

1
2 **Surface Roughness Characterization for Stress**
3 **Concentration Factor Predictions: A Bayesian**
4 **Learning Approach**

5
6 **Jingyi ZHANG ***

7
8 (Telephone number: +86 13572558759, E-mail: echo4869@stu.xjtu.edu.cn)

9 ¹ * Department of Engineering Mechanics, SVL, Xi'an Jiaotong University, Xi'an ,710049,
10 China

11
12
13
14 **ABSTRACT**

15
16 The surface roughness has an important influence on the fatigue life of the structures. The fatigue life reduces due to the stress concentration caused by surface roughness. The stress concentration governs the fatigue crack initiation and propagation. The accurate acquisition of the stress concentration factor of rough surfaces is a key issue in determining fatigue life. Nevertheless, semi-empirical models may be biased for various machining processes. Besides, finite element method simulations cannot give explicit expression of the stress concentration factor. Bayesian learning can construct accurate prediction models which offering a number of additional advantages. In this paper, based on several data pairs constructed by finite element method, the correlation expression between the stress concentration factor and statistical roughness parameters of surfaces is established quickly through Bayesian learning. Compared with some other semi-empirical models, the accuracy and stability of the proposed method are certified. This paper provides a simple and efficient approach to determine the stress concentration factor for rough surfaces under different processing conditions.

17 *Keywords: [surface roughness, stress concentration factor, statistical roughness parameters,*
18 *Bayesian learning]*

19
20
21 **1. INTRODUCTION**

22
23 The fatigue life of structures is known to influence the service life of the structures highly. Moreover, the fatigue life depends on the surface quality, which will lead to the fatigue crack. There are three parameters to describe surface quality: (i) surface roughness, (ii) residual stress, (iii) microstructure. For some materials, especially aluminium alloy, surface roughness is the dominant parameter for fatigue life. Surface roughness introduces local stress concentration governing the crack initiation and propagation. The experiments show that the fatigue strength increases with a decrease in the surface roughness [1]. Therefore, the effect of surface roughness on the fatigue life cannot be neglected. Surface roughness is described through statistical roughness parameters. Traditional fatigue life predictions use empirical reduction factors to calculate the effect of surface roughness [2-5]. However, it is not a simple way due to many materials and machining processes. The empirical reduction factors need to be found for the same machining process through time-consuming testing. Beyond that method, the effect of statistical roughness parameters on the fatigue life is proposed in several approaches. In those methods, surface roughness is considered as causing stress concentration to reduce the crack initiation and propagation life. The effect can be described through the fatigue stress concentration factor K_f . It is defined as

24
25
26
27
28
29
30
31
32
33
34
35

$$K_f = \frac{\sigma_D^{(\text{smooth})}}{\sigma_D^{(\text{rough})}} \quad (1)$$

And it can be expressed by the stress concentration factor K_t according to

$$K_f = 1 + q(K_t - 1) \quad (2)$$

This consideration requires the calculation of the stress concentration factor to evaluate the fatigue life of engineering components [6].

There are two methods to study the influence of surface roughness on the stress concentration factor. One method is a semi-empirical formula, and the other is a numerical simulation. Some researchers have treated surface roughness as a series of microscopic notches. A stress concentration factor K_t is introduced. Neuber [7] considered that the surface topography is characterized by adjacent continuous gaps and proposed to estimate K_t according to

$$K_t = 1 + n\sqrt{\lambda R_z \rho^{-1}} \quad (3)$$

where λ refers to the ratio between spacing and height of surface irregularities, and n represents the conditions (where $n=1$ represents the shear and $n=2$ represents tensile). Nevertheless, the actual value of λ is hard to define for generic surface textures. Arola and Ramulu [8] suggested a different formula, where

$$K_t = 1 + n(R_a / \rho)(R_y / R_z) \quad (4)$$

The surface topography is simplified as an ideal sinusoidal micro-notch, and for AISI 430 CR steel, this expression can estimate the fatigue stress concentration factor (K_f) more accurately. Also, Andrews and Sehitoglu [9] treated the surface topography as a more common semi-elliptical notch and gave the expression as

$$K_t = 1 - 0.719 \exp \left[-0.476 \left(\frac{b}{a} \right) \right] \times \left[1 + 2 \left(\frac{c}{\rho} \right)^{1/2} \right] \quad (5)$$

where the notch width and notch spacing are denoted as a and b , the notch depth and tip radius are symbolized by c and ρ . The above methods simplify the surface topography to microscopic notches, making it difficult to estimate the surfaces under different processing conditions accurately. Another method to study the problem calculates K_t from finite element method (FEM) simulations of the measured surface topography. As et al. [10,11] used the FEM description of the surface profile to calculate K_t , the surface topography was measured using a white light interferometry microscope. Suraratchai et al. [12] fitted the measured machined surface topography and analyzed K_t by FEM simulations. However, the FEM simulations cannot give the explicit expression of the stress concentration factor. Bayesian learning method performs well in constructing prediction models. Tripping et al. [13,14] introduced the bayes theory and its combination with relevance vector machine systematically. Whereafter, Sivia at al. [15]and Gelman et al. [16] had respectively written a comprehensive exposition of the derivation of Bayesian learning. Sergio [17] stated Bayesian learning from a Bayesian and optimization perspective in machine learning. Shufei Ge et al. [18] derived a Markov chain Monte Carlo (MCMC) algorithm to the Bayesian learning. About Markov chain Monte Carlo (MCMC) algorithm, Andrieu et al. and Walter et al. [19,20] had detailed introductions in the articles. And Bhattacharya et al. [21] completed Gibbs posterior inference in Bayesian learning. Meanwhile, Hagiwara et al. [22,23] further developed hierarchical Bayesian model on multimodal information. At present, Bayesian learning has been widely used in a number of fields. Yuen et al. [24] summarized developments of Bayesian model in civil engineering. Ching et al. [25] and Kamaritotis et al. [26] modeled in detail the structural health monitoring with Bayesian learning using Gibbs sampler. Huang et al. [27], Zhang et al. [28] and Li et al. [29] detected structural damage based on hierarchical sparse Bayesian learning. Chatterjee et al. [30] determined unknown coefficients by Bayesian learning in optimization framework based robust design. Fox et al. [31] imposed a two-level regression model based Bayesian estimation in an item response theory model. Xiaochen Ma et al. [32] combined sparse Bayesian learning with least squares estimation for materials' measurement of echo reduction.

