A parameter prediction model of running-in based on surface topography

Gengpei Zhang, Xiaojun Liu and Wenlong Lu

Abstract
Although surface topography is a directive and useful method to evaluate the quality of worn components, the common belief that the surface roughness after running-in is independent of the nature of initial roughness makes it difficult to investigate the surface topography of worn components after running-in based on unworn surface before running-in process. Trying to build the connection of surface topography before and after running-in, this article assumed running-in as a black box and adopted support vector machine as machine learning method to simulate the complex process. Through the training and validation of the simulation, the predictive model of surface topography after running-in based on unworn surface topography was established, which indicated the existence of a correlation between the surface topographies before and after running-in. The model analysis revealed that the hybrid properties of surface topographies before and after running-in have strong correlation. Especially, the correlation between input parameter $Sk$ and output parameter $Sq$ is a particular one among all relationships of parameter pairs.

Keywords
Surface topography, running-in wear, support vector machine

Introduction
As the very first stage of the whole wear process, running-in wear plays an important role in extending the service life of components and improving operation performance of the machine system. Moreover, surface topography is an important feature of wear components. Therefore, surface topography is the best choice for investigation of running-in. However, the traditional belief of independence of surface roughness before and after running-in is widely accepted, which makes it difficult to understand the properties of surface topography after running-in based on unworn surface topography. Nevertheless, researchers have made a lot of effort in this field to study the relevance of surfaces before and after running-in.

A polynomial expression model about correlations of the linear wear rate and the surface roughness change in running-in was proposed. The study investigated the change of cylinder liner surface topography in the early stage of engine life based on two-dimensional (2D) parameters. The surface topographical change during running-in was studied and a theoretical model based on 2D surface height distributions was established. The fractal model of running-in wear was developed, which expressed the evolution of the fractal parameter during running-in. The investigation about the surface roughness evolutions of wear particles and wear components under lubricated rolling wear condition was carried out. The wear behavior of reciprocating sliding friction during running-in and steady wear process was analyzed. A model predicting the running-in performance of the rolling/sliding surfaces subjected to mixed-lubrication line contact was established.

Although impressive progress has been made, there were still two issues that required further research. First, the models based on few characteristic parameters (such as the fractal parameter and the height distribution parameter) were incapable of representing the feature change of complex processes such as...
running-in. Meanwhile, most of them were derived from 2D surface profile and the three-dimensional (3D) information about surface topography was neglected. Second, due to the reduction of concerned friction factors, the simplified analytical models could not efficiently characterize the wear process.

To solve the above issues, first, the comprehensive 3D characterization of surface topography is indispensable; therefore, the areal surface evaluation parameters were adopted. Second, the running-in process was assumed as a black box to avoid the simplification of the model.

Support vector machine (SVM) is a powerful black box method for solving problems in non-linear classification, function estimation and density estimation. SVM has been widely used and made progress in surface analysis and wear research. Three modeling methodologies including SVM were used to investigate the surface in face milling and it was found that the feed has the largest influence on surface roughness. The vision system for strongly reflected metal’s surface defects detection was established based on multi-class SVM. SVM was used to build the mechanical polishing model of predicting surface quality and determining processing parameters, showing great performance in modeling complex process with many factors. The models based on SVM were used to relate the wear rate and technological parameters of the wear-resistant drip moulding. Combining with principal component analysis, SVM was carried out as a predictive model of monitoring tool wear. In sum, SVM is a mature and efficient method for modeling complex systems.

Therefore, the predictive model was established based on SVM. Through the training and validation of the model, the predictive model was established, which was then used to analyze the connection between surface topography before and after running-in.

**Fundamental**

**SVM model of running-in**

As the schema of actual running-in shows in Figure 1, there are too many friction factors (such as load, sliding speed, lubricant kinematic viscosity, temperature and so on) participating in running-in wear. Although the influence of the main friction factors on running-in is predominant, the influence of the other friction factors is not completely negligible to establish the running-in model. Nevertheless, due to the change of the other friction factors during running-in process, it is difficult to establish a reliable running-in model based on them. Considering the stability of the output friction factors after running-in, it is reasonable to model the running-in process as a black box rather than a simplified process. Thus, the model error introduced by ignoring the influence of the other friction factors can be diminished by assuming running-in as a black box model.

