Article

Application of Generalized Regression Neural Network and Gaussian Process Regression for Modelling Hybrid Micro-Electric Discharge Machining: A Comparative Study

Siddhartha Kumar Singh 1, Harlal Singh Mali 1, Deepak Rajendra Unune 2, Szymon Wojciechowski 3,* and Dominik Wilczyński 3

1 Malaviya National Institute of Technology, Jaipur 302017, India; sid.sks@gmail.com (S.K.S.); harlal.singh@gmail.com (H.S.M.)
2 The LNM Institute of Information Technology, Jaipur 302031, India; deepunune@gmail.com
3 Faculty of Mechanical Engineering, Poznan University of Technology, 60-965 Poznan, Poland; dominik.wilczynski@put.poznan.pl
* Correspondence: szymon.wojciechowski@put.poznan.pl

Abstract: Micro-Electric Discharge Machining (µ-EDM) is one of the widely applied micromanufacturing processes. However, it has several limitations, such as a low cutting rate, difficult debris removal, and poor surface integrity, etc. Hybridization of the µ-EDM is proposed as an alternative to overcome the process limitations. Conversely, it complicates the process nature and poses a challenge for modelling and predicting critical process responses. Therefore, in this work, two distinct, nonparametric, previously unreported, workpiece material independent models using a Generalized Regression Neural Network (GRNN) and Gaussian Process Regression (GPR) were developed and compared to assess their performance with limited training data. Various smoothing factors and kernels were tested for GRNN and GPR, respectively. The prediction of models was compared in terms of the mean absolute percentage error, root mean square error, and coefficient of determination. The results showed that GPR outperforms GRNN and accurately predicts the µ-EDM process responses. The GRNN’s performance was better for less stochastic output with a discernible pattern than other outputs. The Automatic Relevance Determination (ARD) squared exponential kernel was found to be the best performing kernel among those chosen. GPR models can be used with reasonable accuracy to predetermine critical process outputs as they have R² values above 0.90 for both training and validation data for all outputs. This work paves the way for future industrial implementation of GPR to model and predict the outputs of complex hybrid machining processes.

Keywords: GPR; GRNN; application; micro-EDM

1. Introduction

Continuing miniaturization in numerous technical disciplines like the electronics industry or microsystems technology demands components with microfeatures such as holes, and channels, etc. Micro-electric discharge machining (µ-EDM) is one of the leading processes widely used for microfeature fabrication. Different kinds of materials, regardless of their hardness and toughness, could be machined easily using µ-EDM. However, the process is slow, and fabrication of deep features or high aspect ratios is a challenge for µ-EDM due to inefficient debris removal, low discharge energy, poorer machining stability, etc. To mitigate such challenges, hybridization of µ-EDM is proposed in the literature either by using energy assistance or through the combined application of one or more machining processes in material removal; to enhance specific process outputs. In energy-assisted EDM, the source could be a laser, a vibrating device, a magnet, etc. Although hybridization certainly improves the process, it also severely complicates the nature of the process, and it becomes challenging to establish a mathematical model for such a process. In the µ-EDM,
the material removal arises due to melting and evaporation as a spark is generated among the tool and workpiece in a dielectric fluid. Adding other energy sources poses difficulties in estimating the time and tool required for a particular job and in estimating the final surface characteristics of the manufactured feature. Surfaces, particularly in the microdomain, play a critical role and thus have to be tailored according to need and application.