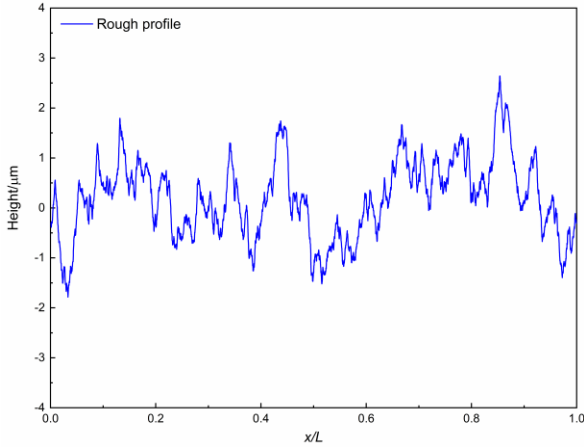
This paper establishes the correlation expression between the stress concentration factor and statistical roughness parameters of surfaces through Bayesian learning. The datasets of K_t are first constructed by FEM simulations of surface topography. Subsequently, Bayesian learning with Gibbs sampler is employed to predict expression between stress concentration factor and statistical roughness parameters. The predicted values are in good agreement with the FEM simulations compared with some other semi-empirical models, which certified the accuracy and stability of the proposed method.

2. MATERIAL AND METHODS

2.1 Finite element method

2.1.1 Rough surfaces

89 The rough surfaces are generated randomly through an open-source code in MATLAB [33]. The usability of this code has
 90 been experimentally verified by the proposer. According to the study [5], machined surface roughness (R_a) values over
 91 $0.1 \mu\text{m}$, which strongly influences fatigue life. With the surface roughness (R_a) being less than $0.1 \mu\text{m}$, this effect
 92 diminishes as cracks initiate due to persistent slip bands or grain boundaries. Therefore, when the surface topography is
 93 generated, the value of R_a is set to be greater than $0.1 \mu\text{m}$. Moreover, the lower frequency cutoff and the upper frequency
 94 cutoff are incorporated through a slight modification. The surfaces can also be assigned from actual rough topographies
 95 through a white light interferometry microscope. The profile of the generated rough surface is shown in Fig. 1, where x
 96 represents the horizontal position of the rough central axis, and L represents the width of the rough surface. x/L
 97 represents the dimensionless relative position.
 98



99
 100 **Fig. 1 The profile of generated surface topography**

101 Surface roughness is characterized through statistical roughness parameters such as R_a (average roughness), R_y (peak-
 102 to-valley height roughness), R_z (10-point roughness), and ρ (the average profile valley radius). These parameters are
 103 defined in terms of the profile height distribution (z) recorded, in respect to the mean line, over an assessment length (L)
 104 according to
 105

$$106 \quad R_a = \frac{1}{L} \int_0^L |z(x)| dx \quad (6)$$

$$107 \quad R_y = |z_{max} - z_{min}| \quad (7)$$

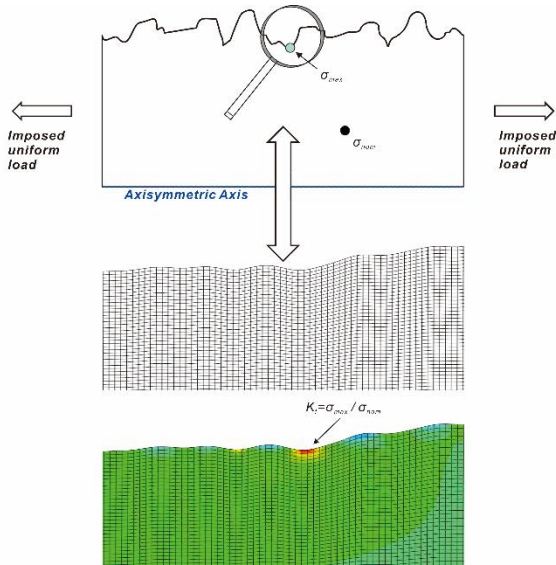
$$108 \quad R_z = \frac{1}{5} \left[\sum_{i=1}^5 (z_i)_{max} + \sum_{j=1}^5 |(z_j)_{min}| \right] \quad (8)$$

109 where $(z)_{max}$ and $(z)_{min}$ are the five higher local maxima and lower local minima, respectively, of the profile height
 110 distribution (z).
 111

112 **2.1.2 Finite element method simulations**

113 In FEM simulations, the assumption of isotropic and homogeneous materials is introduced. When calculating the stress
 114 concentration factor, only the linear elastic behavior of the material should be considered. Firstly, a smooth finite element
 115 model is established by using the commercial software ABAQUS. Then, to realize the surface of model is the above
 116 profile generated by MATLAB, MATLAB programming is used to change the surface node coordinates of the generated
 117 model.
 118

