

Method for solving the nonlinear inverse problem in gas face seal diagnosis based on surrogate models

Yuan YIN, Weifeng HUANG (✉), Decai LI, Qiang HE, Xiangfeng LIU, Ying LIU

State Key Laboratory of Tribology in Advanced Equipment, Tsinghua University, Beijing 100084, China

✉ Corresponding author. E-mail: huangwf@tsinghua.edu.cn (Weifeng HUANG)

© Higher Education Press 2022

ABSTRACT Physical models carry quantitative and explainable expert knowledge. However, they have not been introduced into gas face seal diagnosis tasks because of the unacceptable computational cost of inferring the input fault parameters for the observed output or solving the inverse problem of the physical model. The presented work develops a surrogate-model-assisted method for solving the nonlinear inverse problem in limited physical model evaluations. The method prepares a small initial database on sites generated with a Latin hypercube design and then performs an iterative routine that benefits from the rapidity of the surrogate models and the reliability of the physical model. The method is validated on simulated and experimental cases. Results demonstrate that the method can effectively identify the parameters that induce the abnormal signal output with limited physical model evaluations. The presented work provides a quantitative, explainable, and feasible approach for identifying the cause of gas face seal contact. It is also applicable to mechanical devices that face similar difficulties.

KEYWORDS surrogate model, gas face seal, fault diagnosis, nonlinear dynamics, tribology

1 Introduction

A gas face seal is a type of noncontacting mechanical face seal, such as the type designed to work with a pair of end faces separated by a gas film. However, various faults can occur inside the seal and result in failures [1]. Many studies [2–7] have reported the use of online monitoring techniques to obtain real-time signals from mechanical face seals (including gas face seals). However, quantitatively determining the cause of an abnormal observation found by the sensor(s) remains challenging.

Physical models for gas face seals have long been applied in design tasks [8–10]; however, these models are absent in diagnosis tasks, which should have been able to provide quantitative and explainable expert knowledge (unlike machine learning approaches [11–15] suitable for systems with no or incomplete physical models). This scenario is mainly attributed to computational reasons. The physics-based diagnosis task can be abstracted as a process of determining the input that results in an output consistent with the signals observed through the physical model or a process of solving an inverse problem. When

the model is highly nonlinear and has no explicit inverse expressions, a large number of evaluations are needed to find the correct input (particularly when multiple parameters are undetermined). However, the evaluation of the physical model is costly. Such difficulties are common in systems subject to complex numerical models and do not have explicit equations for the inverse problems.

The gas face seal dynamics model, which directly governs the leakage and friction of a seal, is taken as a typical example. The model discretizes the transient Reynolds equation in space and time and couples it with dynamic equations [16,17]. Solving the model demands considerable time and storage costs. Some existing dynamic models that linearize the transient Reynolds equation for efficiency improvement, namely, semianalytical models [18,19], are effective when applied to conditions close to a particular presupposed state. However, they are not applicable in diagnosis tasks, where the seal state may deviate over an extensive range.

An alternative approach is using surrogate models (metamodels or response surface methodology) [20,21] to replace the physical model for massive iterations by fitting a limited count of existing results. The surrogate

model technique has already been widely recognized as an optimal design tool in various fields [21–25]. In terms of applications in the field of mechanical face seals (including gas face seals), Liu et al. [26] established a quadratic polynomial response surface model fitting a seal ring vibration model for structure optimization. Additional studies are needed to demonstrate the applicability of surrogate models in mechanical face seals.

In the present work, targeting the task of determining the fault parameters of a gas face seal using the monitored acoustic emission signals is established to solve the inverse problem of the seal dynamic model. This method combines the advantage of the physical models and surrogate models. The method starts with a small initial database and uses rapid surrogate models instead of the original physical model to speculate the input with massively iterative optimizers. It also uses the physical model results as a feedback mechanism to overcome the limited accuracy of the surrogate models with small databases. The method is demonstrated in both simulated and experimental cases.

2 Model of gas face seal dynamics

A typical gas face seal structure is shown in Fig. 1. A stator with a planar end face is flexibly mounted on a fixed seat through springs and a secondary seal, and the rotor is rigidly mounted on the rotating shaft. The rotor has spiral grooves on its face (Fig. 2). The pressure is high at the outer radius and low at the inner radius. It is sealed by a pair of relatively rotating faces. The following are brief descriptions of the model.

(1) Geometry

As shown in Fig. 3, the stator moves with three degrees of freedom, denoted as z_s (axial displacement, which increases with separation), γ_{sx} (tilt angle around X axis), and γ_{sy} (tilt angle around Y axis). Besides the rotation speed ω , the motion of the rotor, denoted as z_r , γ_{rx} , and γ_{ry} , is assumed to be unaffected by the forces or moments. Here, z_r is regarded as 0, and γ_{rx} and γ_{ry} fluctuate periodically because of the perpendicularity error between the rotor face and the rotation axis. Their values are given by the angles around axes X and Y at $t = 0$ (t refers to the time), denoted as γ_{rx0} and γ_{ry0} , respectively.