Past and contemporary research in the area of hybrid \( \mu \)-EDM has been chiefly aimed toward demonstrating process capabilities, whether it be laser hybridization [1], magnetic field-assisted hybrid \( \mu \)-EDM [2], ultrasonic or low-frequency vibration-assisted [3–5] or a combination of micro-EDM with electrophoretic deposition [6]. Attempts toward modelling the process have also been made using methods including the finite element method (FEM) [7], fluid dynamics [8], graph theory [9], and empirical modelling [10,11]. However, the comparison of different methods for modelling and predicting machining response characteristics like cutting rate, tool wear rate, etc., are not reported significantly [9,11]. Unune et al. [10] reported predicted values using empirical models and were found to be within 7% of experimental results. In recent work, a comparison was made between ANN-Particle Swarm Optimization (PSO) and FEM-based models for predicting process outputs of the micro-EDM drilling process without any hybridization [12]. They found that ANN-PSO could more accurately predict the material removal rate and dimensional deviation. In another work, the authors found the adaptive neuro-fuzzy inference system (ANFIS) model to perform better than ANN in predicting micro-EDM responses, viz., metal removal rate, surface roughness, etc. and tool wear ratio [13]. However, there was no hybridization of the micro-EDM process in both those works. At the macro scale, numerical and thermal modelling of hybrid electric discharge and arc machining was attempted to consider latent heat and temperature-dependent workpiece properties [14]. Further literature on the implementation of recent machine learning techniques such as Backpropagation Neural Network (BPNN), Generalized Regression Neural Network (GRNN), Gaussian Process Regression (GPR), etc., to model this complicated process could not be found.

The prime benefit of different machine learning techniques lies in modelling and predicting highly complicated nonlinear processes and phenomena. ANNs are very popular and have been extensively used for modelling complex manufacturing process outputs such as microchannel dimensions [15], intelligent manufacturing [16], shear strength [17], and material removal rate [18], etc. However, ANNs require a large number of parameters to be correctly tuned. This limits their usage in cases where the training dataset is limited. On the other hand, in GRNN, only one parameter needs to be optimized to fit the model. Moreover, GRNN has shown better efficacy for predictive modelling [19]. This makes GRNN a suitable candidate for modelling processes such as hybrid \( \mu \)-EDM, which are highly energy-intensive and slow processes (thus restricting the feasibility of performing a large number of experiments) but in which the prediction of different process outputs is also critical.

GPR, like GRNN, offers a probabilistic and nonparametric approach to modelling. But certain hyper-parameters are still required to be optimized for better model fitting. GPR has been used to model and predicts many complex phenomena such as predicting oxygen consumption in steel making process [20], deposit shape in cold metal transfer based wire arc additive manufacturing [21], predicting the ultimate tensile strength in friction stir welding [22], predicting weld quality uncertainty in hybrid-tandem gas metal arc welding [23], etc. Although GPR has not been used to model hybrid \( \mu \)-EDM, it has been used to model Wire-EDM [24,25] and EDM [26]. The authors in these articles find GPR more accurate than BPNN and find it suitable for modelling with small datasets. Thus, it is only natural to attempt to model hybrid \( \mu \)-EDM using GPR.

In light of the discussed literature, it is clear that GRNN and GPR techniques could be applied for modelling complex hybrid machining. Such studies are rarely available in the literature. The application of GRNN and GPR for modelling vibration-assisted micro-EDM has not been reported previously. Thus, this article explores the efficacy of two different nonparametric modelling techniques, GRNN and GPR, for predicting the
different significant responses in a hybrid μ-EDM process, which has not been previously attempted. Therefore, in this work, blind microholes of various depths were initially fabricated using μ-EDM on different materials by varying the energy input and vibration frequency. Post-fabrication, three key process outputs were measured, including hole drilling rate, rate of tool wear, and centerline average surface roughness of the bottom surface. These outputs are then modelled using GRNN and GPR, and a comparison is made in terms of mean absolute percentage error (MAPE), root mean square error (RMSE), and coefficient of determination ($R^2$).

2. Microfeature Fabrication Process

Figure 1 displays a schematic diagram of the machining setup and shows the microhole fabrication process using μ-EDM. In this process, the electrode, also referred to as the tool, is given a negative charge while the workpiece with a positive charge. This electrode is attached to a spindle and rotated at a selected rotation per minute (RPM). The workpiece is placed upon a vibrating device (which is the additional source of energy). A suitable dielectric fills the gap between them. As they are brought closer to each other, at a certain distance, the dielectric breaks down, and a spark is initiated between the two. The high heat generated due to the spark melts and evaporates material from both the electrode and workpiece. As a result, a cavity of the shape of the rotating electrode is formed in the workpiece while some portion of the electrode is worn out.

