

## Research Article

# SVR Prediction Algorithm for Crack Propagation of Aviation Aluminum Alloy

Jincai Chang <sup>1</sup>, Zhihang Wang <sup>1</sup>, Qingyu Zhu,<sup>2</sup> and Zhao Wang<sup>1</sup>

<sup>1</sup>College of Science, North China University of Science and Technology, Tangshan, Hebei 063210, China

<sup>2</sup>Avic China Aero-Polytechnology Establishment, Beijing 100028, China

Correspondence should be addressed to Zhihang Wang; [wzh8240@126.com](mailto:wzh8240@126.com)

Received 22 July 2020; Revised 18 September 2020; Accepted 3 November 2020; Published 28 November 2020

Academic Editor: Zhaoqing Wang

Copyright © 2020 Jincai Chang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Aluminum alloy material is an important component material in the safe flight of aircraft. It is very important and necessary to predict the fatigue crack growth between holes of aviation aluminum alloy materials. At present, the investigation on the prediction of the cracks between two holes and multiholes is a key problem to be solved. Due to the fact that the fatigue crack growth test of aluminum alloy plate with two or three holes was carried out by the MTS fatigue testing machine, the crack length growth data under different test conditions were obtained. In this paper, support vector regression (SVR) was used to fit the crack data, and the parameters of SVR are optimized by the grid search algorithm at the same time. And then the model of SVR to predict the crack length was established. Discussion on the results shows that the prediction model is effective. Furthermore, the crack growth between three holes was predicted accurately through the model of the crack law between two holes under the same load form.

## 1. Introduction

As we all know, the aluminum alloy material studied in this paper, as an important component material in the aviation field, is widely used in the aviation industry [1] and plays an important role in ensuring the safety of the aviation aircraft flight. Therefore, we should attach great importance to carry out relevant research work on it.

The investigation on fatigue crack growth prediction has a history of several decades. The study of fatigue originated from the embryonic period of the European industrial revolution in the first half of the 19<sup>th</sup> century. As early as the middle of the 19<sup>th</sup> century, Wohler, a German railway engineer, put forward the concept of stress-life ( $S-N$ ) curve and fatigue limit and pointed out the influence degree of factors affecting materials fatigue. After that, some researchers developed Wohler's investigation [2–5] from 1870 to 1890. After decades of development, there are much research on fatigue crack propagation, which is microinvestigation on materials and mechanism analysis by the finite element method. As we all know, fatigue failure of materials needs to

go through three stages: crack initiation, crack stable propagation, and crack instability propagation. The total fatigue life is also composed of these three parts. Generally, the  $(da/dN) - \Delta K$  curve in the double logarithmic coordinate system is used to describe the fatigue crack growth rate.

In Figure 1, zone 1 is a low rate region, zone 2 is a medium rate region, and zone 3 is a high rate region. In the medium rate region (i.e., zone 2),  $(da/dN) - \Delta K$  has a good logarithmic linear relationship, which is the focus of fatigue crack growth research in this paper.

The rapid development of fracture mechanics in the early 20<sup>th</sup> century provides a theoretical basis for the study of fatigue crack growth. In 1961, Paris et al. [6] put forward the Paris formula suitable for the medium rate region on the basis of the fracture mechanics method to express the law of crack propagation, which is the most widely used method in engineering. Besides, Alshamma and Jassim [7] confirmed all parameters of the Paris Law in the experiment and found the crack growth speed through the experiment and analysis. On the basis of the previous studies, the investigation on

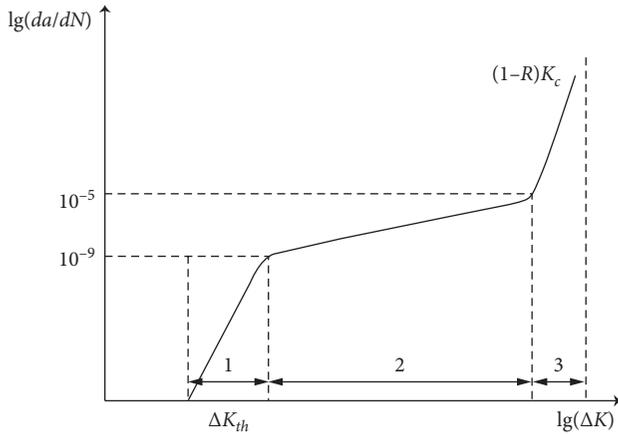


FIGURE 1:  $(da/dN) - \Delta K$  curve.

crack growth and fracture is in the ascendant in recent years [8–12], and the main progress is described as follows: Lei [13] takes 2024-T4 aluminum alloy as the investigation object and uses the finite element method to investigate the influence rule and mechanism of overload condition on the change trend of subsequent fatigue crack growth rate ( $da/dN$ ). For the problem of crack propagation in a plate with a single hole central crack, Zhang and Wang [14] introduced a new Z fracture criterion which can well predict the crack propagating direction of mode I crack in carbon-fiber reinforced composite laminates and proposed new concepts of in-plane average strain, in-plane dilatational strain energy density factor, and reciprocal characteristic function. Furthermore, Hajimohamadi and Ghajar [15] investigated the analytical solution of the stress field and stress intensity in an infinite plane with elliptical holes with an unequal length of prefabricated cracks. For the crack growth problem of two holes plate, Bhargava and Jangid [16] proposed a mathematical model of two unequal-collinear crack growth. Additionally, Singh et al. [17] proposed an improved strip saturation model based on the combination of two internal electric saturation bands and studied two-dimensional (2D) arbitrarily polarized semipermeable dielectric analytically. For the crack growth problem of the porous plate. Zhao et al. [18] applied the basic principle of the complex stress function method and its approximate superposition method to the solution of stress intensity factor of collinear multiple cracks in infinite plates and made a preliminary exploration on the solution of porous cracks. Moreover, Li et al. [19] used the Eshelby inclusion theory and weight function method to give the approximate analytical solution of stress intensity factor of the typical porous multicrock problem and combined with Paris crack growth formula to predict the fatigue crack growth life. Zhu et al. [20–22] conducted an in-depth study on the stress intensity factor and crack opening displacement of interpore crack by using the complex variable function and relevant integral method and then obtained the relevant laws of stress intensity factor and crack opening displacement. Li et al. [23–25] carried out theoretical innovation on the numerical calculation methods used in the fatigue crack growth problem, providing the possibility of

solving such problems more accurately in theory. These investigations on the prediction of fatigue crack growth are based on the empirical formula, analytical method and the finite element method. Based on the existing mechanism investigation and crack growth performance, the analytical expression of crack growth law is obtained. In view of the fact that dealing with the actual aviation crack data which is an urgent proposition to find a way to find a way to get the crack growth law by data-driven method [26], the intelligent algorithm model of aviation fatigue crack growth based on the data-driven method will be built in the focus of this paper.

