pulse discrimination method was used to classify the signals into four pulse categories, viz. “Open”, “Normal”, “Arc” and “Short” pulses. The pulses were classified by calculating the discharge duration of each detected pulse based on the gradient of the voltage. A voltage gradient greater than $+3.75V/\mu s$ was taken to calculate the “rise time” of the pulse. For the gradient values between zero and $+3.75V/\mu s$ the “ignition delay” time was calculated and for a negative gradient, the “breakdown time” was computed. Typical discharge duration of $10\mu s$ was taken

as “Normal” category. The rise time, ignition delay time and the breakdown durations were added to determine single pulse

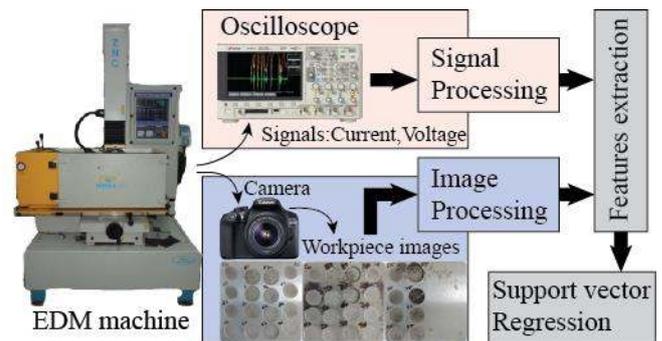


Fig. 1. Experimental setup.

machining (SP mach.) time and compared with the typical discharge duration for classification.

The total number of pulses for each pulse categories were recorded and their percentages were calculated. It was ensured that there were at least five pulse train samples captured during an experiment for estimation of the percentages of the pulse categories. The average SP mach. time was calculated by registering the pulse duration of each category and dividing it with total number of pulses. Several datasets for each experimental condition was used to extract the signal features used for predicting the dimensional features of the machined part.

2.3. Dimensional Features by Image Processing

The machining was performed with a copper rod of 10mm diameter. The tool impression on the workpiece was found to have significant amounts of undercut. Thus the desired circular profile was not obtained. In order to characterize these deviations, three geometric features of the obtained profiles, viz. machined area, amount of undercut and radius of the least square circle (LSC) fit for the machined profile were taken as investigation metrics. All these features were extracted through image processing from the workpiece image data. Fig. 2 shows the steps of the image processing to find out the dimensional features from the machined profile.

The captured images of the workpiece were calibrated to the real coordinates. 42 pixels corresponds to 1mm for all images. Each experimental profile was first imported in ImageJ software. The RGB picture was first converted to greyscale and then converted to a binary image by a greyscale threshold. Since the surface textures of the machined profiles for various experiments were different, the threshold values for each profile was manually selected. The threshold image resulted in the selection of the boundaries of the profiles. The coordinates of the boundary were extracted from the calibration. The identified boundary profile were clustered to form the regions (inside and outside) to calculate the inside area. The total number of pixels were counted and multiplied with the pixel calibration data to extract the “Area” feature. The computation

of the area and the extraction of the coordinate was performed using MATLAB software.

The extracted coordinates of the boundary were used to fit an LSC to estimate the radius of the machined profile. Geometric distance optimization method was used for fitting the curve as discussed in the next section.

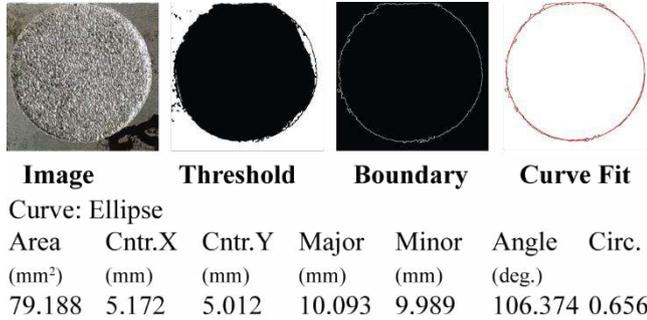


Fig. 2. Extracted Dimensional Features by Image Processing.

The radius of the LSC is taken as the second geometric feature for the investigations.