As an important feature of wear components, surface topography appears in various stages of running-in: an input feature in the initial stage, changing during the process as interior feature and an output feature after running-in. The information provided by surface topography is very meaningful to the investigation of running-in. Therefore, surface topography can be regarded as the input variables and output variables of the black box model to investigate running-in.

The frequently used approach of modeling a black-box process is the artificial neural network (ANN). Due to the existence of local minima in
the implementation of ANN, its performance is occasionally unsatisfactory. More importantly, because the ANN model requires a large number of data samples and the running-in experiments are time-consuming and complicated, there is difficulty in collecting enough data samples to establish running-in model based on ANN. Therefore, significant data reduction should be achieved without degrading the performance of simulation. Moreover, there exists a local optimum problem in the algorithm of ANN; however, the unique global optimum of the model is indispensable. The introduction of SVM in modeling running-in process can efficiently solve the problems. SVM is a learning method that has specific algorithm that is referred to as kernel method and is a well-known tool for classification and regression tasks. It has properties such as good generalization ability, few adjusting parameters and no requisite for experimentation for the purpose of finding the learning machine architecture.

Based on the above analysis, the regression of running-in process based on surface topography can be simply done by SVM. According to Figure 2, the running-in black box model based on SVM, the input and output of training data set can be collected from actual running-in process. After the model training, to validate the reliability, the model is tested by the input and output of testing data set, which can also come from actual running-in process. Finally, the surface topography after running-in process can be provided by the validated model according to the surface topography before running-in. Since the SVM model indirectly takes the other frictional factors into consideration by the data manipulation of training data from actual running-in process, the running-in black box model based on SVM has better performance than simplified models or common black box models. To increase the adaptability of the prediction model, the main friction factors that do not change significantly during running-in process (including sliding speed, load and lubricant kinematic viscosity) were taken into consideration when modeling the running-in process as input variables.

### SVM regression

ANN evolved from applications and extensive experimentation to the theory, while SVM evolved from the sound theory to the implementation and experiments. The fundamental parts of SVM date back to the discrimination work presented. The main idea is to construct a ‘flattest’ hyperplane in a high-dimensional space within the given data, which can be used for classification or regression of non-linear issues.

The learning of SVM regression is essentially a quadratic programming problem. A set of data \((x_1, y_1), \ldots, (x_k, y_k)\) provides a sample set, where \(x_i \in \mathbb{R}^n\) is the input and \(y_i \in \mathbb{R}\) the target output. Regression is a process to fit a function such as equation (1):

\[
f(x) = (w \cdot x) + b
\]

After training, the corresponding \(y\) can be found from function \(f(x)\) for the \(x\) observations not in the sample. The precision of a \(\varepsilon\)-support vector regression (\(\varepsilon\)-SVR) is established with the Vapnik control algorithm, given a specified tolerance error \(\varepsilon\). If the error of the sample is \(\xi\) and \(|\xi| \leq \varepsilon\), ignore the error loss; otherwise consider the loss as \(|\xi| - \varepsilon\). First, map the sample into a high-dimensional feature space by a non-linear mapping function and transform the problem of this non-linear function estimate into a linear regression problem in high-dimensional feature space. Assuming

---

**Figure 2.** The running-in black box model based on SVM. SVM: support vector machine.
that $\varphi(x)$ is the mapping function that transfers the sample space into a high-dimensional feature space, the problem of solving the parameters of $f(x)$ is transferred into solving the minimization problem equation (2), given the constraints described by equation (3).

$$\min \|w\|^2/2$$  \hspace{1cm} (2)

subject to

$$y_i - [w \cdot \varphi(x)]_i + b_i \leq \varepsilon \quad i = 1, 2, \ldots, l$$