On the other hand, the support vector regression (SVR) algorithm is an extension of the support vector machine algorithm in regression problem, which was first proposed by Drucker et al. [27]. After years of development, the SVR algorithm is widely used in all aspects of scientific investigation. The application of the SVR algorithm in crack growth and life prediction is as follows: based on grey theory and support vector regression method, Yang et al. [28] proposed GMSVR model and parameter optimization method of the artificial bee colony and applied it to FCG prediction of 7075 aluminum alloy. Furthermore, Song et al. [29] used XFEM and SVR to predict the fatigue life of plate cracks. SVR algorithm has a large number of kernel functions to use, so it can be very flexible to solve various nonlinear regression problems. In addition, the SVR algorithm is based on the principle of structural risk minimization, which avoids overlearning problem and has strong generalization ability. Moreover, it is a convex optimization problem, so the local optimal solution must be the global optimal solution. The SVR algorithm is used to study the crack growth of the aluminum alloy plate. We just use the SVR model established from the data to get the prediction result of crack growth, which avoids the discussion of crack growth mechanism based on the empirical formula. Therefore, it is feasible to apply the SVR algorithm to the crack growth prediction of aluminum alloy plate, which has the advantage of solving such problems.

In order to predict the interhole crack growth length of the aeronautical aluminum alloy plate by data-driven method, it is necessary to collect the crack growth length and the corresponding cycle number under different load spectrum and initial crack conditions. The support vector regression (SVR) [30, 31] algorithm, especially, is a data-driven based method, which is used to fit the crack growth data obtained from the fatigue load test in this paper. Under the same load form, a model which can predict the crack rule between three circular holes by the crack rule between two holes is established, and its fitting effect is tested. The aviation aluminum alloy plate with two or three holes is the test materials of the fatigue load test. First of all, the data of crack length and cycle number of plate crack growth under constant amplitude load and variable amplitude load are collected in this experiment. After that, to get the general rule of crack growth, data exploration, data preprocessing, and so on are carried out. Finally, the applicable support vector regression algorithm is used to fit the crack growth data to get the prediction results of the model and evaluate

the effect of the model. This model can be used in the field of fatigue crack growth prediction of the porous edge of aluminum alloy materials for aircraft. It can predict the crack growth length by the number of cycles and can also predict the number of cycles by the crack growth length through the model. The establishment of the crack prediction model for the aviation aluminum alloy plate can predict three-hole crack rule from two-hole crack and solve the prediction problem of unknown crack growth rule between holes. It is a breakthrough in crack prediction between holes based on a data-driven method. This model not only provides the basis for the prediction of the crack law of aviation aluminum alloy materials but also can be applied to the prediction of safe flight and life of aviation aircraft and guide the prediction of the crack growth law between holes of actual aviation aluminum alloy plate.

## 2. Test Content

**2.1. Test Introduction.** In this section, to investigate the crack growth law of aviation aluminum alloy plate with two holes and three holes under constant amplitude load and variable amplitude load, respectively, is one purpose of this experimental investigation. And, on the other hand, to use the suitable support vector regression algorithm to fit and analyze the experimental data is another purpose of this experimental investigation. Besides, MTS fatigue testing machine system is used in this test, which is used in fatigue test and crack growth test of typical aviation connection structure with multiple cracks. Moreover, the crack growth behavior was measured by a high power optical microscope. Figures 2 and 3 show the MTS fatigue testing machine system.

In addition, the most commonly used aluminum-lithium alloy on the active advanced aircraft is the experimental materials used in this test. In order to explore the propagation law of cracks between holes in the aviation aluminum alloy plate, the aviation aluminum alloy plate is divided into two holes and three holes. Besides, the schematic diagram of the two types of aviation aluminum alloy plates used is shown in Figures 4 and 5.

Besides, the diameters of the small circular holes in the aluminum alloy plate specimens are 4 mm, while the distance between points B and C in the figure and the distance between points D and E are 12 mm, which is shown in Figures 4 and 5. The parameters of the fatigue load test are shown in Table 1.

Points A, B, C, D, E, and F in Figures 4 and 5 preset the initial crack length, respectively, so as to obtain the crack growth in the test.

The aircraft aluminum alloy plate we use has two holes and three holes. For the load, there are two categories of constant amplitude load and variable amplitude load. For the statistics of the initial crack prefabrication of aviation aluminum alloy plates, we divided the constant amplitude and variable amplitude into two categories, respectively, and then distinguished the configuration of two-hole and three-hole cracks in each category; then, there is the following statistical table of crack prefabrication. Among them, there are 11 constant amplitude plate specimens and 9 variable



FIGURE 2: MTS fatigue testing machine system.



FIGURE 3: Test diagram of two-hole plate specimen.

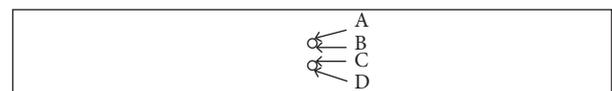


FIGURE 4: Schematic diagram of two-hole plate specimen.

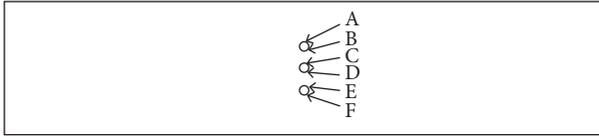


FIGURE 5: Schematic diagram of three-hole plate specimen.

TABLE 1: Table of fatigue load test parameters.

	Load range (kN)	Load median (kN)	Load amplitude (kN)	Median of variable amplitude load (kN)	Amplitude of variable amplitude load (kN)
Constant amplitude load	1.3–13	7.15	5.85	—	—
Variable amplitude load	1.3–13	7.15	5.85	14.3	11.7

TABLE 2: Initial preset cracks of two-hole specimen with constant amplitude.

Serial number	A	B	C	D
01	1	1	0	0
02	1	1.5	0	0
03	1	1	1	1
04	1	1.5	1	1
05	1	1.5	1.5	1
06	0.5	1.5	0.5	0.5
07	0.5	1.5	1.5	0.5
08	0.5	0.5	0.5	0.5
09	0.5	1.5	0	0
10	0.5	1.5	1.5	0.5
11	0.5	0.5	0	0

TABLE 3: Initial preset cracks of three-hole specimen with constant amplitude.

A	B	C	D	E	F
0.5	0.5	1.5	1.5	0.5	0.5

amplitude plate specimens. As shown in Tables 2 and 3, the statistics of crack preset of plate specimen under constant amplitude load (the crack length refers to the length of crack propagation with the initial crack tip as the origin, unit: mm):

It can be seen that Tables 4 and 5 (unit: mm) show the statistics of the preset cracks of the plate specimen under variable amplitude load.

The vertical line of points B and C is taken as the axis of symmetry for two-hole aluminum alloy plate in this paper. It can be seen that point A corresponds to point D and point B corresponds to point C. If the initial crack length of the four points is the same, it is regarded as the case of initial crack

TABLE 4: Initial preset cracks of two-hole specimen with variable amplitude.

Serial number	A	B	C	D
01	0.5	0.5	0	0
02	0.5	1.5	0	0
03	0.5	0.5	0.5	0.5
04	0.5	1.5	0.5	0.5
05	0.5	1.5	1.5	0.5

TABLE 5: Initial preset cracks of three-hole specimen with variable amplitude.