As there was undercutting in the machined profile due to a very low depth of cut, it was irregular and deviated from the desired circular profile. An ellipse was more appropriate fit to the machined profile. The ellipse was also fitted using least square method and the characteristics were computed. The difference between the tool area and area of the obtained ellipse was considered as the amount of undercut in the machined profile. The amount of undercut was identified as the third geometric feature.

2.4. Circle fitting with extracted coordinates

The most important feature that evaluates the dimensional accuracy is the circularity error for the given machined profiles. The coordinates were extracted from the processed files where the boundaries were estimated. A least square circle (LSC) is fitted geometrically through the extracted coordinates [4]. First, an equation of a circle is defined by eq. (1)

$$a\mathbf{P}^T\mathbf{P} + \mathbf{b}^T\mathbf{P} + c = 0 \tag{1}$$

Where a and c are scalars, \mathbf{P} is a $2 \times m$ matrix consisting of m set of coordinates extracted from the image data, and \mathbf{b} is a vector of the coefficients b_1 and b_2 .

Considering $\mathbf{z} = (z_1, z_2)^T$, the coordinate of the centre, and r , the radius of the LSC, the above coefficients are related as below.

$$\mathbf{z} = (z_1, z_2) = \left(-\frac{b_1}{2a}, -\frac{b_2}{2a} \right) \quad r = \sqrt{\frac{\|\mathbf{b}\|}{4a^2} - \frac{c}{a}}$$

To minimize the geometric distance, the error squared distance is given by, $d_i^2 = (\|\mathbf{P}_i - \mathbf{z}\| - r)^2$, d_i denotes the deviation of radius corresponding to the i^{th} row of vector \mathbf{P} , i.e. the i^{th} coordinate. The first term in the parenthesis computes the

Euclidian distance between the center of LSC and the i^{th} point. Let $\mathbf{u} = (z_1, z_2, r)^T$, the minimization objective function is given by

$$\sum_{i=1}^m d_i(\mathbf{u})^2 = \min. \tag{2}$$

The non-linear least squared function, eq. (2), is solved iteratively using Gauss-Newton method. Considering \mathbf{h} as the correction vector, the computer program updates $\mathbf{u}^{i+1} = \mathbf{u}^i + \mathbf{h}$

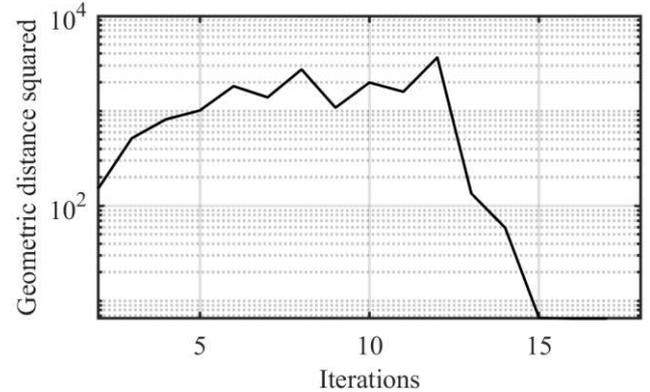


Fig. 3. Plot of the convergence of the geometric distance squared.

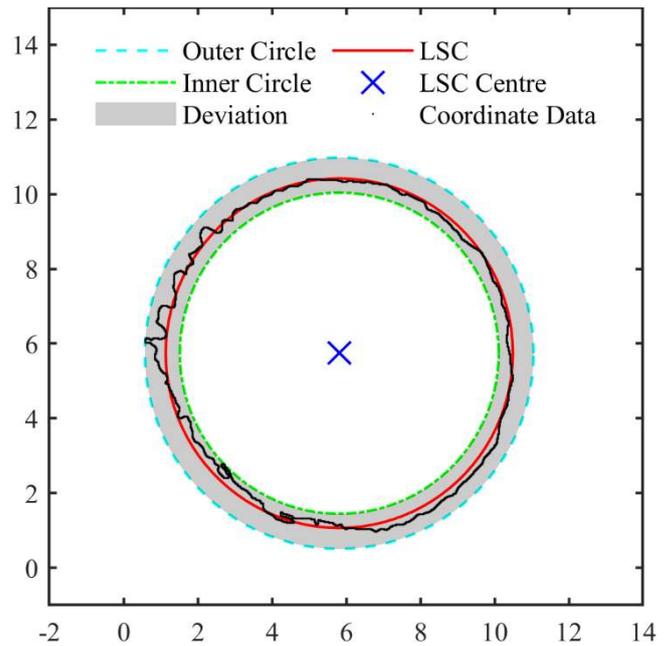


Fig. 4. Graph of LSC: centre, the outer and inner circles, and deviation (mm).