$$\mathbf{U}_s = \begin{bmatrix} z_s \\ \gamma_{sx} \\ \gamma_{sy} \end{bmatrix}, \quad (1)$$

$$\mathbf{U}_r = \begin{bmatrix} z_r \\ \gamma_{rx} \\ \gamma_{ry} \end{bmatrix} = \begin{bmatrix} 0 \\ \gamma_{rx0} \cos(\omega t) + \gamma_{ry0} \sin(\omega t) \\ -\gamma_{rx0} \sin(\omega t) + \gamma_{ry0} \cos(\omega t) \end{bmatrix}, \quad (2)$$

$$h = h_0 + h_{grv} + \begin{bmatrix} 1 & y & -x \end{bmatrix} (\mathbf{U}_s - \mathbf{U}_r), \quad (3)$$

where \mathbf{U}_s and \mathbf{U}_r denote the motion of the stator and the

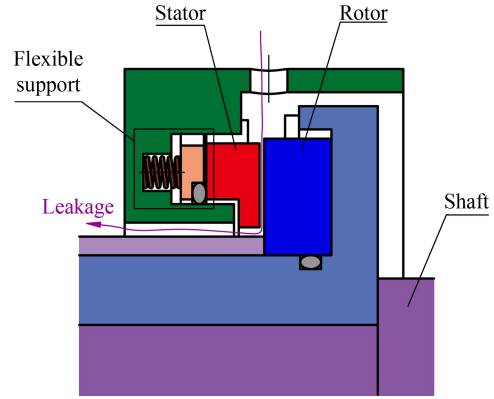


Fig. 1 Typical structure of a spiral groove gas face seal.

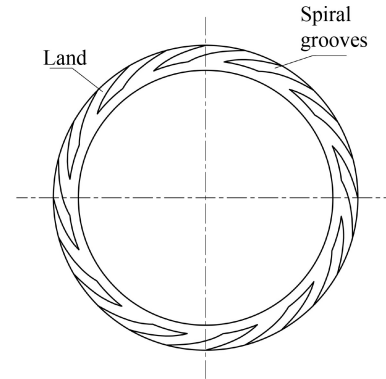


Fig. 2 Spiral grooves on the rotor face.

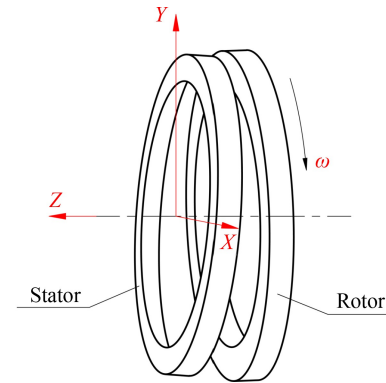


Fig. 3 Motion of the seal rings.

rotor, respectively, h is the gas film thickness defined on the annulus area $\Omega = \{(x, y) | r_i \leq \sqrt{x^2 + y^2} \leq r_o\}$, between the inner radius r_i and the outer radius r_o , and h_{grv} denotes the contribution of the grooves (see Fig. 2), which values the depth h_{spr} in the groove and 0 otherwise, as

$$h_{grv} = \begin{cases} h_{spr}, & \text{in the spiral grooves,} \\ 0, & \text{on the land.} \end{cases} \quad (4)$$

(2) Lubrication

We consider the gas to be ideal and inertia free with constant temperature and viscosity and the flow to be

laminar and nonslipping at the boundary. Under these conditions, the absolute gas pressure in the film, p , is governed by the transient average flow Reynolds equation [27] with Dirichlet boundary conditions, given by

$$\nabla \cdot \left(\frac{\varphi p h^3}{12\mu} \nabla p \right) = \frac{\omega \partial(ph)}{2\partial\theta} + \frac{\partial(ph)}{\partial t}, \quad (5)$$

$$\begin{cases} p|_{r=r_i} = p_i, \\ p|_{r=r_o} = p_o, \end{cases} \quad (6)$$

where a point is denoted as (r, θ) in the polar coordinate system, and is adopted for convenience, μ denotes the gas dynamic viscosity, p_i and p_o are the boundary pressures at the inner radius r_i and outer radius r_o , respectively, and φ is a flow factor given by [27]

$$\varphi = 1 - 0.9 \exp\left(-0.56 \frac{h}{\sigma_a}\right), \quad (7)$$

where σ_a is the standard deviation of surface height.

(3) Contact

Contact does not occur when the gas face seal operates properly. However, a wide range of faults or low-speed operation may result in contact. Fan et al. [28] stated that light face contact can be considered elastic. Given the summit heights obeying a Gaussian distribution, the contact pressure p_c is simulated as [29]

$$p_c = p_c(h) = \frac{4}{3} E \eta_s R_s^{\frac{1}{2}} \int_h^{\infty} \frac{(w-h)^{\frac{3}{2}}}{\sigma_s} \phi\left(\frac{w-\bar{w}_s}{\sigma_s}\right) dw, \quad (8)$$

where $\phi(\cdot)$ denotes the probability density function of the standard normal distribution, and w denotes local surface height. The properties of the roughness of the contacted faces are described with E , η_s , R_s , \bar{w}_s , and σ_s , referring to elastic modulus, asperity density, average asperity radius, mean of asperity height, and standard deviation of asperity height of an equivalent contact pair.

Face contact generates acoustic emissions, which have been considered an effective monitoring approach by researchers [1,4–6]. The short-time root-mean-square (RMS) is modeled as [28,30]

$$V = KV_0 = K \sqrt{\iint_{\Omega} \frac{2}{5} \omega r E \eta_s \int_h^{\infty} \frac{(w-h)^2}{\sigma_s} \phi\left(\frac{w-\bar{w}_s}{\sigma_s}\right) dw}, \quad (9)$$

where K is the coefficient describing the transmission from the acoustic emission source (with RMS V_0) to the signal sent out by the sensor, assumed to be constant, and V is the RMS value of the signal.