A	B	C	D	E	F
0.5	0.5	1.5	0.5	0.5	0.5

symmetry. Similarly, the vertical line of point C and point D is taken as the axis of symmetry for three-hole aluminum alloy plate. It can be seen that point A corresponds to point F, point B corresponds to point E, and point C corresponds to point D. Whether the initial crack is symmetrical or not will have different results in the study of crack propagation mechanism. In the data exploration stage, the symmetry of the initial crack will directly affect the speed of the subsequent crack growth. Therefore, it is necessary to discuss whether the crack is symmetrical or not, whether it is in the relevant mechanical mechanism or the collected data exploration images. In the case of the above initial crack configuration, the type of the initial crack length at each hole edge of the two-hole aluminum alloy plate is involved in this paper. In this way, the random configuration of the initial crack can be considered. In this paper, the initial crack random configuration here means that the subsequent crack length adopts the cumulative crack length between holes. The initial crack configuration has certain randomness, which weakens the influence of the initial crack length configuration between holes on the crack growth law, so more attention is paid to the growth law of the crack accumulation length. With such a random initial crack configuration, a general prediction model of inter-pore crack length under the same load spectrum (considering only the cumulative crack length) can be established. It can provide the basis for the follow-up investigation work through the above analysis.

**2.2. Test Process.** Firstly, the abovementioned constant amplitude fatigue load spectrum and variable amplitude fatigue load spectrum (see Table 1) which are suitable for laboratory use are adopted. Secondly, the fatigue crack growth test of porous plate specimen under constant amplitude load spectrum is carried out. The fatigue crack growth data of multiple cracks are collected, and the growth rule is analyzed. Finally, the fatigue crack growth test of plate specimen with porous edge cracks under variable amplitude load is carried out. Therefore, the data of crack length and number of cycles are collected to provide test data for the

prediction of crack length. According to the previous solution, it also can even establish the life of fatigue multicroack propagation law.

**2.3. Experimental Data.** In this section, the test data of crack growth length and cycle times of 20 groups of aluminum alloy specimens are obtained, which are saved in 20 Excel files, respectively, through the above fatigue loading test. Some original data of crack growth for the abovementioned aluminum alloy specimens are presented in this paper. The initial cracks at points A, B, C, and D in the constant amplitude two-hole test specimen 03 are all 1.0 mm test data, some of which are shown in Table 6 (the data in the first column of the table is the number of cycles, and the rest is in mm).

The initial crack configuration of the constant amplitude three-hole specimen is as follows: the initial crack at points A, B, E, and F is 0.5 mm and the initial crack at point C and point D is 1.5 mm. Some of the data are shown in Table 7 (the data in the first column of the table is the number of cycles, and the rest is in mm).

For the sample with variable amplitude, it should be clear that the effect of retardation after the application of overload will occur. The intuitionistic crack growth image of variable amplitude specimen is obviously different from that under constant amplitude load, which will experience a long slow growth period, thus prolonging the life of the component, while the crack growth rate under constant amplitude load will continue to increase, and its life is shorter than that under variable amplitude load. These features are visually shown in Figures 6–9.

However, the initial cracks at points A, B, C, and D in the two-hole variable amplitude test specimen 03 (complement) are all 0.5 mm test data, some of which are shown in Table 8 (the data in the first column of the table is the number of cycles, and the rest is in mm).

The initial crack configuration of the variable amplitude three-hole specimen is as follows: the initial crack at points A, B, D, E, and F is 0.5 mm and the initial crack at point C is 1.5 mm. Some of the data are shown in Table 9 (the data in the first column of the table is the number of cycles, and the rest is in mm).

### 3. The Theory of Support Vector Regression

As we know, support vector regression (SVR) is a development of the support vector machine algorithm, which was first extended to support vector regression by Drucker et al. [27]. Support vector regression is a small specimen learning method with a solid theoretical basis. And the traditional process from induction to deduction can be avoided by it. By this means, SVR simplified greatly the usual regression problem. The support vector regression algorithm avoids “dimension disaster” in a sense and has good robustness. The main principles of SVR are as follows.

Given training samples  $T = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_m, y_m)\}$ ,  $y_i \in \mathbb{R}$ . The regression model shaped as  $(f(\mathbf{x}) =$

TABLE 6: Partial test data of constant amplitude two-hole test specimen 03.

	A	B	C	D
13000	0.6	0.9	0.9	0.9
13200	0.6	0.9	0.9	1
13400	0.7	1	1.1	1
13600	0.8	1	1.1	1.05
13800	0.8	1.2	1.2	1.05
		...		
18900	5	4.7	5	5.6
18925	5.2	4.7	5	5.9
18950	5.5	4.7	5	6.2
18975	5.7	4.7	5	6.6
19000	6	4.7	5	7

TABLE 7: Partial test data of constant amplitude three-hole test specimen.

	A	B	C	D	E	F
5000	0	0	0	0	0	0
28000	0	0	0	0	0	0
28500	0	0	0	0.1	0	0
29000	0	0	0	0.2	0	0
29500	0	0	0	0.3	0	0
			...			
64250	3.4	2	8	7.2	2.8	6.6
64500	3.7	2	8	7.2	2.8	7.5
64750	4	2	8	7.2	2.8	7.6
64973	6.2	2	8	7.2	2.8	10.2
64976	17.2	2	8	7.2	2.8	16.8

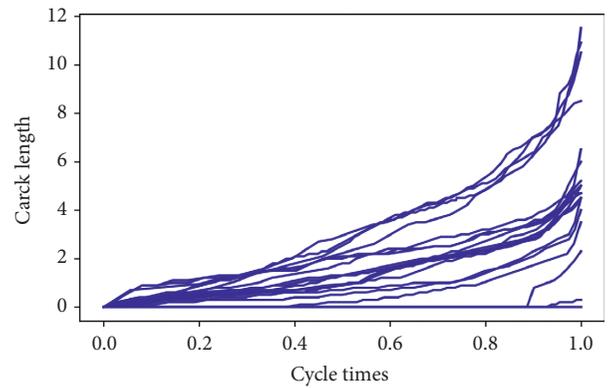


FIGURE 6: Normalized data of two holes with constant amplitude.

$w^T \mathbf{x} + b)$  is the learning objective function. Moreover, the model parameters which make  $f(\mathbf{x})$  as close as possible to  $y$ ,  $w$ , and  $b$  are to be determined.

Next, support vector regression [30] assumes that we can tolerate the maximum deviation of  $\epsilon$  between  $f(\mathbf{x})$  and  $y$ . That is to say, the loss is only calculated when the absolute value of the difference between  $f(\mathbf{x})$  and  $y$  is greater than  $(\epsilon)$ . As shown in Figure 10, this is equivalent to building a spacing band (i.e., the part sandwiched between the two

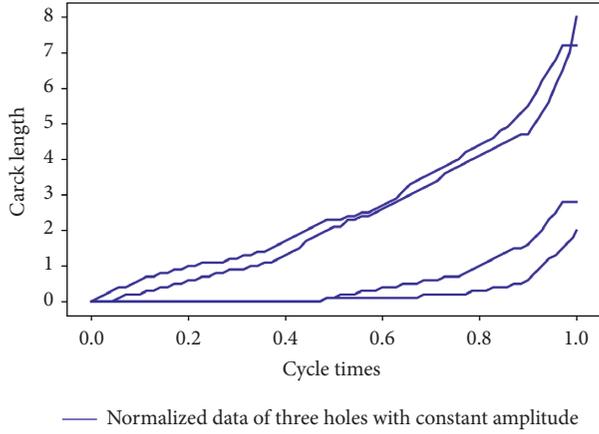


FIGURE 7: Normalized data of three holes with constant amplitude.