![Figure 1. Schematic of microfeature fabrication on μ-EDM.](image)

3. Experimentation & Process Outputs

This study selected AISI 316 SS, Co29Cr6Mo, and BeCu as work materials. A WC electrode of ø0.5 mm was used as a tool. Experimental runs were carried out on DT-110i multi-purpose micromachine, as shown in Figure 2. A constant tool speed of 1500 rpm and feed rate of 0.2 mm/min were used during machining. An electromagnetic actuator was used to vibrate the work material. The constant electrode RPM and feed rate were selected based on previous pilot experiments conducted for different materials. The machine was found to function best at the selected constant parameters.

The discharge energy (DE) in RC circuit based micro-EDM typically depends on the voltage and capacitance values. DE is a product of voltage and capacitance square. The vibrational frequency has been identified to be playing an important role in improving micro-EDM performance [4]. However, work material and aspect ratio have not been researched well in the available literature. Therefore, to conduct experimentation, the energy input, vibrational frequency, and microhole aspect ratio were selected as the control factors; in addition, the work material is considered a categorical factor. The details regarding the control factors are provided in Table 1. The aspect ratio is taken as the ratio of intended hole depth to the electrode diameter. The Central Composite Design (CCD) approach was used to plan 31 experimental runs for control factors, as shown in Table 2. In addition, three more experiments were conducted, one on each material, at new design
points to help in model validation. Figure 3 shows the fabricated microholes on the AISI 316 SS workpiece.

\[
DR = \frac{\text{Actual depth of drilled hole(\(\mu\)m)}}{\text{Machining time(min)}} 
\]

(1)

\[
TWR = \frac{\text{Length of tool wornout(\(\mu\)m)}}{\text{Actual depth of drilled hole(\(\mu\)m)}} 
\]

(2)

Drill rate (\(DR\)), tool wear rate (\(TWR\)), and arithmetical mean height of surface roughness (\(Ra\)) were the process outputs. \(DR\) and \(TWR\) were calculated using Equations (1) and (2), respectively. \(DR\) is basically the depth of hole drilled per unit time, and \(TWR\) is the length of tool worn out per unit depth of hole drilled. Machining time was verified using a digital stopwatch, while tool wear was measured as tool height loss pre- and post-machining. The arithmetical mean height of surface roughness (\(Ra\)) of the hole bottom and depth of the drilled hole was measured using the Profilm 3D-FILMETERS\textsuperscript{®} white light Interference profilometer. Three measurements were made at different hole locations and averaged to obtain the mean \(Ra\) value.

![Micro-EDM machine used for experimentation.](image)

**Figure 2.** Micro-EDM machine used for experimentation.

![Fabricated microholes on the AISI 216 SS workpiece.](image)

**Figure 3.** Fabricated microholes on the AISI 216 SS workpiece.

<table>
<thead>
<tr>
<th>Table 1. Parameters and their levels.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>(M) Material</td>
</tr>
<tr>
<td>(v) Vibration frequency (Hz)</td>
</tr>
<tr>
<td>(E) Energy (mJ)</td>
</tr>
<tr>
<td>(AR) Aspect ratio</td>
</tr>
</tbody>
</table>
Table 2. Input–output table.