(4) Dynamics

The motion of the flexibly mounted stator is governed by

$$\mathbf{M}_s \ddot{\mathbf{U}}_s + \mathbf{C}_s \dot{\mathbf{U}}_s + \mathbf{K}_s \mathbf{U}_s = \mathbf{P}_f + \mathbf{P}_c + \mathbf{P}_b + \mathbf{P}, \quad (10)$$

where the inertia of the stator and the support provided by

the springs and secondary seal (modeled as linear damping and stiffness) are considered as

$$\begin{cases} \mathbf{M}_s = \text{diag}(m, J, J), \\ \mathbf{C}_s = \text{diag}(c_{sz}, c_{sy}, c_{sy}), \\ \mathbf{K}_s = \text{diag}(k_{sz}, k_{sy}, k_{sy}). \end{cases} \quad (11)$$

The inertia properties are described by the inertia matrix \mathbf{M}_s (consisting of mass m , moment of inertia J); and the damping and stiffness are described by the damping matrix \mathbf{C}_s (consisting of axial and angular damping c_{sz} and c_{sy}) and the stiffness matrix \mathbf{K}_s (consisting of axial and angular stiffness k_{sz} and k_{sy}). The damping and stiffness are considered to be potentially abnormal, modeled by being scaled by a scaling coefficient Q_s , as

$$[k_{sz}, k_{sy}, c_{sz}, c_{sy}] = Q_s [k_{sz0}, k_{sy0}, c_{sz0}, c_{sy0}]. \quad (12)$$

The generalized forces in Eq. (10) are defined as

$$\mathbf{P}_f = \begin{bmatrix} \iint_{\Omega} p d\Omega & \iint_{\Omega} p y d\Omega & -\iint_{\Omega} p x d\Omega \end{bmatrix}^T, \quad (13)$$

$$\mathbf{P}_c = \begin{bmatrix} \iint_{\Omega} p_c d\Omega & \iint_{\Omega} p_c y d\Omega & -\iint_{\Omega} p_c x d\Omega \end{bmatrix}^T, \quad (14)$$

$$\mathbf{P}_b = \begin{bmatrix} F_b & 0 & 0 \end{bmatrix}^T, \quad (15)$$

$$\mathbf{P} = \begin{bmatrix} F & M_x & M_y \end{bmatrix}^T. \quad (16)$$

\mathbf{P}_b has only one component, F_b , which comes from the pressure at the back of the stator and the preload. \mathbf{P}_b is equal to \mathbf{P}_f when $\mathbf{U}_s - \mathbf{U}_r = 0$ ($\mathbf{P}_c = 0$, as in a noncontacting seal at the designed state), which is how h_0 is determined. \mathbf{P} contains the faults modeled by the forces or moments exerted on the stator.

The six parameters listed in Table 1 correspond to the common faults in gas face seal applications. These parameters are considered for identification. The axial compactness change can be represented as an axial force F exerted on the stator. The asymmetry of the stator support is abstracted into the moments M_x and M_y . The stiffness and damping of the stator support may differ from their design performance because of assembly faults or aging. Thus, they are modeled using a scaling coefficient Q_s . The rotor tilt components γ_{rx0} and γ_{ry0} commonly occur because of assembly errors.

The fixed parameters are listed in the Appendix.

Table 1 Suspended causes and their respective ranges

Dimension index j	Parameter	Lower bound l_j	Upper bound u_j
1	Axial force exerted, F/N	-100	100
2	Moment exerted around X , $M_x/(\text{N} \cdot \text{m})$	-5	5
3	Moment exerted around Y , $M_y/(\text{N} \cdot \text{m})$	-5	5
4	Support stiffness and damping scale Q_s	0	100
5	Rotor tilt around X at $t = 0$, γ_{rx0}/mrad	-0.5	0.5
6	Rotor tilt around Y at $t = 0$, γ_{ry0}/mrad	-0.5	0.5

3 Establishment of surrogate models

3.1 Test design and initial database

A surrogate model must be established based on an existing database consisting of physical model results. The evaluation of a physical model is expensive. Thus, the evaluation sites must be carefully decided, a process known as test design. Popular test design methods for computer experiments include full factorial design, fractional factorial design, central composite design, Box–Behnken design, Latin hypercube design (LHD), maximum entropy design, minimum integrated mean square error design, maximin/minimax design, and ϕ_p -design [20,31–33]. The LHD approach has a good space-filling ability and can be implemented conveniently. Thus, it is used in the present work.

For M_0 sites with P variables ($P = 6$ here), the LHD scheme can be performed as follows [31].

A matrix \mathbf{D} is generated. Each column of which is a random permutation of 0 to $(M_0 - 1)$.

$$\mathbf{D} = \{d_{ij}\}_{M_0 \times P}. \quad (17)$$

Given that the parameters are uniformly distributed, the parameters are then randomly generated according to the level indices d_{ij} as

$$v_{ij} = l_j + \frac{d_{ij} + \delta_{ij}}{M_0} (u_j - l_j), \quad i = 1, 2, \dots, M_0, \quad j = 1, 2, \dots, P, \quad (18)$$

where v_{ij} contains F , M_x , M_y , Q_s , γ_{rx0} , and γ_{ry0} for ($j = 1, 2, \dots, 6$) at site i , and u_j and l_j are the upper and lower bounds, respectively. δ_{ij} is uniformly distributed from 0 to 1.

$M_0 = 100$ sites are designed using LHD, as summarized in Table 2. The initial database is generated by evaluating the physical model at these sites.