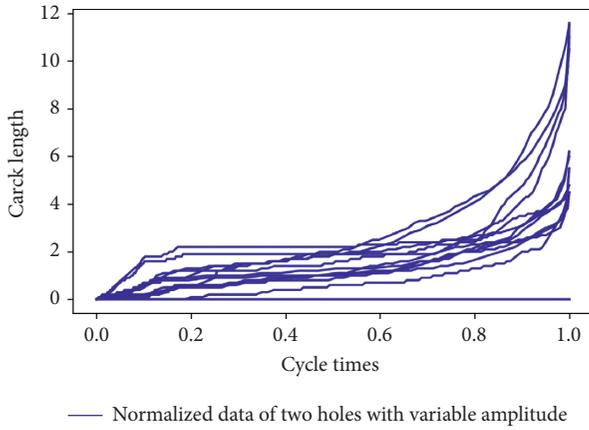


FIGURE 8: Normalized data of two holes with variable amplitude.

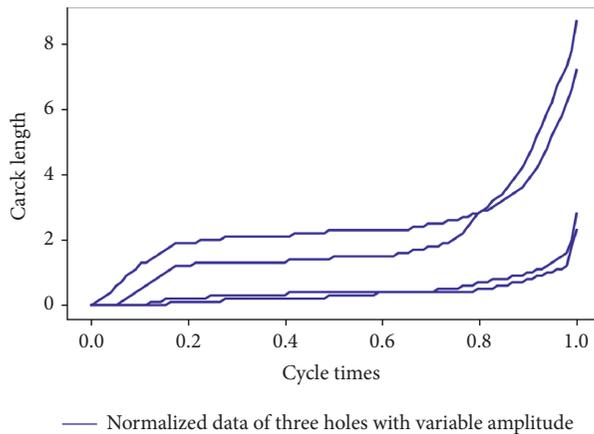


FIGURE 9: Normalized data of three holes with variable amplitude.

dashed lines in the figure) with a width of  $(\varepsilon)$  centered on  $f(\mathbf{x})$  (i.e., the solid line in the figure). On the contrary, if the training samples fall into this interval, they are considered to be correctly predicted.

Therefore, the SVR problem can be formalized as

TABLE 8: Partial test data of variable amplitude two-hole test specimen 03 (complement).

	A	B	C	D	
11000	0	0.1	0.1	0.1	
11500	0	0.1	0.1	0.1	
12000	0.1	0.1	0.1	0.1	
12500	0.1	0.1	0.1	0.1	
13000	0.1	0.2	0.2	0.1	
		...			
15750	0.5	0.5	0.6	0.6	
16000	0.6	0.5	0.6	0.6	Added a variant
16500	0.6	0.5	0.6	0.6	
		.....			
42100	6.3	4.8	6.2	7.2	
42200	6.5	4.8	6.2	7.5	
42300	7	4.8	6.2	7.8	
42400	8	4.8	6.2	9	
42448	17.1	4.8	6.2	18.2	

TABLE 9: Partial test data of variable amplitude three-hole test specimen.

	A	B	C	D	E	F	
5000	0	0	0	0	0	0	
10000	0	0	0	0	0	0	
15000	0	0	0	0	0	0	
20000	0	0	0	0	0	0	
25000	0	0	0	0	0	0	
		...					
46000	0	0.1	1.8	1.1	0.2	0	
47000	0	0.1	1.9	1.2	0.2	0	Added a variant
48000	0	0.1	1.9	1.2	0.2	0	
		...					
132500	6.3	2.8	7.2	8.7	2.3	5.3	
133000	8	2.8	7.2	8.7	2.3	5.9	
133250	9.1	2.8	7.2	8.7	2.3	6.7	
133500	10.5	2.8	7.2	8.7	2.3	7.7	
133532	17.3	2.8	7.2	8.7	2.3	18	

$$\min_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}\|^2 + b, \quad (1)$$

where  $b = C \sum_{i=1}^m l_{\varepsilon}(f(\mathbf{x}_i) - y_i)$  and  $C$  is the regularization constant. It can be seen from Figure 11 that  $l_{\varepsilon}$  is the  $\varepsilon$ -insensitive loss function:

$$l_{\varepsilon}(z) = \begin{cases} 0, & \text{if } |z| \leq \varepsilon, \\ |z| - \varepsilon, & \text{otherwise.} \end{cases} \quad (2)$$

By introducing relaxation variables  $\xi_i$  and  $\tilde{\xi}_i$ , (1) can be rewritten as

$$\min_{\mathbf{w}, b, \xi_i, \tilde{\xi}_i} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^m (\xi_i + \tilde{\xi}_i),$$

$$f(\mathbf{x}_i) - y_i \leq \varepsilon + \xi_i, \quad (3)$$

$$\text{s.t. } y_i - f(\mathbf{x}_i) \leq \varepsilon + \tilde{\xi}_i,$$

$$\xi_i \geq 0, \tilde{\xi}_i \geq 0, \quad i = 1, 2, \dots, m.$$

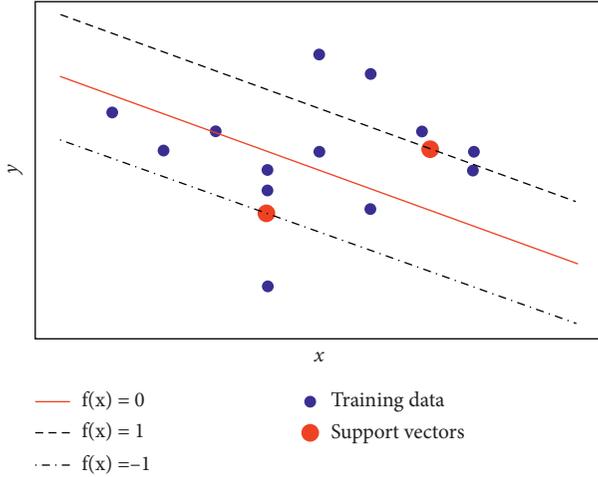
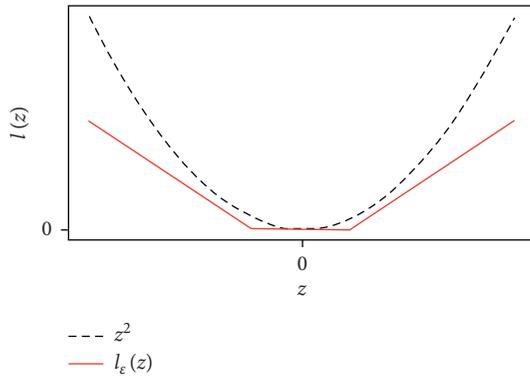


FIGURE 10: Schematic diagram of support vector regression.