and finds the solution \mathbf{u}^* . Developing the function,

$$\mathbf{d}(\mathbf{u}) = (d_1(\mathbf{u}), d_2(\mathbf{u}), \dots, d_m(\mathbf{u}))^T$$

Using Taylor series to expand the developed function around the solution \mathbf{u}^* , approximated as $\hat{\mathbf{u}} + \mathbf{h}$, the function becomes

$$\mathbf{d}(\hat{\mathbf{u}} + \mathbf{h}) \approx \mathbf{d}(\hat{\mathbf{u}}) + J(\hat{\mathbf{u}})\mathbf{h} \approx 0 \tag{3}$$

Where, J represents the Jacobian matrix. Solving eq. (3) the correction vector \mathbf{h} is given by

$$J(\hat{\mathbf{u}})\mathbf{h} \approx -\mathbf{f}(\hat{\mathbf{u}}) \tag{4}$$

The Jacobian is given by

$$J(\mathbf{u}) = \begin{bmatrix} \frac{u_1 - x_1}{\sqrt{(u_1 - x_1)^2 + (u_2 - y_1)^2}} & \frac{u_1 - y_1}{\sqrt{(u_1 - x_1)^2 + (u_2 - y_1)^2}} & -1 \\ \vdots & \vdots & \vdots \\ \frac{u_1 - x_m}{\sqrt{(u_1 - x_m)^2 + (u_2 - y_m)^2}} & \frac{u_1 - y_m}{\sqrt{(u_1 - x_m)^2 + (u_2 - y_m)^2}} & -1 \end{bmatrix}$$

The convergence of the square geometric distance is shown in Fig. 3. In the figure, the initial value of $(z_1, z_2, r)^T$ for Gauss-Newton minimization was taken as $[3,5,3]^T$ and the convergence criterion was $\Delta d < 0.0001$.

The LSC radius is taken as the feature radius for the given data set. As shown in Fig. 4, the LSC for one of the data sets is plotted in continuous red line. The centre of the LSC is shown in the blue cross mark at $(5.8147, 5.7425)$ and its radius is 4.68 mm. The maximum deviation of 0.87 mm.

3. Theory and Modelling

3.1. Support Vector Machine Regression

Support vector machine regression (SVR) is a non-parametric regression model where the regression hyperplane is determined by optimizing the distances from the nearby data points known as support vectors. Non-linear SVR formulation is obtained by considering kernel-functions that considers the predictor interactions. The polynomial kernel-function is given by

$$g_{i,j} = \mathbf{G}(\mathbf{x}_i, \mathbf{x}_j) = (1 - \mathbf{x}_i^T \mathbf{x}_j)^q, \quad q \in \{2, 3, \dots\}$$

Where, $g_{i,j}$ is the entry in the Gram’s matrix, \mathbf{G} is the polynomial kernel-function, $\mathbf{x}_i, \mathbf{x}_j$ are the predictor variables, and q is the degree of the polynomial function. The solution is obtained by minimizing the predictor coefficients \mathbf{a}

$$\min_{\mathbf{a}} f(\mathbf{a}) = \frac{1}{2} \mathbf{a}^T \mathbf{Q} \mathbf{a} - \mathbf{e}^T \mathbf{a}; \text{ constraints } 0 < \alpha_i < C;$$

$$i = \{1, 2, \dots, m\}; \quad \mathbf{y}^T \mathbf{a} = 0; \text{ where, } Q_{i,j} = y_i y_j g_{i,j}$$

Here \mathbf{e} is a vector of ones, C is the upper bound and \mathbf{y} denotes the target variable classes for m number of predictors. The minimization for the cost function f is performed by sequential minimal optimization method [5].