119 Fig. 2 shows the example of the final generated finite element model and simulation performed to determine K_t .
 120



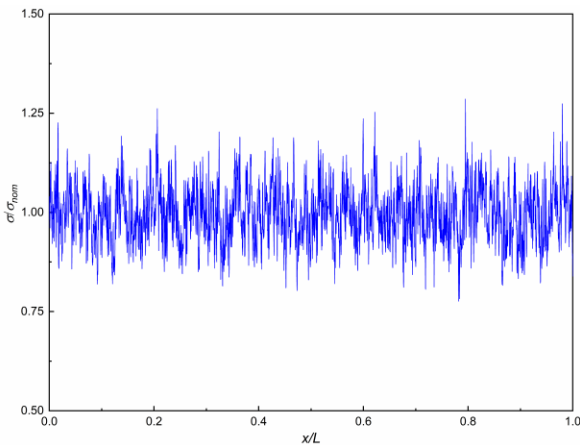
121
122 **Fig. 2 Finite element model of determining the stress concentration factor**
123

124 For the turning process, the specimens are rotated in a lathe. This process produces circumferential grooves through a
125 repeated way, which can be modelled by an axisymmetric model. The height of the model is 5000 μm , and the nominal
126 length of the rough surface is 20000 μm . Uniform tensile load is applied to both ends of the boundary. Young's module E
127 is 72 GPa, and Poisson's ratio ν is 0.33. In FEM simulations, axisymmetric quadrilateral elements (CAX4) are adopted to
128 discretize the model, in which the meshes are gradually coarsened along the r -axis from top to bottom. For one of rough
129 surfaces ($R_a=0.192818$, $R_y=1.191023$, $R_z=1.179942$, $\rho=569.1562$), the minimum mesh size of the adjacent surface
130 topography is 0.4 μm . As preparation, the accuracy of our computational results has been guaranteed by convergence
131 tests.

132 The stress concentration factor of each point on the surface topography is defined as

133
$$K = \frac{\sigma}{\sigma_{nom}} \quad (9)$$

134 where σ is the Von Mises equivalent stress, and σ_{nom} is the nominal Von Mises equivalent stress of the cross-section, as
135 shown in Fig. 3, where x represents the horizontal position of the rough central axis, and L represents the width of the
136 rough surface. x/L represents the dimensionless relative position.
137



138
139 **Fig. 3 The stress concentration factor K of each point on the surface topography**
140

141 And the maximal Von Mises equivalent stress obtained by the calculation is then divided by the nominal Von Mises
142 equivalent stress due to the applied load to determine the stress concentration factor K_t classically, expressed as

143
$$K_t = \frac{\sigma_{max}}{\sigma_{nom}} \quad (10)$$

144 where σ_{max} is the maximal Von Mises equivalent stress along with the surface topography.
145

146 **2.1.3 Data**

147
148
149
150
151
152
153
154
155
156

By FEM simulations, the statistical roughness parameters, and corresponding stress concentration factor of 41 different specimens have obtained, which constitutes the datasets, as shown in Table 1. The datasets consist of 41 data pairs, including R_a , R_y , R_z , ρ these four independent variables that characterize the properties of rough surface and K_t which describe the stress concentration caused by surface roughness. The value of R_a is set to be greater than 0.1 μm and less than 6.4 μm . The range of R_y , R_z , ρ and K_t are (1.19, 44.81), (1.17, 44.53), (16.14, 569.16) and (1.04, 2.73) to verify that this method is applicable to rough surfaces under various processing conditions.

Table 1. Database constructed by FEM simulations

Number	$R_a/\mu\text{m}$	$R_y/\mu\text{m}$	$R_z/\mu\text{m}$	$\rho/\mu\text{m}$	K_t
1	3.555986	26.30374	26.11672	16.99651	2.153361439
2	4.2303	22.98283	22.69772	25.07418	2.105857375
3	5.440313	31.41201	31.00602	17.03947	2.638873261
4	0.772174	4.512271	4.475618	137.147	1.153786981
5	4.025152	25.22172	25.04139	23.22475	1.963668871
6	5.583996	30.33734	29.96099	19.03229	2.519425105
7	4.726791	30.61076	30.35646	26.28119	1.983688209
8	0.331772	2.402115	2.379683	246.3604	1.09061872
9	2.4	12.32988	12.24221	36.10878	1.509358561
10	3.693573	26.18617	25.79348	25.35032	2.040652054
11	0.192818	1.191023	1.179942	569.1562	1.03880591
12	6.000485	36.44235	36.27377	17.45057	2.673570711
13	3.252341	23.35157	23.11154	31.73201	1.863292314
14	1.92	9.863903	9.793769	45.11696	1.402105278
15	2.404957	14.19217	13.96398	39.43441	1.640159853
16	2.400194	14.57694	14.50951	38.80922	1.622350349
17	1.721862	11.27853	11.23405	56.88085	1.37554257
18	5.550449	33.70917	33.55324	18.84812	2.535736852
19	7.066363	44.00704	43.72766	16.13747	2.733849542
20	1.018625	6.101202	6.014407	93.90284	1.24962646
21	2.226748	11.31604	11.21579	41.2961	1.480055216
22	0.792587	3.752242	3.695435	101.3596	1.185491742
23	5.69759	40.40472	39.95214	17.55418	2.184433519
24	1.121263	7.949373	7.830165	83.26314	1.299609142
25	4.074499	24.40481	24.05763	41.9364	2.12358941
26	5.337771	34.96346	34.82555	18.41903	2.325390202
27	1.118098	7.006034	6.955942	83.35896	1.247479891
28	5.59575	32.3095	31.89191	16.57165	2.690652606
29	2.346066	16.63724	16.45088	42.45231	1.459829098
30	6.398332	44.80768	44.53094	16.74337	2.546558207
31	3.053939	17.03504	16.92747	39.79985	1.681130644
32	1.894048	12.40639	12.35745	51.71469	1.416110263
33	5.045286	36.30095	36.14158	18.61619	2.129087822