3.2 Kriging model

As a preprocessing step, standardization is first performed to provide the data with a mean of 0 and a variance of 1. This step is applied to each input dimension to give $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_P]$. For the output, V_0 is sampled at $Q = 50$ uniformly distributed time steps in a period, and

the standardization is applied to all time steps, where V_0 is sampled, instead of on each dimension, yielding $\mathbf{y} = [y_1, y_2, \dots, y_Q]$. The training dataset is denoted as $W = \{(\mathbf{x}^{(k)}, \mathbf{y}^{(k)}) | k = 1, 2, \dots, M\}$ or in matrix form as $\mathbf{X} = [\mathbf{x}^{(1)\top} \ \mathbf{x}^{(2)\top} \ \dots \ \mathbf{x}^{(M)\top}]^\top$ and $\mathbf{Y} = [\mathbf{y}^{(1)\top} \ \mathbf{y}^{(2)\top} \ \dots \ \mathbf{y}^{(M)\top}]^\top = [\mathbf{Y}_1 \ \mathbf{Y}_2 \ \dots \ \mathbf{Y}_Q]$.

Multiple methods for establishing surrogate models have been developed, including polynomial response surfaces, support vector machine regression, and the Kriging model [20]. The Kriging model has been frequently used in recent surrogate model studies because of its good performance for problems with strong linearity and small sample. Our preliminary studies demonstrated its advantages. The Kriging model has several variants; in particular, the universal Kriging model is employed here [34].

The output dimension y_j is assumed to be a random variable $Y_j(\mathbf{x})$ depending on \mathbf{x} , composed of a deterministic polynomial model $f_q(\mathbf{x})\beta_j$ and a random variable $Z_j(\mathbf{x})$ submitting to a Gaussian process of \mathbf{x} as

$$Y_j(\mathbf{x}) = f_q(\mathbf{x})\beta_j + Z_j(\mathbf{x}), \quad (19)$$

where $Z_j(\mathbf{x})$ has a mean of 0, a variance of σ_j^2 , and a covariance of

$$\text{Cov}(Z_j(\mathbf{x}), Z_j(\mathbf{x}')) = \sigma_j^2 R(\mathbf{x}, \mathbf{x}'). \quad (20)$$

The estimated output \hat{y}_j and its estimated variance \hat{s}_j^2 are

$$\hat{y}_j = \mathbf{f}\hat{\beta}_j + \mathbf{r}\mathbf{R}^{-1}(\mathbf{Y}_j - \mathbf{F}\hat{\beta}_j), \quad (21)$$

$$\hat{s}_j^2 = \sigma_j^2 \left(1 - [\mathbf{f} \ \mathbf{r}] \begin{bmatrix} 0 & \mathbf{F}^\top \\ \mathbf{F} & \mathbf{R} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{f}^\top \\ \mathbf{r}^\top \end{bmatrix} \right), \quad (22)$$

where R denotes the correlation function, which is used in the correlation matrix of the sample \mathbf{R} and the correlation vector for a general input \mathbf{r} , \mathbf{F} denotes the polynomial model output of the sample, \mathbf{f} denotes the polynomial model output for a general input, and β_j is the Kriging model coefficient found.

$$\mathbf{R} = \begin{bmatrix} R(\mathbf{x}^{(1)}, \mathbf{x}^{(1)}) & R(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) & \dots & R(\mathbf{x}^{(1)}, \mathbf{x}^{(M)}) \\ R(\mathbf{x}^{(2)}, \mathbf{x}^{(1)}) & R(\mathbf{x}^{(2)}, \mathbf{x}^{(2)}) & \dots & R(\mathbf{x}^{(2)}, \mathbf{x}^{(M)}) \\ \vdots & \vdots & \ddots & \vdots \\ R(\mathbf{x}^{(M)}, \mathbf{x}^{(1)}) & R(\mathbf{x}^{(M)}, \mathbf{x}^{(2)}) & \dots & R(\mathbf{x}^{(M)}, \mathbf{x}^{(M)}) \end{bmatrix}, \quad (23)$$

$$\mathbf{r} = [R(\mathbf{x}, \mathbf{x}^{(1)}) \ R(\mathbf{x}, \mathbf{x}^{(2)}) \ \dots \ R(\mathbf{x}, \mathbf{x}^{(M)})], \quad (24)$$

$$\mathbf{F} = \begin{bmatrix} f_q(\mathbf{x}^{(1)}) \\ f_q(\mathbf{x}^{(2)}) \\ \vdots \\ f_q(\mathbf{x}^{(M)}) \end{bmatrix}, \quad (25)$$

$$\mathbf{f} = f_q(\mathbf{x}), \quad (26)$$

Table 2 Sites designed by LHD for the initial dataset

No.	F/N	$M_x/(N \cdot m)$	$M_y/(N \cdot m)$	Q_s	γ_{rx0}/mrad	γ_{ry0}/mrad
1	-92.0	-2.02	-3.09	82.9	0.147	-0.101
2	-60.8	-4.16	2.95	10.7	0.052	0.498
3	63.5	-3.35	-1.86	12.6	-0.075	-0.231
...
99	17.3	-1.61	3.94	17.7	-0.144	-0.433
100	82.5	4.66	0.90	18.3	-0.290	-0.378

$$\hat{\beta}_j = (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{R}^{-1} \mathbf{Y}_j, \quad (27)$$

$$\hat{s}_j^2 = \frac{1}{M} \mathbf{Y}_j - \mathbf{F} \hat{\beta}_j^T \mathbf{R}^{-1} \mathbf{Y}_j - \mathbf{F} \hat{\beta}_j. \quad (28)$$

A Nugget effect [35] is appended to ensure stability: \mathbf{R}^{-1} is calculated as $(\mathbf{R} + \mathbf{N})^{-1}$, where a tiny diagonal matrix $\mathbf{N} = 10^{-3} \mathbf{I}$ is added before calculating the inverse.