3.2. Data preparation, model predictors and targets

The master dataset having the input conditions, the extracted signal parameters and the values of the extracted dimensional features were prepared and imported in the MATLAB software. Principal component analysis (PCA) was performed on the master dataset to understand the significance of each variable. The Pareto chart, as shown in Fig.5, identifies the significant variables that explain 95% variance in the master

dataset. The Pareto chart demarcates the predictors of the dataset. Table 2. shows the final predictors and targets of for building the SVR models.

Three different datasets were extracted from the master dataset, each containing one of the target features. Each column in the data sets of the predictors as well as targets, was normalized. The normalization eliminates the influence of the magnitude of the data on the training of the model. In the columns of the datasets like “open per cent”, “normal per cent” are fractions while the targets such as “radius” or “area” are ten to hundred times more than these predictor columns. The columns were normalized by re-scaling its mean to zero and the standard deviation to one.

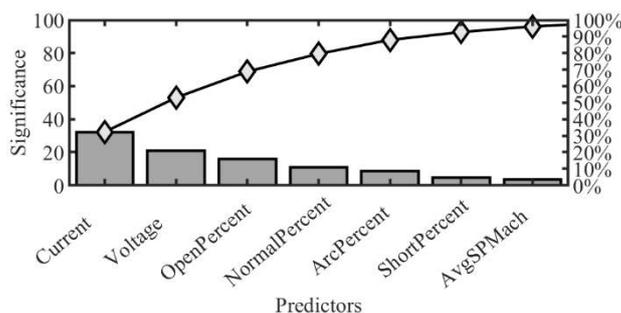


Fig. 5. Pareto chart showing significant predictors.

Table 2. Inputs predictors and target variables for the model

Predictors	Target
Current	Area (Image process)
Voltage	Undercut (Curve fit)
Open Pulse percent	Radius of LSC (Circle fit)
Normal Pulse Percent	
Arc Pulse percent	
Short Pulse percent	
SP mach. time (average)	

This process also ensures that the shape properties of the distribution of the original data for the variable is conserved.

Each of the three datasets, formed to predict the respective targets, was partitioned to hold out twenty per cent data used as test points at random. The rest of the data was used to build the model. As the test point data were avoided for building the SVR models, it was used to evaluate the performance of the model for unknown conditions.

4. Results and Discussion

4.1. Model Performance

After the data-partition, described in section 3.2, the remainder dataset for the model was again partitioned to form a training set with seventy per cent of the data and remaining for validation. The dataset was used to build the model with 5-fold cross-validation. This process generates 5 of training data

from the dataset with various combinations of training and validation data-points selected at random. The model was trained and validated five times with the generated data sets, and the k-fold loss is calculated as the average root mean square error (RMSE) in the five-fold validation of the model.

The overall performance index with the 5-fold cross-validation for each SVR models given in Table 3. As shown in the table, three separate models were developed for predicting “Area”, “Radius” and “Undercut”. The model parameters like “Kernel-function” and “Regularization” were adjusted until minimum RMSE value was obtained. The “Kernel-function”, the parameter which determines the order of the predictors, were varied as “linear”, “quadratic”, “cubic” and “Gaussian”. Higher orders for the “kernel-function” were ignored to avoid overfitting for the dataset. “Regularization”, a penalty term was also used for each model and increased in steps of 0.3 for each training model, to ensure generalization.

Table 3. Performance and properties of the SVR models

Sl. No.	Model Properties	Values
1	Area model	
	Kernel function	Cubic
	Regularization	6.6
	Validation	5 fold Cross-Validation
2	Radius model	
	Kernel function	Quadratic
	Regularization	4.0
	Validation	5 fold Cross-Validation
3	Undercut model	
	Kernel function	Cubic
	Regularization	2.3
	Validation	5 fold Cross-Validation
	RMSE	7.1956 mm ²
	RMSE	0.8136 mm
	RMSE	5.6709 mm ²

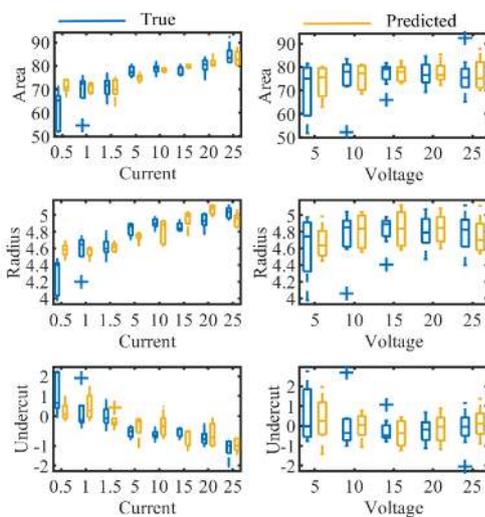


Fig. 6. Predictions and true responses of model validation for two predictors.