<table>
<thead>
<tr>
<th>Expt. No.</th>
<th>M</th>
<th>v</th>
<th>E</th>
<th>AR</th>
<th>DR (µ/min)</th>
<th>TWR (µm/µm)</th>
<th>Ra (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>62.31</td>
<td>1.02</td>
<td>5.646</td>
</tr>
<tr>
<td>2</td>
<td>+1</td>
<td>+1</td>
<td>+1</td>
<td>−1</td>
<td>137.53</td>
<td>0.25</td>
<td>8.051</td>
</tr>
<tr>
<td>3</td>
<td>+1</td>
<td>−1</td>
<td>−1</td>
<td>−1</td>
<td>75.03</td>
<td>0.10</td>
<td>3.275</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>+1</td>
<td>0</td>
<td>94.44</td>
<td>1.16</td>
<td>12.780</td>
</tr>
<tr>
<td>5</td>
<td>+1</td>
<td>+1</td>
<td>−1</td>
<td>+1</td>
<td>165.75</td>
<td>0.10</td>
<td>3.190</td>
</tr>
<tr>
<td>6</td>
<td>−1</td>
<td>+1</td>
<td>−1</td>
<td>+1</td>
<td>14.11</td>
<td>0.38</td>
<td>3.069</td>
</tr>
<tr>
<td>7</td>
<td>+1</td>
<td>−1</td>
<td>−1</td>
<td>+1</td>
<td>59.67</td>
<td>0.10</td>
<td>1.985</td>
</tr>
<tr>
<td>8</td>
<td>−1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>71.61</td>
<td>1.01</td>
<td>5.334</td>
</tr>
<tr>
<td>9</td>
<td>−1</td>
<td>+1</td>
<td>−1</td>
<td>−1</td>
<td>16.75</td>
<td>0.31</td>
<td>2.312</td>
</tr>
<tr>
<td>10</td>
<td>+1</td>
<td>+1</td>
<td>−1</td>
<td>−1</td>
<td>147.17</td>
<td>0.13</td>
<td>2.916</td>
</tr>
<tr>
<td>11</td>
<td>−1</td>
<td>−1</td>
<td>+1</td>
<td>−1</td>
<td>105.30</td>
<td>0.62</td>
<td>9.732</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>62.45</td>
<td>1.03</td>
<td>7.128</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>72.79</td>
<td>1.00</td>
<td>8.136</td>
</tr>
<tr>
<td>14</td>
<td>−1</td>
<td>−1</td>
<td>−1</td>
<td>+1</td>
<td>11.73</td>
<td>0.34</td>
<td>3.331</td>
</tr>
<tr>
<td>15</td>
<td>+1</td>
<td>+1</td>
<td>+1</td>
<td>+1</td>
<td>167.29</td>
<td>0.28</td>
<td>10.150</td>
</tr>
<tr>
<td>16</td>
<td>−1</td>
<td>+1</td>
<td>+1</td>
<td>−1</td>
<td>85.37</td>
<td>0.96</td>
<td>8.550</td>
</tr>
<tr>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>88.88</td>
<td>0.87</td>
<td>7.281</td>
</tr>
<tr>
<td>18</td>
<td>0</td>
<td>−1</td>
<td>0</td>
<td>0</td>
<td>32.08</td>
<td>1.09</td>
<td>7.940</td>
</tr>
<tr>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>+1</td>
<td>73.34</td>
<td>1.05</td>
<td>8.937</td>
</tr>
<tr>
<td>20</td>
<td>0</td>
<td>+1</td>
<td>0</td>
<td>0</td>
<td>103.00</td>
<td>0.93</td>
<td>6.506</td>
</tr>
<tr>
<td>21</td>
<td>−1</td>
<td>−1</td>
<td>+1</td>
<td>+1</td>
<td>82.43</td>
<td>1.28</td>
<td>10.020</td>
</tr>
<tr>
<td>22</td>
<td>+1</td>
<td>−1</td>
<td>+1</td>
<td>−1</td>
<td>139.07</td>
<td>0.24</td>
<td>8.548</td>
</tr>
<tr>
<td>23</td>
<td>+1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>167.08</td>
<td>0.17</td>
<td>9.186</td>
</tr>
<tr>
<td>24</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>72.81</td>
<td>0.97</td>
<td>8.403</td>
</tr>
<tr>
<td>25</td>
<td>−1</td>
<td>+1</td>
<td>+1</td>
<td>+1</td>
<td>84.38</td>
<td>1.22</td>
<td>9.922</td>
</tr>
<tr>
<td>26</td>
<td>0</td>
<td>0</td>
<td>−1</td>
<td>0</td>
<td>16.17</td>
<td>0.38</td>
<td>3.855</td>
</tr>
<tr>
<td>27</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>−1</td>
<td>61.40</td>
<td>1.19</td>
<td>4.146</td>
</tr>
<tr>
<td>28</td>
<td>−1</td>
<td>−1</td>
<td>−1</td>
<td>−1</td>
<td>14.18</td>
<td>0.32</td>
<td>3.308</td>
</tr>
<tr>
<td>29</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>71.89</td>
<td>1.00</td>
<td>6.804</td>
</tr>
<tr>
<td>30</td>
<td>+1</td>
<td>−1</td>
<td>+1</td>
<td>+1</td>
<td>146.75</td>
<td>0.29</td>
<td>7.231</td>
</tr>
<tr>
<td>31</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>96.93</td>
<td>0.73</td>
<td>7.862</td>
</tr>
</tbody>
</table>