34	3.890513	23.19896	22.96311	33.14648	1.892614248
35	4.597006	21.76301	21.43352	17.53594	2.207043075
36	3.431884	20.05454	19.89164	30.91224	1.751842258
37	3.400587	23.04051	22.9307	28.43874	1.743377001
38	4.118261	24.06545	23.86996	25.78116	1.918798908
39	1.505916	7.129261	7.021326	53.36182	1.362512558
40	0.606735	37.02597	36.87318	27.83285	2.536359788
41	0.517418	3.715022	3.676836	199.0674	1.125912072

2.2 Bayesian learning

According to the existing models, several models propose the expression between K_t and the statistical roughness parameters of surfaces, as shown in Table 1.

Table 2. Some empirical models for estimating the stress concentration factor K_t

Empirical models	Literature
$K_t = 1 + 2\sqrt{\lambda R_z \rho^{-1}}$	Neuber
$K_t = 1 + 2R_a R_y (R_z \rho)^{-1}$	Arola-Ramulu

Those models can be rewritten as linear models. Therefore, the expression is assumed as

$$\ln(K_t - 1) = \omega_1 \ln(R_a / \rho) + \omega_2 \ln(R_a / R_y) + \omega_3 \ln(R_a / R_z) + \omega_4 \quad (11)$$

The Bayesian learning [14-17] with Gibbs sampler [18-22] is used to estimate ω .

2.2.1 Bayesian theorem

According to the Bayesian theorem and the proposed methods of early scholars, the expression can be written as a linear combination of function Φ of the form

$$y_i(\mathbf{X}_i, \omega) = \sum_{m=0}^M \omega_m f_m(\mathbf{X}_i) = \omega^T \Phi \quad (12)$$

where $y_i(\mathbf{X}_i, \omega)$ represents the function of predicted K_t , expressed as

$$y_i(\mathbf{X}_i, \omega) = \ln[(K_t)_i - 1] \quad (13)$$

\mathbf{X}_i represents the independent variables, and the Φ is a function of \mathbf{X}_i . The ω_m are the parameters of the model and are generally called weights.

To support additive noise between $y_i(\mathbf{X}_i, \omega)$ and $\omega^T \Phi$, the Eq. (7) will be rewritten as

$$y_i(\mathbf{X}_i, \omega) = \sum_{m=0}^M \omega_m \phi_m(\mathbf{X}_i) + \varepsilon \quad (14)$$

where ε is assumed to be Gaussian distribution expressed as

$$P(\varepsilon | \tau) = N(\varepsilon | 0, \tau^{-1}) \quad (15)$$

and τ follows Gamma distribution as

$$P(\tau) = \Gamma(\tau | a, b) = b^a \tau^{a-1} e^{-b\tau} / \Gamma(a) \quad (16)$$

Also, to complete the specification of hierarchical prior, choose a zero-mean Gaussian prior distribution over ω .

$$P(\omega | \alpha) = N(\omega | 0, \alpha^{-1}) \quad (17)$$

Suitable priors thereover are Gamma distribution, which is expressed as

$$P(\alpha) = \Gamma(\alpha | c, d) = \prod_{j=1}^n d^c \alpha_j^{c-1} e^{-d\alpha_j} / \Gamma(c) \quad (18)$$

In Bayesian theorem, the basic idea is to convert a prior probability density function for the parameters ω into a posterior distribution with the data \mathbf{D} in the form

$$P(\omega | \mathbf{D}) = \frac{P(\mathbf{D} | \omega) P(\omega)}{P(\mathbf{D})} \quad (19)$$

192 Thus, the posterior PDF is derived as below by Bayesian theorem,

$$193 \quad P(\boldsymbol{\omega}, \tau, \boldsymbol{\alpha} | \mathbf{D}) = [P(\mathbf{D} | \boldsymbol{\omega}, \tau) P(\boldsymbol{\omega} | \boldsymbol{\alpha}) P(\tau) \\ P(\boldsymbol{\alpha})] / P(\mathbf{D}) \quad (20)$$

194 **2.2.2 Gibbs sampler**

195 The process for Gibbs sampler is as below[15],

196 (i) select the training data pairs,

197 (ii) initialize the parameters $\boldsymbol{\alpha}(0)$ and $\tau(0)$,

198 (iii) **For** $n=1, 2, \dots, N$, **Do**

199 sample the parameter $\boldsymbol{\omega}(n) \sim P(\boldsymbol{\omega} | \boldsymbol{\alpha}(n-1), \tau(n-1), \mathbf{D})$; sample the parameters $\boldsymbol{\alpha}(n) \sim P(\boldsymbol{\alpha} | \boldsymbol{\omega}(n), \mathbf{D})$ and $\tau(n) \sim P(\tau | \boldsymbol{\omega}(n), \mathbf{D})$

200 - **IF** $\boldsymbol{\omega}(n)$, $\boldsymbol{\alpha}(n)$ and $\tau(n)$ samples reach stationary states

201 **End For**

202 **2.2.3 Burn-in period**

203 In practice, an initial set of samples (burn-in) are often discarded to avoid starting biases [20]. The period required for the Markov chain to reach its stationary state is called the burn-in period. The burn-in period is usually determined by plotting the Markov chain samples through time. Moreover, the samples within the burn-in period are discarded. In this paper, by plotting the Markov chain samples, the burn-in period is obtained by visual inspection.