The validation response plots for each SVR model with “current” and “voltage” predictors are shown in Fig. 6. The variations in the predictions due to the 5-fold validation are shown as boxes and the mean value of the responses given by the horizontal line inside the respective box. The plots show that the variation of the target variables are large for current values between 0.5A to 1.5A, where the validation predictions drift from true responses. For other values of current, all the models perform with good accuracy of prediction. The validation responses for the voltage is also shown.

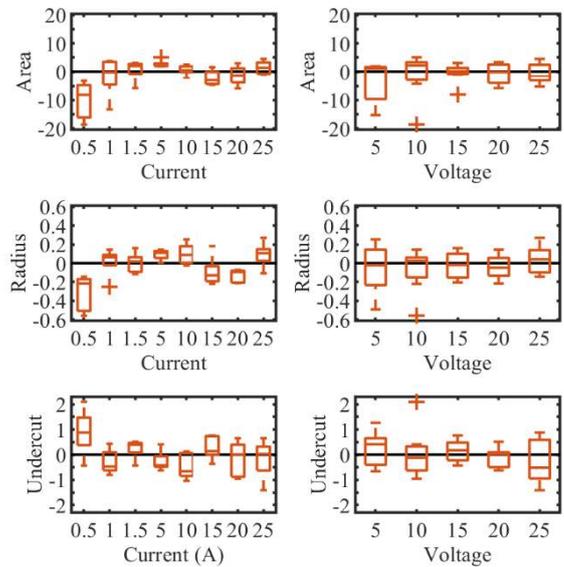


Fig. 7. Validation residuals of current and voltages for respective SVR models.

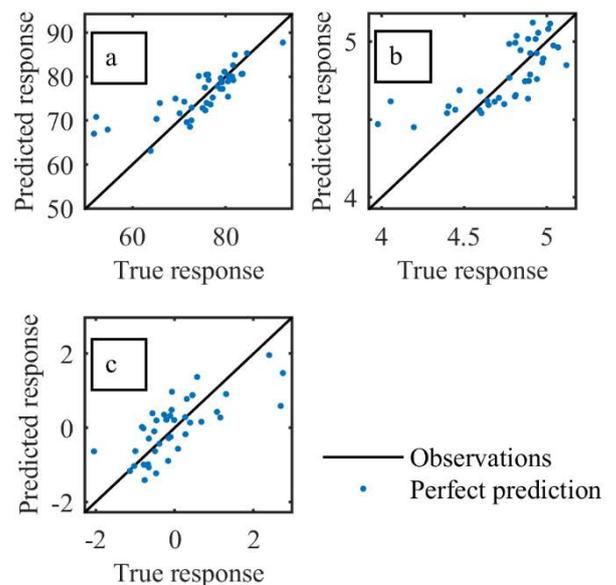


Fig. 8. Validation response for (a) Area, (b) Radius (c) Undercut SVR models.

It can be seen that the mean values of the validation predictions match the true responses. Thus, the voltage has a more

significant influence for predicting the dimensional features of the machined profile. It can be seen from Fig. 7 that the residuals of the current feature for predicting the dimensional feature is drifted and the mean values are away from zero. On the other hand, the voltage residuals for validation prediction have their mean values close to zero. Hence, the reliability of the dimensional feature prediction with voltage is better.

The overall performance of the validation prediction and true responses are shown in Fig 8. Fig. 8a shows the response tally for “area”. It can be seen that most of the predictions are very near to the observations.