V1

| V2 | −1 | +1 | +1 | −1 | 61.71      | 1.01        | 6.581  |
|    | 0 | −1| 0  | 0  | 8.57       | 0.34        | 2.986  |
|    | +1| 0 | −1| +1 | 233.60     | 0.25        | 7.989  |

4. Modelling

4.1. GRNN Model

GRNN is a neural network that does not require any backpropagation to reach the optimal solution. Data is passed only once through the network, and thus, it is also known as a feed-forward type network. The network consists of four layers, viz. input, pattern, summation, and output, as shown in Figure 4.
The neurons at the input layer correspond to the independent variables that affect the output. Each input neuron is connected to all pattern neurons. The neurons in the pattern layer correspond to vectors in the target set. At this layer, the Euclidean distance \((d_i)\) between the input and training data is calculated, and an exponential activation function is used to determine the weights. The output from the pattern layer is then fed into two summation neurons—\(S_j\) and \(S_d\), also known as numerator and denominator, respectively. The formula for calculating \(S_j\), \(S_d\) and \(d_i\) are provided in Equations (3)–(5), respectively.

\[
S_j = \sum_{i=1}^{n} Yi \exp \left[ -\frac{d_i^2}{2\sigma^2} \right] \tag{3}
\]

\[
S_d = \sum_{i=1}^{n} \exp \left[ -\frac{d_i^2}{2\sigma^2} \right] \tag{4}
\]

where,

\[
d_i^2 = (X - Xi)^T (X - Xi) \tag{5}
\]

The data from these neurons are then divided to calculate the final output, \(Y(X)\), as shown in Equation (6).

\[
Y(X) = \frac{S_j}{S_d} = \frac{\sum_{i=1}^{n} Yi \exp \left[ -\frac{d_i^2}{2\sigma^2} \right]}{\sum_{i=1}^{n} \exp \left[ -\frac{d_i^2}{2\sigma^2} \right]} \tag{6}
\]

\(Xi\) represents the input used during training, while \(X\) represents sample input. \(Yi\) represents the target set used for training, and \(Y(X)\) is the calculated output corresponding to input vector \(X\), while \(n\) is the total number of training datasets. \(\sigma\) is the smoothing factor, also known as the spread, and is the only tunable parameter that determines the accuracy of the network. MATLAB R2018b® was used to perform the GRNN modelling in this work.

4.2. GPR Model

Gaussian Process Regression (GPR), a class of machine learning, obtains nonparametric and probabilistic models for response variables of the process under consideration. Rasmussen and Williams [27] define a Gaussian process as “a collection of random variables, any finite number of which have a joint Gaussian distribution”. A Gaussian factor established the relationship among process parameters and response variables. More details regarding GPR modelling are available in our previous paper [5]. Several categories of kernel functions can be selected for establishing the model. The exponential, squared exponential, quadratic, rational quadratic, and kernel functions are widely used.

If \(y\) is the output for an input vector \(x\), then for a Gaussian process,

\[
y = f(x) + \epsilon \tag{7}
\]
Such that both $f(x)$ and $\epsilon$ are typically distributed. This is to say:

$$
\epsilon \sim N\left(0, \sigma^2\right)
$$

(8)

$$
f(x) \sim N\left(\mu(x), k(x, x')\right)
$$

(9)

where $\mu(x)$ is the mean function and $k(x, x')$ is the covariance function. The process can be simplified, assuming the prior mean function as zero. Hence, for a set of response variables $y$ corresponding to a set of input process parameters $X$, the joint distribution for a set of predictions $y_{\text{new}}$ corresponding to input $X_{\text{new}}$ is:

$$
[y_{\text{new}}] \sim N\left(0, [K(X, X) + \sigma^2IK(X_{\text{new}}, X_{\text{new}})K(X_{\text{new}}, X_{\text{new}})]\right)
$$