204 **2.2.4 The stability of the method**

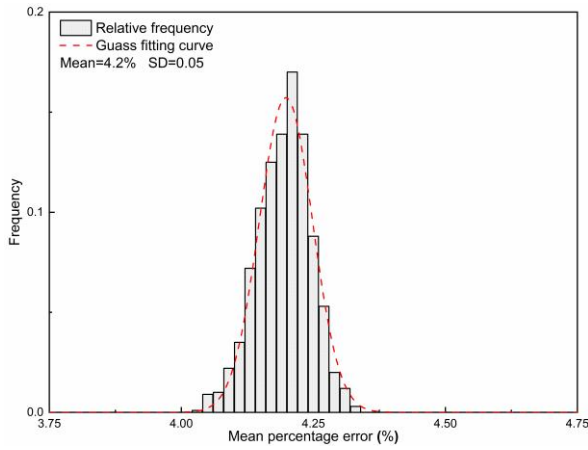
205 The coefficient of determination (R^2) and mean percentage error (MPE) is used to evaluate the accuracy of the proposed method, which are obtained as below. Where $(K_t)_i$ is the i^{th} simulated data, $((K_t)_{pre})_i$ is the i^{th} predicted value. The best value for R^2 is 1.0, and the smaller the value of MPE, the better the predicted result.

$$206 \quad R^2 = \left(\frac{n \sum \mu_{(K_t)_i} \times (K_t)_i - \sum \mu_{(K_t)_i} \sum (K_t)_i}{\sqrt{n \sum \mu_{(K_t)_i}^2 - \left(\sum \mu_{(K_t)_i} \right)^2} \sqrt{n \sum (K_t)_i^2 - \left(\sum (K_t)_i \right)^2}} \right)^2 \quad (21)$$

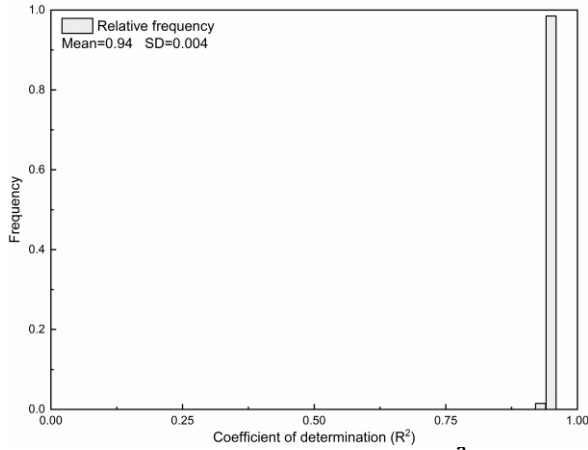
$$207 \quad \text{MPE} = \left(\frac{1}{n} \sum_{i=1}^n \left| (K_t)_i - ((K_t)_{pre})_i \right| / (K_t)_i \right) \times 100\% \quad (22)$$

208 The procedure for a set of experiments is described below. Five pairs are randomly selected from the datasets as training samples, and the rest of the data pairs are testing sets. The predicted model of the stress concentration factor is obtained by the method mentioned in Eq. (20). Comparing the predicted stress concentration factor with the simulated value, R^2 and MPE are calculated.

209 A total of 1000 experiments are conducted. The frequency histogram of MPE and R^2 are plotted in Fig. 4 and Fig. 5, where "Mean" represents the mean value of 1000 experiments, and "SD" represents standard deviation. By the figures, for 1000 experiments, the mean value of MPE is 4.2%, and its standard deviation is 0.05. Meanwhile, for most experiments, the R^2 is around 0.94, and the standard deviation is 0.004. MPE and R^2 remain within the error tolerance range, which indicates the excellent stability of the method.



231
232 **Fig. 4 Mean percentage error (MPE) of 1000 experiments**
233



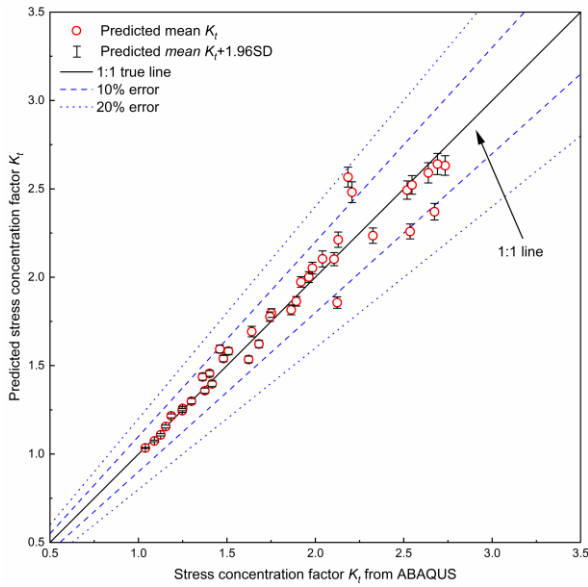
234
235 **Fig. 5 Coefficient of determination(R^2) of 1000 experiments**
236