FIGURE 11:  $(\varepsilon)$ -insensitive loss function.

Therefore, the solution of SVR can be obtained by the Lagrange multiplier method as follows:

$$f(\mathbf{x}) = \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) \mathbf{x}_i^T \mathbf{x} + b. \quad (4)$$

Furthermore, the samples of  $(\hat{\alpha}_i - \alpha_i) \neq 0$  in equation (4) are the support vectors of SVR, which must fall outside the  $\varepsilon$ -interval. Obviously, support vectors in SVR are only a part of the training samples; that is, its solution is still sparse.

If the feature mapping form  $f(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x}) + b$  is considered, then the corresponding solution can be obtained by the Lagrange multiplier method:

$$\mathbf{w} = \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) \phi(\mathbf{x}_i). \quad (5)$$

Substituting equation (5) into  $f(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x}) + b$ , SVR can be expressed as

$$f(\mathbf{x}) = \sum_{i=1}^m (\hat{\alpha}_i - \alpha_i) \kappa(\mathbf{x}, \mathbf{x}_i) + b, \quad (6)$$

where  $\kappa(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j)$  represent the kernel function. The kernel function used in this paper is Gauss kernel function, whose expression is  $\kappa(\mathbf{x}_i, \mathbf{x}_j) = \exp(-(\|\mathbf{x}_i - \mathbf{x}_j\|^2 / 2\sigma^2))$ . Where  $\sigma > 0$  stand for the width of the Gaussian kernel.

## 4. Prediction of Crack Length by SVR

**4.1. Data Standardization.** As we all know, the input data to be normalized is required in the SVR algorithm. Therefore, it is necessary to normalize the data to be processed. The data to be used needs to be determined before normalization. According to the data exploration in the early stage, we can know from the visual graph that the crack length data is in the state of nonexpansion when the number of cycles is small. That is, a large number of zero values in the original data have no impact on the prediction results, so it can be discarded. After a certain length that the crack length reaches, the crack length remains unchanged when the number of cycles in the later period is large. Therefore, this part of the data can be discarded and only part of the data with changing rules can be retained. The influence of point A and point D of the two-hole plate and point A and point F of three-hole plate on interpore crack propagation is ignored in this test. Therefore, the following investigating contents only consider the crack growth between holes. We also add the initial crack length of each crack to each row of crack data, so that the crack data has a certain physical significance.

For convenience, standardization methods in data standardization are as follows:

Transform sequence  $(x_1, x_2, \dots, x_n)$   $x_i^* = ((x_i - \min_{1 \leq j \leq n} x_j) / (\max_{1 \leq j \leq n} x_j - \min_{1 \leq j \leq n} x_j))$ , the new sequence  $(x_1^*, x_2^*, \dots, x_n^* \in [0, 1])$  is dimensionless. The data can be standardized first for data preprocessing.

Using the Matplotlib Library in Python to draw the image and get the relevant statistics for each test specimen. Data preprocessing images between holes cracks of the above specimen are presented in this paper. Each image drawn below does not consider the zero-value data in the front part of the data file, and the number of cycles is processed by the above normalization method. The abscissa is the number of cycles after normalization and the ordinate is the crack length in mm, as expected in Figures 6–9 for the specific image.

After the above-standardized treatment, the crack length is treated as follows: the crack data of point B and point C are added to get the cumulative crack length, which is recorded as  $I_{BC}$ , and the cumulative crack length is about 12 mm for the two-hole crack specimen. Meanwhile, the crack data of points B, C, D, and E are added and record as  $I_{BCDE}$  to get the cumulative crack length, which is about 25 mm for the three-hole crack specimen.

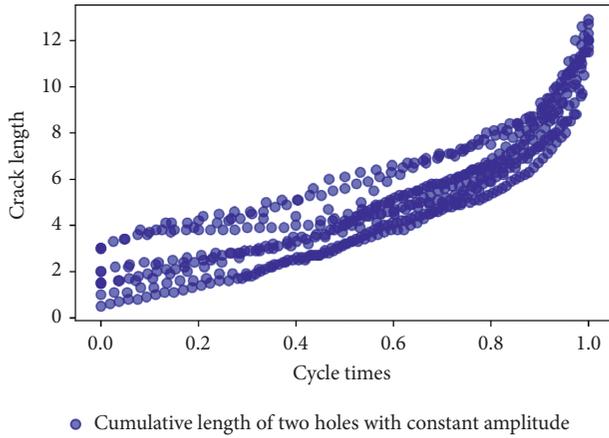


FIGURE 12: Cumulative length of two holes with constant amplitude.

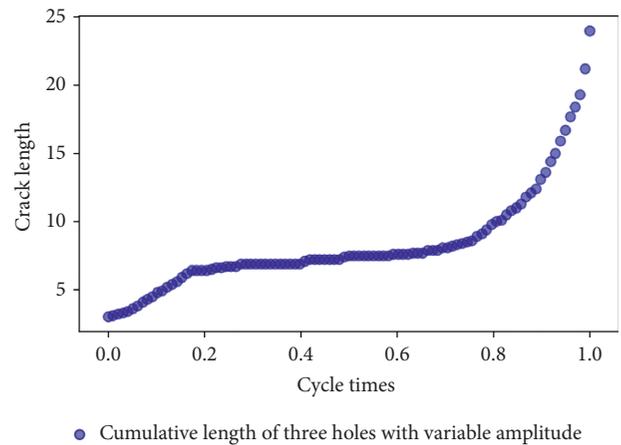


FIGURE 15: Cumulative length of three holes with variable amplitude.

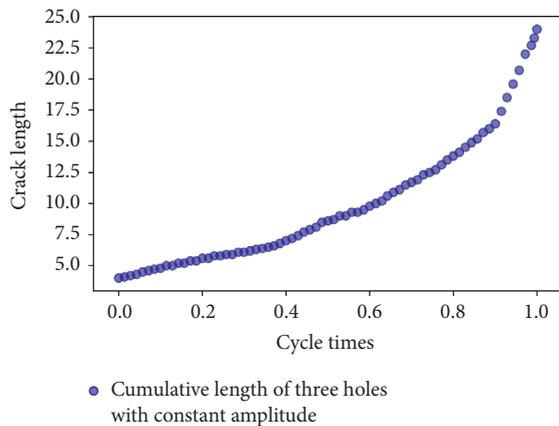


FIGURE 13: Cumulative length of three holes with constant amplitude.