The probability density of y_j at a targeted value y'_j can be estimated as

$$g(\mathbf{x}; \mathbf{X}, \mathbf{Y}_j, y'_j) = \frac{1}{\sqrt{2\pi}\hat{s}_j} \exp\left(-\frac{(y' - \hat{y}_j)^2}{2\hat{s}_j^2}\right). \quad (29)$$

A commonly used correlation function family is the anisotropic Gaussian exponential function, with scaling parameters $\boldsymbol{\theta} = [\theta_1 \ \theta_2 \ \cdots \ \theta_p]$ and exponential parameters $\mathbf{p} = [p_1 \ p_2 \ \cdots \ p_p]$, given by

$$R(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{1}{P} \sum_{i=1}^P (\theta_i \|x_i - x'_i\|^{p_i})\right). \quad (30)$$

Note that the parameters $\boldsymbol{\theta}$ and \mathbf{p} are consistently applied to all dimensions of \mathbf{y} . In the present study, they are restricted to the ranges $1/64 \leq \theta_i \leq 64$ (scaled logarithmically during optimization) and $1 \leq p_i \leq 2$, optimized by minimizing [36]

$$L(\boldsymbol{\theta}, \mathbf{p}) = \frac{M}{2} \sum_{j=1}^Q \ln \hat{s}_j^2 + \frac{Q}{2} \ln |\mathbf{R}|, \quad (31)$$

which is realized with an interior point method [37].

A static surrogate model can be established from the initial database. Eight test sites are randomly set for testing according to a uniform distribution, as summarized in Table 3.

The discrepancy measurement d is defined as

$$d(\mathbf{y}, \mathbf{y}') = \sqrt{\frac{\sum_{j=1}^Q (y_j - y'_j)^2}{\sum_{j=1}^Q (y_j)^2 + \sum_{j=1}^Q (y'_j)^2}}. \quad (32)$$

A comparison of the true output at the test sites and the output estimated by the surrogate model is presented in

Table 3 Sites randomly set for testing

No.	F/N	$M_x/(N \cdot m)$	$M_y/(N \cdot m)$	Q_s	γ_{rx0}/mrad	γ_{ry0}/mrad
1	-56.2	-4.53	1.79	67.9	0.435	-0.116
2	3.9	3.31	-4.65	5.3	0.030	0.171
3	-98.5	-1.17	-4.33	41.7	0.187	0.089
4	86.1	3.46	0.27	9.2	0.154	-0.084
5	40.2	4.10	2.62	26.2	-0.453	0.236
6	-34.4	1.33	2.56	99.1	-0.135	-0.253
7	96.5	2.23	2.53	65.2	-0.427	0.132
8	76.9	-2.27	-0.64	76.6	-0.022	-0.262

Fig. 4(a), along with the discrepancy defined above. The surrogate model reflects the overall trend for several cases but does not perform well in others. Although one can eventually find a database large enough to generate a surrogate model with satisfactory accuracy, this approach is against the initial purpose of decreasing the number of physical model evaluations. Given the small sample size (100 observations for six dimensions) and the strong nonlinearity of the problem, the result can be regarded as successful in terms of the sparse training set.

Although the accuracy is limited by the database size, the estimation for the distribution provides information about the uncertainty, which can be used to support decision-making. The 95% confidence range is plotted in Fig. 4(a). The ratio of the error to the estimated standard deviation $((\hat{y}_j - y_j)/\hat{s}_j)$ is plotted in Fig. 4(b), along with the probability density function of the standard normal distribution, which it should obey. The consistency between the two demonstrates that the probability estimation is reasonable.

We establish a probability-based criterion for determining the sites for physical model evaluations by taking advantage of this feature. Thus, an efficient method is presented by dynamically updating the surrogate model while searching for the desired input using the feedback acquired from the physical model.

4 Probability-based iterative method for solving the inverse problem

The present study aims to determine the causes of the abnormal signals. This inverse problem can be abstracted as a process of searching for an identified $\hat{\mathbf{x}}$ of the input \mathbf{x}' for an observed output y' , or

$$y' = \mathbf{y}(\mathbf{x}') \approx \mathbf{y}(\hat{\mathbf{x}}), \quad (33)$$

where $\mathbf{y}(\mathbf{x})$ denotes the physical model of the gas face seal, which fetches the input parameter \mathbf{x} and returns the standardized fluctuation of V_0 as output.

The genetic algorithm is considered because the mathematical nature of $\mathbf{y}(\mathbf{x})$ (i.e., multimodality, strong nonlinearity) is unclear, and its derivative cannot be analytically deduced [38,39]. Genetic algorithms can handle problems with weak mathematical conditions and undefined derivatives. However, they require a large number of iterations. Thus, they cannot be directly applied to the physical model. Enlightened by the idea of Bayesian optimization [40], we employed the surrogate models to construct the objective function for the genetic algorithm, which can be evaluated rapidly.