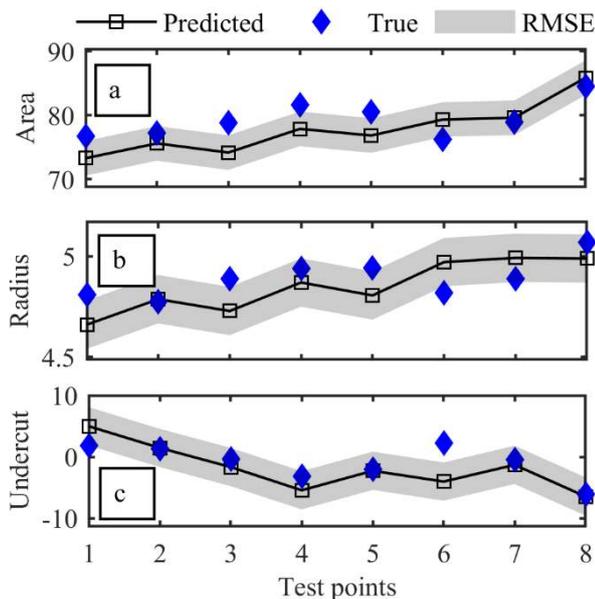


Fig. 9. Predicted responses with test point data for respective SVR models.

Table 4. Root mean square error of the test point data

Root Mean Square Error	Values
Area model	3.9389 mm ² (Lower than validation)
Radius model	0.1776 mm (Lower than validation)
Undercut model	3.6781 mm ² (Lower than validation)

Fig. 8b shows the prediction for “radius”. The predictions are not as good as area, but near to the observation. It implies that the validation of radius model has performed satisfactorily. Fig. 8c shows the response tally for “undercut” prediction and the model has also performed with reasonable accuracy.

4.2. Test performance of the Models

The held-out test points were used to simulate model prediction for unknown conditions. The test points were fed to each model to predict the dimensional features, respectively. The output of these models was recorded. RMSE was

computed with the predicted and the true responses of the test point data.

The model performances, evaluated by the closeness of the predicted and true responses are shown in Fig. 9. The RMSE of each model for the test scenario is given in table 4. It can be seen that for all the dimensional feature prediction models, the RMSE values with test point data are less than the models’ 5-fold cross-validation error, as given in table 3. Thus the developed model has performed better in unknown conditions. Thus, the developed SVR models can be used for predicting the dimensional features of the machined profiles reliably.

5. Conclusion

The present work had attempted to predict the dimensional features like, machined area, undercut, and radius of Die-sinking EDM machined part by formulating support vector regression (SVR) models. The models were constructed by identifying machine inputs (current and voltage) as well as conditional parameters (percentage of ‘open’, ‘normal’, ‘arc’, ‘short’ and average ‘SPmach. time’) as the predictors, and extracted dimensional features (machined area, undercut and machined radius) as targets. The predictor features were extracted through signal processing of the captured pulse data during the experiments. The target features were extracted from the workpiece images by identifying the machined boundary through threshold technique. The coordinate data were extracted from the boundary to fit the least square circle (LSC) by geometric distance optimization method.

Twenty per cent of the dataset were held-out for testing. Three exclusive SVR models for predicting machined area, undercut as well as the radius of the fitted LSC respectively were developed with the training set. For validation, Five-fold cross-validation method was used. Regularization on each model was also used to fit cubic-kernel functions for the area and undercut predictions, as well as to fit quadratic-kernel function for predicting the radius. The models have performed satisfactorily as the root mean square errors with the test dataset were found lower than validation errors for all cases.

The present work will be extended to also predict the surface integrity features of the machined components in future.

Acknowledgements

The research has not been funded by any agency.

References

- [1] A. Caggiano, R. Perez, T. Segreto, R. Teti, P. Xirouchakis, *Procedia CIRP* 2016;42:34–39.
- [2] A. Conde, A. Arriandiaga, J.A. Sanchez, E. Portillo, S. Plaza, I. Cabanes, *Robot. Comput. Integr. Manuf.* 2018;49:24–38.
- [3] J.A. Sanchez, A. Conde, A. Arriandiaga, J. Wang, S. Plaza, *Materials (Basel)*. 2018;11:1–12.
- [4] W. Gander, G.H. Golub, R. Strebler, *Bit* 1994;34:558–578.
- [5] P.H. Chen, R.E. Fan, C.J. Lin, *IEEE Trans. Neural Networks* 2006;17:893–908.