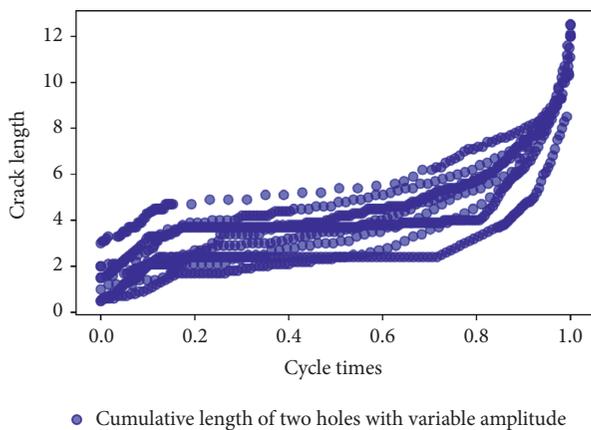


FIGURE 14: Cumulative length of two holes with variable amplitude.

After the abovementioned processing, the preprocessing image of crack growth data of each test piece is obtained. The abscissa is the number of cycles, and the ordinate is the sum of crack length between holes  $I_{BC}$  or  $I_{BCDE}$  data whose unit

is mm in all the following figures. See Figures 12–15 for the specific image.

4.2. Prediction Steps of SVR. Obviously, the steps of using support vector regression to predict the crack length are as follows: first, establish the support vector regression model for the normalized  $I_{BC}$  crack data of the constant amplitude two-hole specimen, obtain the training and testing errors, and then test the effect of the model on the constant amplitude two-hole specimen. Then, the support vector regression model was established and predicted for  $I_{BCDE}$  crack data of normalized constant amplitude three-hole specimen, observe the error and deviation, and verify the prediction effect of the model. Finally, the prediction model is established for the variable amplitude specimen similar to the above prediction steps.

The SVR algorithm in scikit-learn is used to build the SVR model for crack prediction of aviation aluminum alloy plate by adjusting the parameters such as kernel (i.e., specify the kernel type to be used in the algorithm), gamma (i.e., for kernel coefficient), and C (i.e., for penalty parameter C of error term) in this paper.

Furthermore, this paper adopts the grid search optimization algorithm when adjusting the parameters. Grid search optimization algorithm is a kind of exhaustive search algorithm which optimizes the parameters by cross-validation and then obtains the optimal learning algorithm, that is to say, to traverse and search the corresponding permutation and combination values of each parameter in a given range, and each group of combination results constitutes a “grid.” Then, each combination is applied to the training learning algorithm, and the evaluation results are obtained by cross-validation. After traversing and training all the parameter combinations, the grid search algorithm will automatically return the best parameter combination with the highest score, and its corresponding learning model is the optimal regression model.

It is known that the kernel, gamma, and C parameters need to be adjusted in SVR. For the kernel, that is, the type of

the kernel function, we specify using the radial basis kernel function. For gamma and C parameters, we use the GridSearchCV method in scikit-learn to optimize the grid search parameters and get the optimal parameters in this paper, where the value range of gamma is set as {0.01, 0.1, 1, 10, 100}, that is, ( $\gamma \in \{0.01, 0.1, 1, 10, 100\}$ ), and the value range of C is set as {1, 10, 100, 1000}, that is  $C \in \{1, 10, 100, 1000\}$ . According to the range of gamma and C from the above, there are 20 grid points in this method.

**4.3. Evaluating Indicator.** Data set partition standard: for the division of data sets under different load spectra, this paper uses two-hole data under the corresponding load spectrum as training data and three-hole data as test data.

The commonly used cross-validation method in machine learning is used for the aspect of model evaluation in data set partition. In view of the fact that the half-fold cross-validation method [32] is specifically used in this paper, the reliable evaluation of the model effect is obtained by making full use of the data set, so as to verify the performance of the established SVR model for crack prediction of aviation aluminum alloy plate.

Additionally, error analysis indicators, mean square error [33], root mean square error, [34] and other error analysis indicators, are commonly used in regression problems. Variance, standard deviation, mean square error, and root mean square error are the error evaluation indicators used in this paper. The calculation formulas of each indicator are as follows:

Variance ( $S^2$ ):

$$S^2 = \frac{1}{m} \sum_{i=1}^m (y_i - \bar{y})^2. \quad (7)$$

Standard deviation (S):

$$S = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \bar{y})^2}. \quad (8)$$

Mean square error (MSE):

$$\text{MSE} = \frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2. \quad (9)$$

Root mean square error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2}, \quad (10)$$

where  $m$ ,  $y_i$ ,  $\hat{y}_i$ , and  $\bar{y}$  are the total number of samples, the predicted value of the model, the real value, and the sample mean value, respectively.

It is significant to compare variance and standard deviation with MSE and RMSE. For example, if the MSE of prediction error is almost equal to the variance of the target (or RMSE is almost equal to the standard deviation of the target), this indicates that the prediction algorithm is not

effective. By simply averaging the target value to replace the prediction algorithm, almost the same effect can be achieved. If the prediction error RMSE is about half of the actual target standard deviation, this is already pretty good performance. Because the variance and MSE comparison and the standard deviation and RMSE comparison have a certain correlation, thus, the comparison between standard deviation and RMSE is the main error indicator used in this paper.

**4.4. Result Analysis.** According to the previous solution, this paper presents a visual image of the results of predicting the crack length of three-hole aluminum alloy plate specimens under the same load spectrum. In the visualization image, Figures 16 and 17 are the images of test data and prediction data. Besides, Figures 18 and 19 are the learning curve images. In the images in Figures 16 and 17, the red lines are the line graph of test data, and the blue lines are the line graph of support vector regression prediction data.

However, through the optimization of SVR parameters by the grid search algorithm, we get the optimal SVR model for aluminum alloy plate under constant amplitude. The parameters of it are as follows: SVR ( $C=100.0$ ,  $\text{cache\_size}=200$ ,  $\text{coef0}=0.0$ ,  $\text{degree}=3$ ,  $\text{epsilon}=0.1$ ,  $\text{gamma}=10.0$ ,  $\text{kernel}=\text{"RBF"}$ ,  $\text{max\_iter}=-1$ ,  $\text{shrinking}=-\text{true}$ ,  $\text{tol}=0.001$ ,  $\text{verbose}=\text{false}$ ). Among them,  $\text{kernel}=\text{"RBF"}$  is the setting parameter,  $C=100.0$  and  $\text{gamma}=10.0$  are the optimization parameters of the grid search algorithm. And other parameters are the default values. A simple interpretation can be given as follows.

Figures 16 and 18 show the results of constant amplitude three-hole specimen.

The standard deviation is about 10.11 and the RMSE of this model is about 0.64 which is far less than the standard deviation. It is easily found that the effect of this model is very well. The results of half-fold cross-validation are 0.99942329, 0.99940067, 0.9994369, 0.9991129, 0.9994331, and the mean value is 0.9993613737263342, which also shows that the prediction effect of the model is good. In this way, the model can be used for this problem.

Moreover, through the optimization of SVR parameters by the grid search algorithm, we get the optimal SVR model for aluminum alloy plate under variable amplitude. The parameters of it are as follows: SVR ( $C=1000.0$ ,  $\text{cache\_size}=200$ ,  $\text{coef0}=0.0$ ,  $\text{degree}=3$ ,  $\text{epsilon}=0.1$ ,  $\text{gamma}=1.0$ ,  $\text{kernel}=\text{"RBF"}$ ,  $\text{max\_iter}=-1$ ,  $\text{shrinking}=-\text{true}$ ,  $\text{tol}=0.001$ ,  $\text{verbose}=\text{false}$ ). Among them,  $\text{kernel}=\text{"RBF"}$  is the setting parameter, and  $C=1000.0$  and  $\text{gamma}=1.0$  are the optimization parameters of the grid search algorithm. And other parameters are the default values. A simple interpretation can be given as follows.