Therefore, a probability-based iterative method is established. Based on all available results from the physical model, the next point for evaluation is determined by

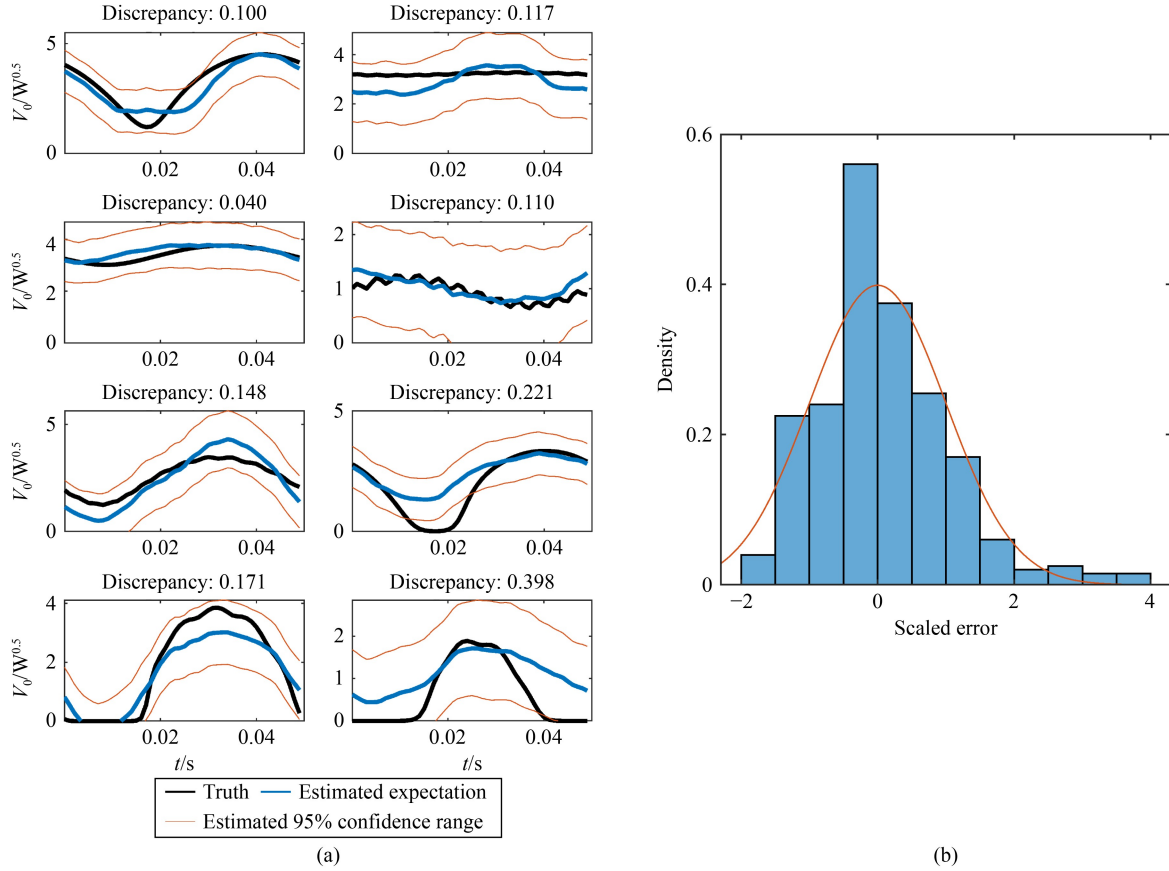


Fig. 4 Estimation of the surrogate model established on the initial database: (a) comparison of the true output and the estimated output; (b) density of the scaled error.

$$\mathbf{x}^* = \arg \max_{\mathbf{x} \in S} G(\mathbf{x}) = \arg \max_{\mathbf{x} \in S} \prod_{j=1}^Q g(\mathbf{x}; \mathbf{X}, \mathbf{Y}_j, \mathbf{y}'_j), \quad (34)$$

where S denotes the feasible region when searching for \mathbf{x} , as listed in Table 1.

The function $G(\mathbf{x})^{1/(PQ)}$ is used as the fitness function for the genetic algorithm, where the power $1/(PQ)$ is applied to prevent an extreme product of probability densities. Compared with using the discrepancy definition in Eq. (32) for the optimization, using this probabilistic expression encourages the algorithm to be sensitive to the error on sites with low randomness, allowing early escape from the local minimum.

Then, the physical model is evaluated at \mathbf{x}^* . The result acts as validation because the surrogate model may not always be sufficiently accurate, particularly when a limited database is used. If the iteration continues, the result is added to the database, thereby updating the surrogate model near the site that it previously identified.

The whole method proceeds in the following cycle (illustrated in Fig. 5):

(1) Prepare the initial database W_0 , as described in Section 3.1. The database is denoted by W , which is dynamically updated.

(2) Terminate and return the $(\mathbf{x}^{(i)}, \mathbf{y}^{(i)})$ in W with minimum $d(\mathbf{y}^{(i)}, \mathbf{y}')$ as the identification if

$\exists (\mathbf{x}^{(i)}, \mathbf{y}^{(i)}) \in W : d(\mathbf{y}^{(i)}, \mathbf{y}') \leq d_T$ or the iteration limit k_T has been exceeded; otherwise, proceed. Here, discrepancy termination threshold d_T and maximum iteration count k_T are set to be 0.02 and 100, respectively.

(3) Establish a surrogate model on the database W (as described in Section 3.2).

(4) Find the \mathbf{x}^* by optimizing Eq. (34) using the genetic algorithm.

(5) Evaluate $\mathbf{y}^* = \mathbf{y}(\mathbf{x}^*)$ at the maximum point found in step 4 and add the newly obtained $(\mathbf{x}^*, \mathbf{y}^*)$ into the database W . Then, go back to step (2).

The infeasible cost of evaluating the physical model numerous times is avoided using the above method, with the surrogate models used instead. Additionally, the relatively low accuracy of the surrogate models with the limited database is overcome using the reliable physical model results as feedback. Therefore, the method combines the rapid evaluation speed of the surrogate models with the accuracy of the physical model.