Due to the high dimension of the feature space, the target function is non-differentiable. In general, SVM regression problems are solved by establishing a Lagrangian function that transforms this problem into a dual programming problem to determine the Lagrangian multipliers $\alpha_i$ and $\alpha'_i$, in maximization problems, i.e. equation (4), with the constraints of equation (5)

$$\max \frac{1}{2} \sum_{i,j=1}^{l} (\alpha_i - \alpha'_i)(\alpha_j - \alpha'_j)(x_i, x_j)$$

$$- \varepsilon \sum_{i=1}^{l} (\alpha_i - \alpha'_i) + \sum_{i=1}^{l} \gamma_i(\alpha_i - \alpha'_i)$$  \hspace{1cm} (4)

subject to

$$\sum_{i=1}^{l} (\alpha_i - \alpha'_i) = 0$$

$$\alpha_i, \alpha'_i \in [0, C]$$  \hspace{1cm} (5)

where $\alpha_i$ and $\alpha'_i$ are Lagrangian multipliers, $\alpha_i, \alpha'_i \geq 0$ and $\alpha_i \times \alpha'_i = 0$. Then, the SVM regression problem is transformed into a quadratic programming problem. The regression equation can be obtained by solving this problem. A kernel function $K(x_i, x_j)$ is introduced and the corresponding regression function can be presented as equation (6)

$$f(x) = \sum_{i=1}^{N} (\alpha_i - \alpha'_i)(\alpha_j - \alpha'_j)K(x_i, x_j) + b$$  \hspace{1cm} (6)

where the kernel function $K(x_i, x_j)$ is the internal product of vectors $x_i$ and $x_j$ in feature spaces $\varphi(x_i)$ and $\varphi(x_j)$. Only a small part of $\alpha_i$ and $\alpha'_i$, determined by solving the quadratic programming problem equation (4), is not zero and its corresponding data points are the support vectors. The regression function of SVM is only decided by these support vectors.

**Experimental**

Black box model of running-in process based on SVM regression is available according to the framework of theory mentioned above. The training data set of surface topography parameter of running-in wear under different wear conditions is indispensable for a model predicted under different work conditions. In this article, these data all came from real running-in experiments, which were presented in following sections.

**Apparatus and material**

The establishment of predictive model was based on a series of running-in experiments, which were conducted by pins and disk. Experiments were conducted on a friction and wear testing machine WWM-1, which is shown in Figure 3(a). Load was implemented by weight and lever. In each running-in experiment, driven by motor, three pins slide on disk and the sliding track diameter of the pins was 46 mm.

In order to obtain the information about surface topography, surfaces of pins before and after running-in process were measured by the measurement instrument of surface topography PGI830. It progressively scans surface profiles of measured surface by contacting stylus and sensor. The PGI830 is shown in Figure 3(b) and its vertical resolution is 0.8 nm. The area of measured surfaces was $1 \times 1$ mm, and the sampling intervals of horizontal and vertical directions were 0.01 mm. After surface measurements, the form errors were removed by plane fitting and then the waviness and roughness were separated by Robust Gaussian regression filtering with cutoff wavelength of 0.25 nm. Then, surface topographies of pins were characterized by the areal surface evaluation parameters, which include height

![Figure 3. The schematic diagram of WWM-1 and PGI830.](image-url)
parameters, spatial parameters, hybrid parameters and functional parameters.

Typical metal materials were used in the experiments to simplify the design of experiments. Due to their widespread use in machinery, metal material 1045 and E52100 (ASTM material code) were selected. The pins made of 1045 were treated to a hardness of 34–38 HRC and then machined to 3 mm in diameter and 17 mm in length. The chemical compositions of pins are: 0.43–0.50% C, 0.15–0.35% Si and 0.60–0.90% Mn. The disks made of E52100 were heat treated to a hardness of 59–62 HRC and then machined to 54 mm in diameter and 10 mm in thickness. Its chemical compositions are: 0.95–1.05% C, 0.20–0.40% Mn, 0.15–0.35% Si and 1.30–1.65% Cr. Considering the influence of lubricant kinematic viscosity on surface topography parameter after running-in, three kinds of lubricants with different kinematic viscosity were used, including white mineral oil 7# (10 mm²/s), diesel engine oil SAE 15W-40 (80 mm²/s) and hydraulic oil L-HM 46 (130 mm²/s).