Figures 17 and 19 show the results of variable amplitude three-hole specimen:

The standard deviation is about 9.56, and the RMSE of this model is about 0.86 which is far less than the standard deviation. It is easily found that the effect of this model is very well. The results of half-fold cross-validation are 0.99930434, 0.99930836, 0.99873057, 0.999125, and 0.9991911, with the mean value of

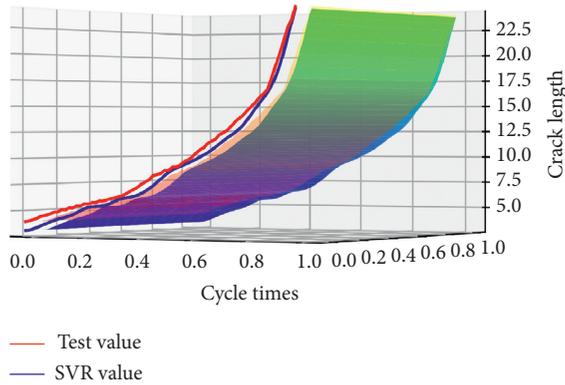


FIGURE 16: Prediction of crack length with constant amplitude by SVR.

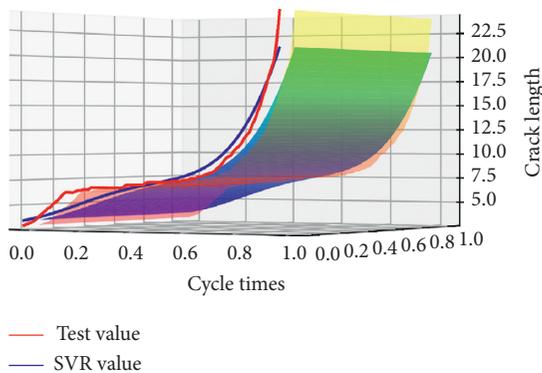


FIGURE 17: Prediction of crack length with variable amplitude by SVR.

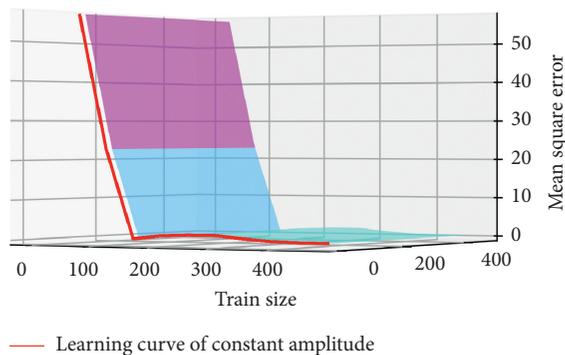


FIGURE 18: Learning curve of constant amplitude.

0.9991318738985194, which also shows that the prediction effect of the model is good. In this way, the model can be used for this problem.

Additionally, through the establishment, implementation, and evaluation of the above models, it can be seen that the constant amplitude data and the variable amplitude data are two different types of models. The crack propagation law between the holes of the two-hole aluminum alloy plate and three-hole aluminum alloy plate has roughly the same trend

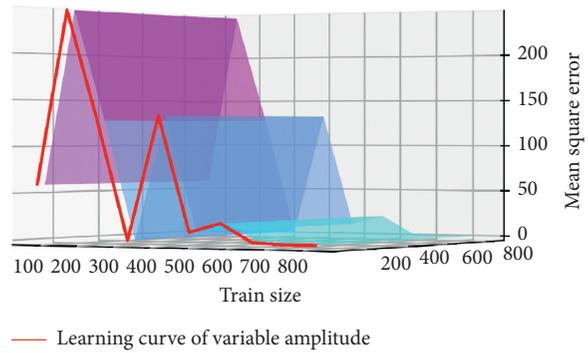


FIGURE 19: Learning curve of variable amplitude.

in each type of model. Therefore, three-hole crack data can be predicted by two-hole crack data under the same load form. In practical application, we can refer to the crack configuration form of the specimen in this paper and find a crack configuration form model similar to the actual situation. In addition, we can get the crack growth prediction results through the model. In addition, in the actual aviation aluminum alloy interhole crack prediction problem, under the same load spectrum situation, only two-hole plate test is needed to predict the corresponding three-hole plate crack propagation through the established SVR model. And the prediction accuracy of this model is high so that it can effectively improve the efficiency of solving such problems and reduce various costs.

### 5. Conclusion

The SVR model can be used to predict the crack between holes in similar aluminum alloy plates. Using different models under different conditions can make the prediction more accurate: use the model with a higher score in constant amplitude model under constant amplitude condition and use the model with a higher score in variable amplitude model under variable amplitude condition. According to the results, we can know that two-hole and three-hole aluminum alloy plate have roughly the same crack propagation law under the same load spectrum. It leads us to predict the crack law between three holes with constant amplitude by the crack law between two holes with constant amplitude. And at the same time, we also can predict the crack law between three holes with variable amplitude by the crack law between two holes with variable amplitude. Alternatively, the data-driven SVR algorithm model for crack growth prediction is a useful supplement to the existing methods for crack growth prediction in this paper. In the practical prediction of interpore crack growth, the results of the method in this paper and the existing methods can be considered comprehensively, and the more reliable prediction results can be obtained through mutual verification. The accuracy of the model can meet the accuracy requirements of the aviation crack growth problem. In view of this, the established model can be modified into the aviation crack growth prediction method to a certain extent and can

guide the actual prediction of the law of the aviation cracks between holes.

### Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

The support from the National Natural Science Foundation of China (nos. 51674121 and 61702184), the Tangshan Team Project (no. 18130209B), and Aviation Fund Project (no. 2017ZD41006) is gratefully acknowledged.