237 3. RESULTS AND DISCUSSION

238
239 Firstly, using the five data from datasets as training samples, the predicted model of stress concentration factor is
240 obtained through the proposed method. The first 500 samples of 10,000 samples obtained by Gibbs sampling are
241 discarded because of the burn-in period. Based on the remaining 9,500 samples, the predicted mean stress concentration
242 factor is calculated from their means and standard deviations. The predicted formula between K_t and statistical roughness
243 parameters of the surfaces is expressed as

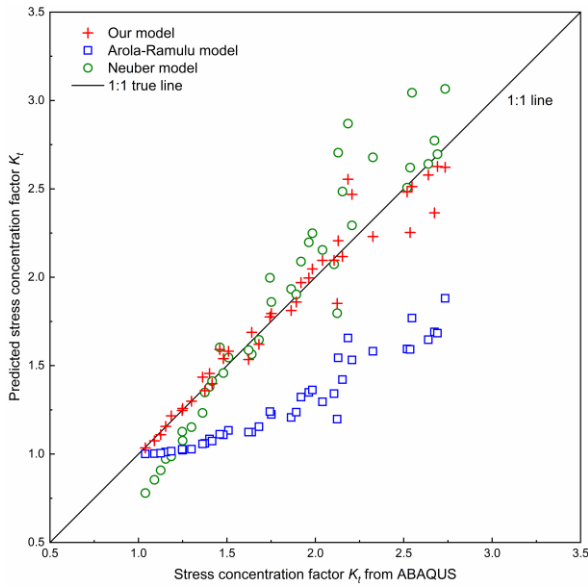
244
$$K_t = 1 + 1.915 \left(\frac{R_a}{R_y} \right)^{-0.223} \left(\frac{R_a}{R_z} \right)^{0.122} \left(\frac{R_a}{\rho} \right)^{0.549} \quad (23)$$

245
246 To make the results more intuitive, each predicted K_t in the 41 datasets are compared with the simulated K_t in Fig. 6. It is
247 showed by open circles from mean K_t and error bars for uncertainty in terms of $K_t + 1.96SD$, where SD is the standard
248 deviation of each predicted K_t , to show the ranges of data. Fig. 6 also includes the true 1:1 line for comparison. It shows
249 that open circles fluctuate around the 1:1 line, and most samples have relative errors smaller than 10%, and all samples
250 smaller than 20%. The maximum error is found to be 15.4%, and the average error is 3.7%. For quantitative the predicted
251 model, R^2 and MPE between predicted and simulated K_t are calculated, obtained as $R^2=0.95$, $MPE=3.76\%$. The large R^2
252 and small MPE together indicate that the predicted model is reasonably accurate.



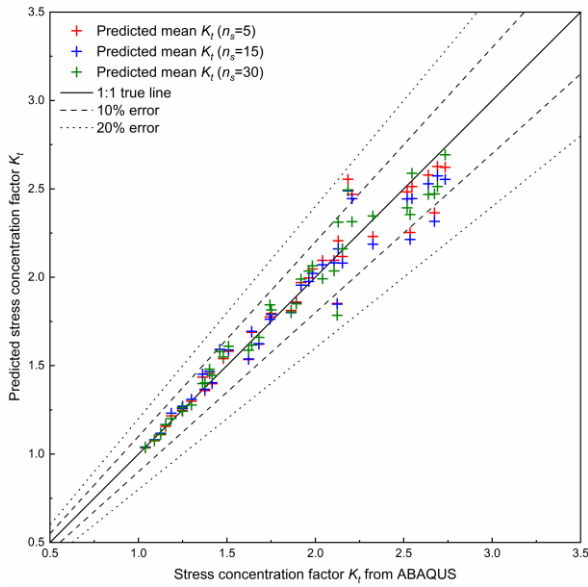
253
254 **Fig. 6 Predicted versus simulated K_t**
255

256 For comparison, Fig. 7 also includes the stress concentration factor predicted by the Neuber model and Arola- Ramulu
257 model. K_t predicted by the proposed method is closer to the 1:1 line, which indicates our predicted values are more
258 consistent with the simulated values than the other two models. The average error of these two models is 9.87% and
259 21.23%, while for the proposed method, the average error is 3.76%.
260



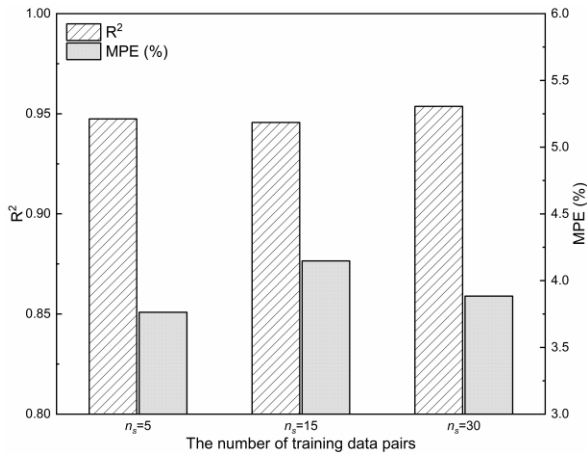
261
262 **Fig. 7 Estimation results of K_t by the proposed method, Neuber model and Arola- Ramulu model**
263

264 To verify the influence of the number of data pairs (n_s) on estimation performance, the randomly 5, 15, 30 groups of
265 datasets are selected as training samples to obtain the corresponding predicted expressions. The predicted K_t are
266 compared with simulated K_t , as shown in Fig. 8. The values predicted by the three correlation expressions all fluctuate
267 around the 1:1 line. Furthermore, R^2 and MPE of three correlation expressions are plotted in Fig. 9. R^2 and MPE are kept
268 around 0.94 and 4%, respectively.