Data acquisition

Since the training data set should be derived from actual running-in process, the speed and the applied load were low in experiments to simulate measurable running-in wear. Training work conditions were obtained by orthogonal design consisting of three factors at three different levels: load (20, 30 and 40 N), sliding speed (20, 40 and 60 r/min) and lubricant kinematic viscosity (10, 80 and 130 mm²/s), generating nine different work conditions. Different sliding times were specified to achieve same sliding distance (240 circles) of all samples under different sliding speed. Moreover, the coefficients of friction of all wear process mentioned in this article have been stable before the process was completed, which meant that all the running-in process were accomplished.

Four types of surface topography of pins were prepared to go through running-in wear under training work conditions. These surface topographies have different characteristics, which are shown in Figure 4. Surface topography in Figure 4(a) has valleys and ridges with single direction, surface with pits and flat top is shown in Figure 4(b), surface in Figure 4(c) is dominated by peaks and multi-direction valleys appear in Figure 4(d).

The combination of 4 types of surface topographies and 9 different work conditions provided a set of 36 different samples as input data set, which generated a set of 36 corresponding surfaces after running-in experiments as output data set. All characteristics of the surface topographies were evaluated by numerical parameters of areal surface evaluation. In sum, the input data include: the parameter evaluation of the unworn surface topographies and three main variables of work condition; the output data is the parameter evaluation of the worn surface topographies after running-in. Then the input data and output data were manipulated to training the predictive model based on SVM.

All surface topography of disks before running-in shared same surface texture feature and surface height as shown in Figure 5(a). Due to the higher hardness, the disks barely wore. The surface topography of disk after running-in is shown in Figure 5(b), which has similar surface texture with unworn surface and confirms the stability of disk surface. The areal surface evaluation of unworn disk and worn disk after running-in was carried out and the comparison of surface parameters is shown in Table 1.

Sa, Sq, Ssk, Sku and Sz represent the height properties of the evaluated surface. Sa and Sq represent the smoothness of the evaluated surface. Ssk and Sku...
represent the symmetry and deviation from bell curve distribution. $Sdq$ and $Sdr$ represent the hybrid properties of evaluated surface, affected by both height and spatial properties of surface. $Sk$, $Spk$ and $Svk$ are important functional parameters that influence the supporting and lubrication properties.

According to the parameter comparison in Table 1, it is found that surface topography of the disk barely changed during the running-in process. Therefore, the influence of surface topography of disks can be ignored.

Table 1. The surface parameters of unworn disk and worn disk.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$Sa$ (µm)</th>
<th>$Sq$ (µm)</th>
<th>$Ssk$</th>
<th>$Sku$</th>
<th>$Sz$ (µm)</th>
<th>$Ssc$ ($1/µm$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unworn surface</td>
<td>0.405</td>
<td>0.51</td>
<td>-0.338</td>
<td>3.07</td>
<td>2.94</td>
<td>0.00592</td>
</tr>
<tr>
<td>Worn surface</td>
<td>0.399</td>
<td>0.494</td>
<td>-0.254</td>
<td>3.17</td>
<td>3.04</td>
<td>0.00561</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$Sds$</th>
<th>$Sdq$</th>
<th>$Sdr$ (%)</th>
<th>$Sk$ (µm)</th>
<th>$Spk$ (µm)</th>
<th>$Svk$ (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unworn surface</td>
<td>613</td>
<td>0.0488</td>
<td>0.12</td>
<td>0.694</td>
<td>0.24</td>
<td>0.299</td>
</tr>
<tr>
<td>Worn surface</td>
<td>609</td>
<td>0.0462</td>
<td>0.107</td>
<td>0.703</td>
<td>0.24</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Results and discussion