### References

- [1] M. Yang, Y. Bi, J. Ren, H. Wang, and W. Jiang, "Study on propagation life of fretting fatigue crack of aerial aluminum alloy," *Equipment Environmental Engineering*, vol. 9, no. 5, pp. 1–5, 2012.
- [2] Z. Zheng and Z. Wang, "Historical review of fatigue study," *Journal Materials Science and Engineering*, vol. 2, pp. 1–6, 1993.
- [3] P. C. Paris, M. P. Gomez, and W. P. Anderson, "A rational analytic theory of fatigue," *The Trend in Engineering*, vol. 13, pp. 9–14, 1961.
- [4] N. Thompson, N. Wadsworth, and N. Louat, "The origin of fatigue fracture in copper," *Philosophical Magazine*, vol. 1, no. 2, pp. 113–126, 1956.
- [5] X. Yang and C. Chen, *Fatigue and Fracture*, pp. 152–182, Huazhong University of Science and Technology Press, Wuhan, China, 2nd edition, 2018.
- [6] X. Ni, X. Li, and X. Wang, "General modification and application of the Paris law for fatigue crack propagation," *Pressure Vessel Technology*, vol. 23, no. 12, pp. 8–15, 2006.
- [7] F. A. Alshamma and O. A. Jassim, "Dynamic crack propagation in nano-composite thin plates under multi-axial cyclic loading," *Journal of Materials Research and Technology*, vol. 8, no. 5, pp. 4672–4681, 2019.
- [8] H. Liu and Z. Zhou, "Dynamic behavior of rectangular crack in three-dimensional orthotropic elastic medium by means of non-local theory," *Applied Mathematics and Mechanics*, vol. 38, no. 2, pp. 173–190, 2017.
- [9] Z. Zhuravlova, "Stress analysis near the tips of a transverse crack in an elastic semi-strip," *Applied Mathematics and Mechanics*, vol. 38, no. 7, pp. 935–956, 2017.
- [10] M. Karimi, A. Ghassemi, A. Atrian, and M. Vahabi, "Compensation of stress intensity factors in hollow cylinders containing several cracks under torsion by electro-elastic coating," *Applied Mathematics and Mechanics*, vol. 40, no. 9, pp. 1335–1360, 2019.
- [11] X. Zhao, Z. Qian, J. Liu, and C. Gao, "Effects of electric/magnetic impact on the transient fracture of interface crack in piezoelectric-piezomagnetic sandwich structure: anti-plane case," *Applied Mathematics and Mechanics*, vol. 41, no. 1, pp. 139–156, 2020.
- [12] Xu Wang and P. Schiavone, "Elastic field near the tip of an anticrack in a decagonal quasicrystalline material," *Applied Mathematics and Mechanics*, vol. 41, no. 3, pp. 401–408, 2020.
- [13] B. Lei, "Study on fatigue crack growth behavior and fracture mechanism of Al-alloy subjected to tensile overloads", M.S. thesis, Chang'an University, Xi'an, China, 2018.
- [14] S. Zhang and W. Yang, "Prediction of mode I crack propagation direction in carbon-fiber reinforced composite plate," *Applied Mathematics and Mechanics (English Edition)*, vol. 25, no. 6, pp. 714–722, 2004.
- [15] M. Hajimohamadi and R. Ghajar, "An analytical solution for the stress field and stress intensity factor in an infinite plane containing an elliptical hole with two unequal aligned cracks," *Applied Mathematics and Mechanics*, vol. 39, no. 8, pp. 1103–1118, 2018.
- [16] R. R. Bhargava and K. Jangid, "Strip-coalesced interior zone model for two unequal collinear cracks weakening piezoelectric media," *Applied Mathematics and Mechanics*, vol. 35, no. 10, pp. 1249–1260, 2014.
- [17] S. Singh, K. Sharma, and R. R. Bhargava, "Modified strip saturated models for two equal collinear cracks with coalesced zones in piezoelectric media," *Applied Mathematics and Mechanics*, vol. 40, no. 8, pp. 1097–1118, 2019.
- [18] J. Zhao, L. Xie, and J. Liu, "Calculation of stress intensity factor for plate containing a number of symmetrical collinear holes," *Machinery Design and Manufacture*, vol. 12, pp. 12–14, 2010.
- [19] Z. Li, Wu Xu, X. Zhang et al., "Experimental and analytical analyses of fatigue crack growth in sheets with multiple holes and cracks," *Acta Aeronautica et Astronautica Sinica*, vol. 39, no. 7, pp. 154–162, 2018.
- [20] Q. Zhu, Bo Song, and Z. Zeng, "Research of opening displacements for two cracks emanating from a circular hole in an infinite plate under pressure on part of crack surface," *Journal of Mechanical Strength*, vol. 41, no. 3, pp. 713–717, 2019.
- [21] Q. Zhu and R. Zhou, "Complex variable function solutions of mixed-mode stress intensity factors for two cracks of arbitrary length emanating from an elliptical hole," *Journal of Mechanical Strength*, vol. 41, no. 4, pp. 983–987, 2019.
- [22] Q. Zhu and J. Jiang, "Complex variable function solutions of stress intensity factors for multiple hole-edge cracks," *Journal of Mechanical Strength*, vol. 42, no. 2, pp. 437–442, 2020.
- [23] J. Li, Y. Cheng, and Z. Li, "Superconvergence of the composite rectangle rule for computing hypersingular integral on interval," *Numerical Mathematics: Theory, Methods and Applications*, vol. 13, no. 3, pp. 770–787, 2020.
- [24] Li Jin and Y. Cheng, "Linear barycentric rational collocation method for solving second-order Volterra integro-differential equation," *Journal of Computational and Applied Mathematics*, vol. 39, pp. 1–9, 2020.
- [25] Li Jin and Y. Cheng, "Linear barycentric rational collocation method for solving heat conduction equation," *Numerical Methods for Partial Differential Equations*, pp. 1–13, 2020.
- [26] J. Chang, Q. Pan, Z. Shen, and H. Qin, "Intelligent prediction of refrigerant amounts based on internet of things," *Complexity*, vol. 2020, Article ID 1743973, 12 pages, 2020.
- [27] H. Drucker, C. J. C. Burges, L. Kaufman et al., "Support vector regression machines," *Advances in Neural Information Processing Systems*, Vol. 9, MIT Press, Cambridge, MA, USA, 1996.
- [28] D. Yang, Y. Liu, S. Li, J. Tao, C. Liu, and J. Yi, "Fatigue crack growth prediction of 7075 aluminum alloy based on the GMSVR model optimized by the artificial bee colony

- algorithm,” *Engineering Computations*, vol. 34, no. 4, pp. 1034–1053, 2017.
- [29] W. Song, Z. Jiang, and H. Jiang, “Predict the fatigue life of crack based on extended finite element method and SVR,” in *Proceedings of the 6th International Conference on Computer-Aided Design, Manufacturing, Modeling and Simulation (CDMMS 2018)*, American Institute of Physics, Busan, South Korea, April 2018.
- [30] Z. Zhou, *Machine Learning*, pp. 121–140, Tsinghua University Press, Beijing, China, 1st edition, 2016.
- [31] H. Li, *Statistical Learning Methods*, pp. 111–154, Tsinghua University Press, Beijing, China, 2nd edition, 2019.
- [32] T. Braun, S. Spiliopoulos, C. Veltman et al., “Detection of myocardial ischemia due to clinically asymptomatic coronary artery stenosis at rest using supervised artificial intelligence-enabled vectorcardiography—a five-fold cross validation of accuracy,” *Journal of Electrocardiology*, vol. 59, pp. 100–105, 2020.
- [33] H. Yao, J. Zhou, L. Lin et al., “Solar radiation prediction method using bat algorithm optimized SVR,” *Renewable Energy Resources*, vol. 36, no. 11, pp. 1612–1617, 2018.
- [34] Z. Shang, “Research on the boarding passengers prediction algorithm of subway trains based on SVR,” *Modern Computer (Professional)*, vol. 7, pp. 42–44, 2019.