(10)

where $I$ is the identity matrix. The posterior predictions as a Gaussian distribution can be given as

$$
y_{\text{new}} \sim N\left(y_{\text{new}}, \text{cov}(y_{\text{new}})\right)
$$

(11)

$$
y_{\text{new}} = K(X_{\text{new}}, X)[K(X, X) + \sigma^2I]^{-1}y,
$$

(12)

then

$$
\text{cov}(y_{\text{new}}) = K(X_{\text{new}}, X_{\text{new}}) - K(X_{\text{new}}, X)K(X, X) + \sigma^2I\right]^{-1}K(X, X_{\text{new}})
$$

(13)

Although the method is a nonparametric modelling, specific hyperparameters dependent on the covariance function are required to be optimized for a better fitting of the model. Further details are available in the study by Rasmussen and Williams [27]. For GPR modelling, MATLAB R2018b® was used.

In this work, different ARD-based kernels were chosen to model as they allow dimensional independence, which means a separate length scale is possible for each predictor. Exponential, squared exponential, Matern 3/2, Matern 5/2, and rational quadratic ARD kernels were chosen and defined below.

ARD Exponential:

$$
K(X_i, X_j) = \sigma_s^2 \exp(-d)
$$

(14)

ARD Squared Exponential:

$$
K(X_i, X_j) = \sigma_s^2 \exp\left(-\frac{1}{2}d^2\right)
$$

(15)

ARD Matern 3/2:

$$
K(X_i, X_j) = \sigma_s^2 \left(1 + \sqrt{3}d\right) \exp\left(-\sqrt{3}d\right)
$$

(16)

ARD Matern 5/2:

$$
K(X_i, X_j) = \sigma_s^2
$$

(17)

ARD Rational Quadratic:

$$
K(X_i, X_j) = \sigma_s^2 \left(1 + \frac{r^2}{2\alpha \sigma_l^2}\right)^{-\alpha}
$$

(18)

where,

$$
r = \sqrt{\sum_{m=1}^{d} \frac{(Xim - Xjm)^2}{\sigma_m^2}}
$$

(19)

$\sigma_s$ is the signal standard deviation, $\sigma_l$ is the specific length scale, $\alpha$ is the scale-mixture factor and $m$ is the number of predictors.
4.3. Model Comparison & Discussion

4.3.1. Drill Rate

The first attempt at modelling drill rate was made using GRNN. As the smoothing factor is the only parameter on which the prediction accuracy of a GRNN depends, thus, its optimization is crucial. The smoothing factor varies between 0 and 1. A significantly lower value of smoothing factor near 0 may show better results while training but cannot predict accurately as they have poor generalization. On the other hand, a smoothing factor of 1 may have a better conception but would also have a much higher value of error [5]. In addition, the same smoothing factor may not be adequate for all models, and thus, for every model, a search must be performed to determine the appropriate value of the smoothing factor. The model was trained using the 31 data points with different smoothing factors. Then, to test the model adequacy, the output for the three validation experiments (V1, V2 & V3) were compared with experimental results. The results are shown in Figure 5, in which Sigma denotes the smoothing factor. It is observed that GRNN is not much adept at predicting these values. Gross under-prediction is observed for the validation experiment performed on BeCu irrespective of the smoothing factor, while severe over-prediction is observed for experimentation on Co29Cr6Mo. At lower smoothing factor values, GRNN fails to predict the trend, and it was only after the value of the smoothing factor was above 0.6 that the prediction began to resemble the experimental trend. It was found that the smoothing factor of 0.6 (shown in black) provided the best fit for the validation experiments as it had the least mean absolute error (MAE) and mean square error (MSE), and thus, it was chosen as the final value for the model. The resulting predicted values for all the experiments and their comparison with the actual experimental values are shown in Figure 6.