269
270
271

Fig. 8 Predicted K_t obtained by different n_s ($n_s=5, 15, 30$) versus simulated K_t



272
273
274

Fig. 9 R^2 and MPE calculated by different n_s ($n_s=5, 15, 30$)

275

It indicates no significant effect on the predicted results with decreasing the number of training samples in the appropriate range. In other words, the estimated error of the proposed method is on the reasonable range when the number of data pairs decreases. The results demonstrate that the proposed method performs well under limited training sets.

276

277

278

The above analysis proved that the proposed method could efficiently obtain the stress concentration factor explicit and has better accuracy than the previous models. Moreover, this method is quite convenient to be conducted through only a few FEM simulations or tests for certain machining processes.

279

280

281

282

4. CONCLUSION

283

284

This article proposes a novel Bayesian learning to obtain correlation expression between the stress concentration factor and statistical roughness parameters. The training samples and testing samples are selected from the database, which is constructed by FEM simulations. Using the Bayesian learning with Gibbs sampler, the correlation expression between the stress concentration factor and statistical roughness parameters of the surfaces is estimated. The accurate correlation expression gives good agreements with direct FEM simulations, which demonstrates the feasibility of using the proposed method. Compared to other models, the proposed method has the advantages of high accuracy and broad adaptability. In conclusion, this work demonstrates the viability of using machine learning to predict the stress concentration factor. It provides a simple and efficient approach to obtain the predicted stress concentration factor for the rough surfaces under a specific type of processing.

285

286

287

288

289

290

291

292

293

AUTHORS' CONTRIBUTIONS

Authors have declared that no competing interests exist.

CONSENT (WHERE EVER APPLICABLE)

All authors declare that “written informed consent was obtained from the patient (or other approved parties) for publication of this case report and accompanying images.” A copy of the written consent is available for review by the Editorial office/Chief Editor/Editorial Board members of this journal.

REFERENCES

1. Bayoumi M. R.; Abdellatif A. Effect of surface finish on fatigue strength. *Engineering Fracture Mechanics*. 1995; 51 (5):861-870.
[https://doi.org/10.1016/0013-7944\(94\)00297-U](https://doi.org/10.1016/0013-7944(94)00297-U).
2. Maiya P. Geometrical characterization of surface roughness and its application to fatigue crack initiation. *Materials Science and Engineering*. 1975; 21:57-62.
[https://doi.org/10.1016/0025-5416\(75\)90198-6](https://doi.org/10.1016/0025-5416(75)90198-6).
3. Maiya P. Effects of surface roughness and strain range on the low-cycle fatigue behavior of type 304 stainless steel. *Scr Metall;(United States)*. 1975; 9 (11).
[https://doi.org/10.1016/0036-9748\(75\)90424-X](https://doi.org/10.1016/0036-9748(75)90424-X).
4. Wiesner C.; Künzi H. -U.; Ilschner B. Characterization of the topography of turned surfaces and its influence on the fatigue life of Al-7075. *Materials Science and Engineering: A*. 1991; 145 (2):151-158.
[https://doi.org/10.1016/0921-5093\(91\)90244-H](https://doi.org/10.1016/0921-5093(91)90244-H).
5. Novovic D.; Dewes R.; Aspinwall D.; Voice W.; Bowen P. The effect of machined topography and integrity on fatigue life. *International Journal of Machine Tools and Manufacture*. 2004; 44 (2-3):125-134.
<https://doi.org/10.1016/j.ijmachtools.2003.10.018>.
6. Pilkey W. D.; Pilkey D. F. *Peterson's Stress Concentration Factors*. New York : John Wiley & Sons; 2008.
7. Neuber H. Theory of stress concentration for shear-strained prismatical bodies with arbitrary nonlinear stress-strain law. *Journal of Applied Mechanics*. 1961.
<https://doi.org/10.1115/1.3641780>.
8. Arola D.; Ramulu M. An examination of the effects from surface texture on the strength of fiber reinforced plastics. *Journal of Composite Materials*. 1999; 33 (2):102-123.
<https://doi.org/10.1177/002199839903300201>.
9. Andrews S.; Sehitoglu H. A computer model for fatigue crack growth from rough surfaces. *International Journal of fatigue*. 2000; 22 (7):619-630.
[https://doi.org/10.1016/S0142-1123\(00\)00018-9](https://doi.org/10.1016/S0142-1123(00)00018-9).
10. Ås S. K.; Skallerud B.; Tveiten B. W.; Holme B. Fatigue life prediction of machined components using finite element analysis of surface topography. *International Journal of Fatigue*. 2005; 27 (10-12):1590-1596.
<https://doi.org/10.1016/j.ijfatigue.2005.07.031>.
11. Ås S. K.; Skallerud B.; Tveiten B. W. Surface roughness characterization for fatigue life predictions using finite element analysis. *International Journal of Fatigue*. 2008; 30 (12):2200-2209.
<https://doi.org/10.1016/j.ijfatigue.2008.05.020>.
12. Suraratchai M.; Limido J.; Mabru C.; Chieragatti R. Modelling the influence of machined surface roughness on the fatigue life of aluminium alloy. *International Journal of fatigue*. 2008; 30 (12):2119-2126.
<https://doi.org/10.1016/j.ijfatigue.2008.06.003>.
13. Tipping M. Sparse Bayesian Learning and the Relevance Vector Machine. *Journal of Machine Learning Research*. 2001; 1:211-244.
14. Bishop C. M.; Tipping M. Variational relevance vector machines. *arXiv preprint arXiv:13013838*. 2013.
15. Sivia D.; Skilling J. *Data analysis: a Bayesian tutorial*. London : OUP Oxford; 2006.
16. Gelman A.; Carlin J. B.; Stern H. S.; Dunson D. B.; Vehtari A.; Rubin D. B. *Bayesian data analysis*. Boca Raton : CRC press.; 2013.
17. Sergios T. *Machine learning: A bayesian and optimization perspective*. 2nd ed. Elsevier Ltd; 2020.
18. Ge S, Wang S, Nathoo F S, et al. Online Bayesian learning for mixtures of spatial spline regressions with mixed effects. *Journal of Statistical Computation and Simulation*, 2021: 1-37.
<https://doi.org/10.1080/00949655.2021.2002329>
19. Andrieu C.; De Freitas N.; Doucet A.; Jordan M. I. An introduction to MCMC for machine learning. *Machine learning*. 2003; 50 (1-2):5-43.