5 Results and discussion

5.1 Validation of simulated cases

The simulated outputs of the sites listed in Table 3 were

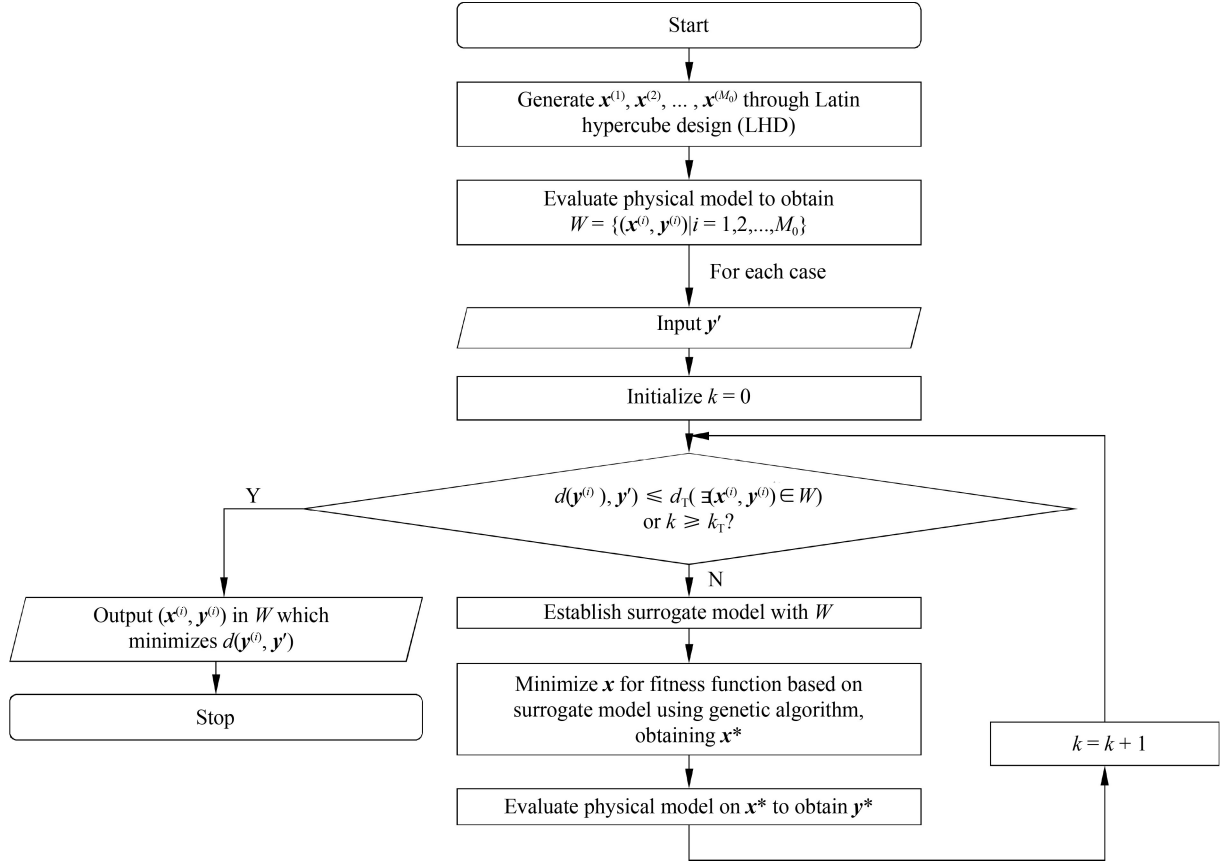


Fig. 5 Schematic flowchart of the inverse problem solution procedure.

provided to test the proposed method by identifying the axial force F , moment $M = \sqrt{M_x^2 + M_y^2}$, support stiffness and damping scale Q_s , and rotor tilt $\gamma_r = \sqrt{\gamma_{rx0}^2 + \gamma_{ry0}^2}$. The moments around the two axes were combined and the tilts around the two axes were also merged because identifying their phases does not make sense in practice.

The identification results of all cases are presented in Table 4 and Fig. 6. For cases 1, 2, 3, 5, and 6, early terminations occurred because highly similar outputs (with discrepancy less than $d_T = 0.02$) were found. For the other cases, the iterations were terminated because the maximum iteration count ($k_T = 100$) was reached. Moreover, the x corresponding to the minimum discrepancy found was returned. Figure 7 exhibits the iteration process toward the target y' in each case. Each iteration costs approximately 5.5 min on average, within which approximately 4.5 min is the cost for solving the physical model with an 8×3 GHz computer.

The method shows the following performance:

(1) The identifications for M are generally close to the respective ground truths. Except for case 2, where the absolute error reaches 1.51 N·m, all cases exhibit an absolute error of M less than 0.8 N·m.

(2) The identifications for F fall into great errors in cases 2 and 4. This scenario is mainly attributed to a weakness in following the high-order oscillation features

in contact signals.

(3) The identifications for Q_s and γ_r show considerable errors. However, the existing studies on seal dynamics concluded [16,41] that the support stiffness and damping affect the relative motion of seal faces by impeding the tracing ability, thereby scaling the effect of rotor tilt. This finding means that the effects of Q_s and γ_r on the contact are highly similar. The analysis of $Q_s\gamma_r$ is appended in Table 4 and Fig. 6, showing relatively small errors. Therefore, the weakness in recognizing Q_s and γ_r is due to their physical confounding.