After that, GPR was used to model the drill rate. Different ARD-based kernels were chosen to model as they allow dimensional independence, which means for each predictor, a separate length scale is possible. Exponential, squared exponential, Matern 3/2, Matern 5/2, and rational quadratic ARD kernels were tested, and the results for the three validation experiments, similar to GRNN, were evaluated to test the models. The results are shown in Figure 7. It can be observed that all kernels show better efficacy at demonstrating the general trend than GRNN models with different smoothing factors. From these different kernels, the ARD squared exponential (shown in black) was chosen for the final model as it has the lowest MAE and MSE. The predicted values of this model, prediction intervals for 95% confidence level, and corresponding experimental findings are displayed in Figure 8.

Figure 5. GRNN model testing for DR with different smoothing factors.
After that, GPR was used to model the drill rate. Different ARD-based kernels were chosen to model as they allow dimensional independence, which means for each predictor, a separate length scale is possible. Exponential, squared exponential, Matern 3/2, Matern 5/2, and rational quadratic ARD kernels were tested, and the results for the three validation experiments, similar to GRNN, were evaluated to test the models. The results are shown in Figure 7. It can be observed that all kernels show better efficacy at demonstrating the general trend than GRNN models with different smoothing factors. From these different kernels, the ARD squared exponential (shown in black) was chosen for the final model as it has the lowest MAE and MSE. The predicted values of this model, prediction intervals for 95% confidence level, and corresponding experimental findings are displayed in Figure 8.

Figure 6. Experimental and GRNN predicted values for DR training data.

Figure 7. GPR model testing for DR with different kernels.

Figure 8. Experimental and GRNN predicted values for DR.

Figure 9 presents the comparison of these two models in terms of MAPE, RMSE, and $R^2$. The extension "-V" represents the value of these indices for validation experiments only. The histograms show that GPR performs better than GRNN at predicting DR. A substantial difference between these two models is distinct for all the characteristic indices. GPR shows a good correlation with the experimental results for the training data and the validation data. However, it should also be noted that RMSE increases almost five-fold for the validation data, even for GPR; however, it follows the experimental trend as seen in Figure 8.
only. The histograms show that GPR performs better than GRNN at predicting DR. A substantial difference between these two models is distinct for all the characteristic indices. GPR shows a good correlation with the experimental results for the training data and the validation data. However, it should also be noted that RMSE increases almost five-fold for the validation data, even for GPR; however, it follows the experimental trend as seen in Figure 8.

Figure 9. Comparison histogram of model characteristics.

4.3.2. Tool Wear Rate

For modelling tool wear rate, a similar procedure was followed. First, different smoothing factors were tested for GRNN; the best value was chosen based on the least error for validation data. It was found that the error was minimum when the smoothing factor was 1, as seen in Figure 10. For TWR, too, it is observed that GRNN fails to follow the experimental trend as in the case with DR. Figure 11 shows the GRNN predicted values and their comparison with the experimental values. GPR again excels at modelling, and in fact, the difference among the various kernels is minute, as can be observed from Figure 12. Still, the ARD squared exponential has the best performance in terms of MSE and MAE.

The GPR predicted results and the corresponding 95% confidence level prediction interval, along with the experimental data, are shown in Figure 13. The histogram of model performance indices, as shown in Figure 14, also highlights this difference.

Figure 10. GRNN model testing for TWR with different smoothing factors.
Figure 11. Experimental and GRNN predicted values for TWR training data.

Figure 12. GPR model testing for TWR with different kernels.

Figure 13. Experimental and GPR predicted values for TWR training data.
4.3.3. Surface Roughness

When it comes to the roughness of the machined surface, like TWR, a smoothing factor of 1 was again observed to be the optimum value and was used to predict the outputs using GRNN for model comparison. The results are shown in Figures 15 and 16. Unsurprisingly, the ARD squared exponential came out as the best kernel function, as seen in Figure 17, and was used to predict the outputs, as shown in Figure 18.

Although GPR still outperforms GRNN, the difference between the two models is not as significant as in the case of the other two outputs. This is evident from the performance indices histogram, as shown in Figure 19. The GRNN model has a usable R² value for the training data and the validation experiments. MAPE and RMSE do not fall too far behind the GPR model. GRNN, in this case, proves to be better at modelling surface roughness than tool wear or drill rate.