- 353 <https://doi.org/10.1023/A:1020281327116>.
- 354 20. Walter J. -C.; Barkema G. An introduction to Monte Carlo methods. *Physica A: Statistical Mechanics and its*
355 *Applications*. 2015; 418:78-87.
- 356 <https://doi.org/10.1016/j.physa.2014.06.014>.
- 357 21. Bhattacharya I, Martin R. Gibbs posterior inference on multivariate quantiles. *Journal of Statistical Planning and*
358 *Inference*, 2021; 218:106–121.
- 359 <https://doi.org/10.1016/j.jspi.2021.10.003>
- 360 22. Hagiwara Y, Taguchi K, Ishibushi S, et al. Hierarchical Bayesian model for the transfer of knowledge on spatial
361 concepts based on multimodal information. *Advanced Robotics*. 2021.
- 362 <https://doi.org/10.1080/01691864.2021.2004224>
- 363 23. Liu Y, Li L, Zhao S. Efficient Bayesian updating with two-step adaptive Kriging. *Structural Safety*, 2022, 95: 102172.
- 364 <https://doi.org/10.1016/j.strusafe.2021.102172>
- 365 24. Yuen K. -V. 2010. Recent developments of Bayesian model class selection and applications in civil engineering.
366 *Structural Safety* 32 (5):338-346.
- 367 <https://doi.org/10.1016/j.strusafe.2010.03.011>.
- 368 25. Ching J.; Muto M.; Beck J. L. Structural model updating and health monitoring with incomplete modal data using Gibbs
369 sampler. *Computer - Aided Civil and Infrastructure Engineering*. 2006; 21 (4):242-257.
- 370 26. Kamariotis A, Chatzi E, Straub D. Value of information from vibration-based structural health monitoring extracted via
371 Bayesian model updating. *Mechanical Systems and Signal Processing*, 2022; 166: 108465.
- 372 <https://doi.org/10.1016/j.ymssp.2021.108465>
- 373 27. Huang Y.; Beck J. L.; Li H. Bayesian system identification based on hierarchical sparse Bayesian learning and Gibbs
374 sampling with application to structural damage assessment. *Computer Methods in Applied Mechanics and Engineering*.
375 2017; 318:382-411.
- 376 <https://doi.org/10.1016/j.cma.2017.01.030>.
- 377 28. Zhang Z, Sun C, Guo B. Transfer-learning guided Bayesian model updating for damage identification considering
378 modeling uncertainty. *Mechanical Systems and Signal Processing*, 2022, 166: 108426.
- 379 <https://doi.org/10.1016/j.ymssp.2021.108426>
- 380 29. Li J, Huang Y, Asadollahi P. Sparse Bayesian learning with model reduction for probabilistic structural damage
381 detection with limited measurements[J]. *Engineering Structures*, 2021; 247: 113183.
- 382 <https://doi.org/10.1016/j.engstruct.2021.113183>
- 383 30. Chatterjee T, Chowdhury R. An efficient approximation-based robust design optimization framework for large-scale
384 structural systems. Butterworth-Heinemann, 2020.
- 385 31. Le Folgoc L.; Delingette H.; Criminisi A.; Ayache N. Quantifying registration uncertainty with sparse bayesian
386 modelling. *IEEE transactions on medical imaging*. 2016; 36 (2):607-617.
- 387 <https://doi.org/10.1109/TMI.2016.2623608>.
- 388 32. Ma X, Li J, Huang Y, et al. Measurement of echo reduction for passive-material samples using sparse Bayesian
389 learning and least squares estimation. *The Journal of the Acoustical Society of America*, 2021; 150(5): 3251-3262.
- 390 <https://doi.org/10.1121/10.0006753>
- 391 33. Kanafi M. M. Surface generator: artificial randomly rough surfaces. MATLAB Central File Exchange. 2020. Accessed
392 24 Jan. 2021.
- 393 Available: [https://www.mathworks.com/matlabcentral/fileexchange/60817-surface-generator-artificial-randomly-rough-](https://www.mathworks.com/matlabcentral/fileexchange/60817-surface-generator-artificial-randomly-rough-surfaces)
394 [surfaces](https://www.mathworks.com/matlabcentral/fileexchange/60817-surface-generator-artificial-randomly-rough-surfaces).