The combination of 4 surface topographies and 9 work conditions generated 36 running-in tests, which had different evolution of coefficient of friction (COF). The main patterns of COF evolution in three different speeds is shown in Figure 6, which shows that the COF of all tests had been stable before reaching the total sliding distance and indicates that the running-in process of all tests was completed. Moreover, the percentage of running-in
Areal surface evaluation parameters include 26 parameters from height parameters, spatial parameters, hybrid parameters and functional parameters.\textsuperscript{28,29} Through the one-to-one parameter pairs training based on surface topography, the matrix of one-to-one predictive models was established, as shown in Figure 7, so there were 26 × 26 predictive models established. The symbol ‘\( \rightarrow \)’ in Figure 7 indicates predict, the parameter on the left side of ‘\( \rightarrow \)’ is the unworn surface parameter and the parameter on the right side of ‘\( \rightarrow \)’ is the worn surface parameter.

Based on the matrix of predictive models, the effect of their prediction was tested by actual data obtained from running-in experiments. When the relative error of prediction result from an established model was less than 15%, there was an evident correlation of prediction result from an established model was compared with that of the corresponding input parameters; (c) when the input parameter increased while output parameter decreased, the two parameters are unrelated;\( ^{\text{**}} \) indicates that the two parameters are related;\( ^{\text{!!}} \) indicates the unworn surface before running-in;\( ^{\text{!!}} \) indicates the worn surface after running-in.

The precision of the four predictive models mentioned above was further tested by surface parameters obtained from 10 actual running-in experiments, in which the lubricant kinematic viscosity, load and sliding speed were different. The sliding distance of the testing set was same as that of the training set. The comparison between the relative error of testing results based on SVM and ANN is shown in Figure 8.

The relative error of testing result of four predictive models based on SVM was less than 8%, which indicated good performance of parameter prediction of running-in model based on SVM. Meanwhile, the relative error of prediction result of four predictive models based on ANN was much larger than that of SVM. Therefore, it can be concluded that the application of SVM in modeling the surface topography of running-in is better than that of ANN. Moreover, the correlation between the surface topographies before and after running-in can be confirmed.

To investigate the connection between surface topography before and after running-in based on the validated models, the influence of unworn surface parameters on worn surface parameters after running-in is analyzed. Analyzing samples were predicted by running-in models and the prediction result is shown in Figure 9. The input data of these samples is not displayed in Figure 9, each input parameters increased linearly within the measured range, respectively, generating six samples.

According to the evolution of output parameters with the increase of input parameters shown in Figure 9, it was found that: (a) the influences of input \( S_a \) and \( S_k \) on output \( S_q \) are opposite—output \( S_q \) decreases while input \( S_k \) increases, output \( S_q \) increases while input \( S_a \) increases; (b) outputs \( S_z \), \( S_{dq} \) and \( S_{vk} \), all increase with the increase in their corresponding input parameters; (c) when the input parameters are about less than the mid-value of the measured range in this article, output parameters are
less sensitive to the increase of their corresponding input parameters and (d) according to Figure 9(d), the output $S_{vk}$ does not change with the increase of input $S_k$ when the input value is about less than the mid-value of the measured range.

**Conclusions**

To investigate the connection of surface topography before and after running-in process, a reliable prediction model of surface topography after running-in based on unworn surface topography was established, which indicated that there is a certain correlation between the surface topographies before and after running-in. The analysis of experiment result showed that the hybrid property of surface topographies before and after running-in have strong correlation. Moreover, the influence of input parameter $S_k$ on output parameter $S_q$ is different from the other parameter pairs.

**Funding**

The paper was supported by National Natural Science Foundation of China (Grant nos 51005089 and 51075169).
References


Appendix

Notation

$S_a$ average roughness
$S_{dq}$ the root mean square surface slope
$S_{dr}$ the percentage of additional surface area contributed by the texture as compared to an ideal plane the size of the measurement region.
$S_{ds}$ the density of summits
$S_k$ the core roughness depth
$S_{ku}$ kurtosis
$S_{pk}$ the peak height above the core roughness
$S_q$ root mean square roughness
$S_{sc}$ the mean summit curvature for the various peak structures, and
$S_{sk}$ skewness
$S_{vk}$ the valley depth below the core roughness
$S_z$ the max height of surface