This improvement in the performance of GRNN is possible since, unlike the other two outputs, a discernible pattern is evident in the output data of surface roughness, as seen in Figure 20. The discharge energy has a far more significant influence on surface roughness than other outputs, and thus, an almost linear relationship between the two is observed. This ‘simplification’ helps GRNN perform better than other outputs as such a pattern is not observable in the data for drill rate and tool wear rate. However, despite this enhanced performance, it still lags behind GPR in every aspect.
Figure 16. Experimental and GRNN predicted values for Ra training data.

Figure 17. GPR model testing for $R_a$ with different kernels.

Figure 18. Experimental and GPR predicted values for Ra training data.
Although GPR still outperforms GRNN, the difference between the two models is not as significant as in the case of the other two outputs. This is evident from the performance indices histogram, as shown in Figure 19. The GRNN model has a usable $R^2$ value for the training data and the validation experiments. MAPE and RMSE do not fall too far behind the GPR model. GRNN, in this case, proves to be better at modelling surface roughness than tool wear or drill rate.

Figure 19. Comparison histogram of model characteristics for $Ra$.

This improvement in the performance of GRNN is possible since, unlike the other two outputs, a discernible pattern is evident in the output data of surface roughness, as seen in Figure 20. The discharge energy has a far more significant influence on surface roughness than other outputs, and thus, an almost linear relationship between the two is observed. This 'simplification' helps GRNN perform better than other outputs as such a pattern is not observable in the data for drill rate and tool wear rate. However, despite this enhanced performance, it still lags behind GPR in every aspect.

Figure 20. Experimental Ra data points stacked against energy.

5. Conclusions

In this article, two distinct approaches, namely GRNN and GPR, were applied for modelling and prediction of process performance of a complex microscale hybrid machining process. The performance of GRNN and GPR in the prediction of process outputs with limited training data was analyzed in terms of MAPE, RMSE, and $R^2$. Initially, the experiments were planned using the CCD approach of RSM and then were conducted to fabricate microholes using vibration-assisted hybrid $\mu$-EDM on different materials to eliminate the workpiece dependency of models. The energy input to the machine, vibration frequency, and aspect ratio of microfeature were varied to conduct the experiments. Their effects were assessed on the rate of drilling, rate of tool wear, and roughness of the machined surface are crucial outputs of the process. The following inferences may be drawn from the conducted work:

- Among the two approaches tested, the prediction accuracy of the GPR model outperformed GRNN in the case of all output parameters.
- GRNN developed models for drill rate and tool wear rate cannot be used for predictions as the correlation for validation data points is extremely low, and MAPE and RMSE errors are extremely high.
For surface roughness, GRNN performance improved significantly but still lags behind GPR. It was observed that this improvement is possible because surface roughness is less stochastic in nature and has a direct relationship with energy input.

An ARD squared exponential kernel was observed to be the best performing kernel for all output models using GPR.

Based on the work done in this article, it may be concluded that GPR is a much better candidate than GRNN for modelling seemingly stochastic outputs of complex machining processes even with limited training data.

6. Future Work

Considering the proven efficacy of GPR in this research work, GPR models can be implemented in the future, particularly for vibration-assisted micro-EDM, to predetermine with reasonable accuracy the time and tool requirements for performing a particular drilling operation. Using more training data, the model may be extended in the future for other operations performed using hybrid micro-EDM, such as micromillings, to develop a single holistic model to predict the output of all operations performed with vibration-assisted hybrid µ-EDM. The extension of this work could also be in the direction of combining the approach of Big Data along with GPR modelling by extracting any possible real-time continuous data from different machining processes.


Funding: This research was funded by the Department of Science and Technology, Government of India, through the grant DST-SERB EMR/2016/003372, and Ministry of Education and Science (Poland) grant numbers: 0611/SBAD/0115, 0614/SBAD/1547.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References
6. Tsui, H.-P.; Hsu, S.-Y. Study on Fe-Based Metallic Glass Micro Hole Machining by Using Micro-EDM Combined with Electrophoretic Deposition Polishing. Processes 2022, 10, 96. [CrossRef]