(4) For five of the eight cases, the iteration is terminated to find a highly compatible match (three of which have less than 10 iterations). For the other cases, the requirement of early termination is not met; however, outputs with discrepancies of less than 0.05 are found within the limit of 100 iterations. This finding shows a balance between opportunism and robustness.

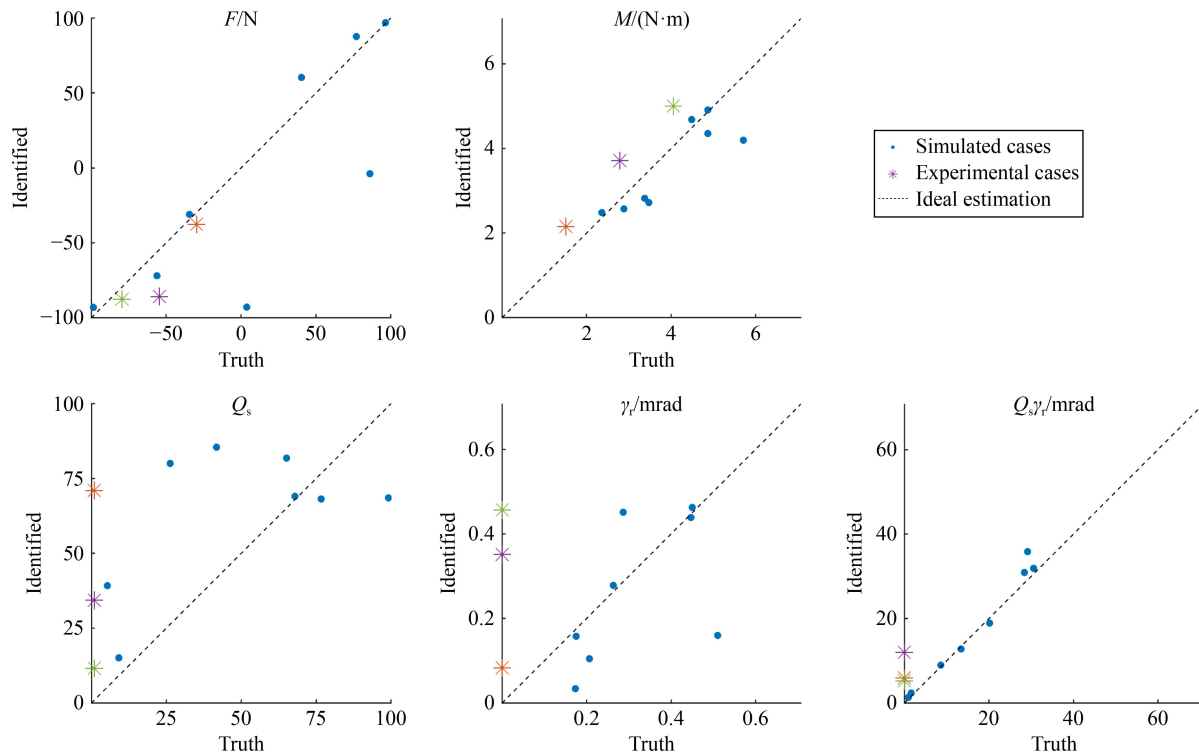
In summary, the validation using simulated cases shows that the method has good performance in identifying M and $Q_s\gamma_r$ in the task of rooting the causes of face contact in limited iterations.

5.2 Validation of experimental cases

Experimental data were employed to validate the performance of the proposed method. In our previous

Table 4 Validation of the simulated cases by comparing the true and identified parameters

No.	Parameter type	F/N	$M/(N \cdot m)$	Q_s	γ_r/mrad	$Q_s\gamma_r/\text{mrad}$	Iteration count	Discrepancy
1	Truth	-56.2	4.87	67.9	0.450	30.6	34	0.019
	Identified	-71.9	4.36	69.0	0.462	31.9		
	Error	-15.7	-0.51	1.1	0.012	1.4		
2	Truth	3.9	5.71	5.3	0.174	0.9	8	0.009
	Identified	-92.8	4.20	39.2	0.034	1.3		
	Error	-96.7	-1.51	33.8	-0.139	0.4		
3	Truth	-98.5	4.49	41.7	0.207	8.6	9	0.006
	Identified	-92.9	4.68	85.5	0.105	9.0		
	Error	5.5	0.20	43.7	-0.102	0.3		
4	Truth	86.1	3.47	9.2	0.175	1.6	43/100	0.045
	Identified	-3.8	2.72	15.1	0.158	2.4		
	Error	-89.9	-0.75	5.9	-0.017	0.8		
5	Truth	40.2	4.87	26.2	0.510	13.4	7	0.017
	Identified	60.4	4.91	80.0	0.160	12.8		
	Error	20.2	0.04	53.8	-0.350	-0.6		
6	Truth	-34.4	2.89	99.1	0.287	28.4	35	0.019
	Identified	-31.0	2.58	68.5	0.451	35.9		
	Error	3.3	-0.31	-30.6	0.165	2.5		
7	Truth	96.5	3.37	65.2	0.447	29.1	14/100	0.040
	Identified	96.9	2.82	81.8	0.439	35.9		
	Error	0.4	-0.55	16.6	-0.008	6.7		
8	Truth	76.9	2.36	76.6	0.263	20.2	56/100	0.048
	Identified	87.6	2.49	68.1	0.278	19.0		
	Error	10.7	0.13	-8.5	0.015	-1.2		

**Fig. 6** Actual values of the queried parameters and their respective identified values. M and $Q_s\gamma_r$ are identified well. Q_s and γ_r are difficult to be discriminated because of physical reasons.

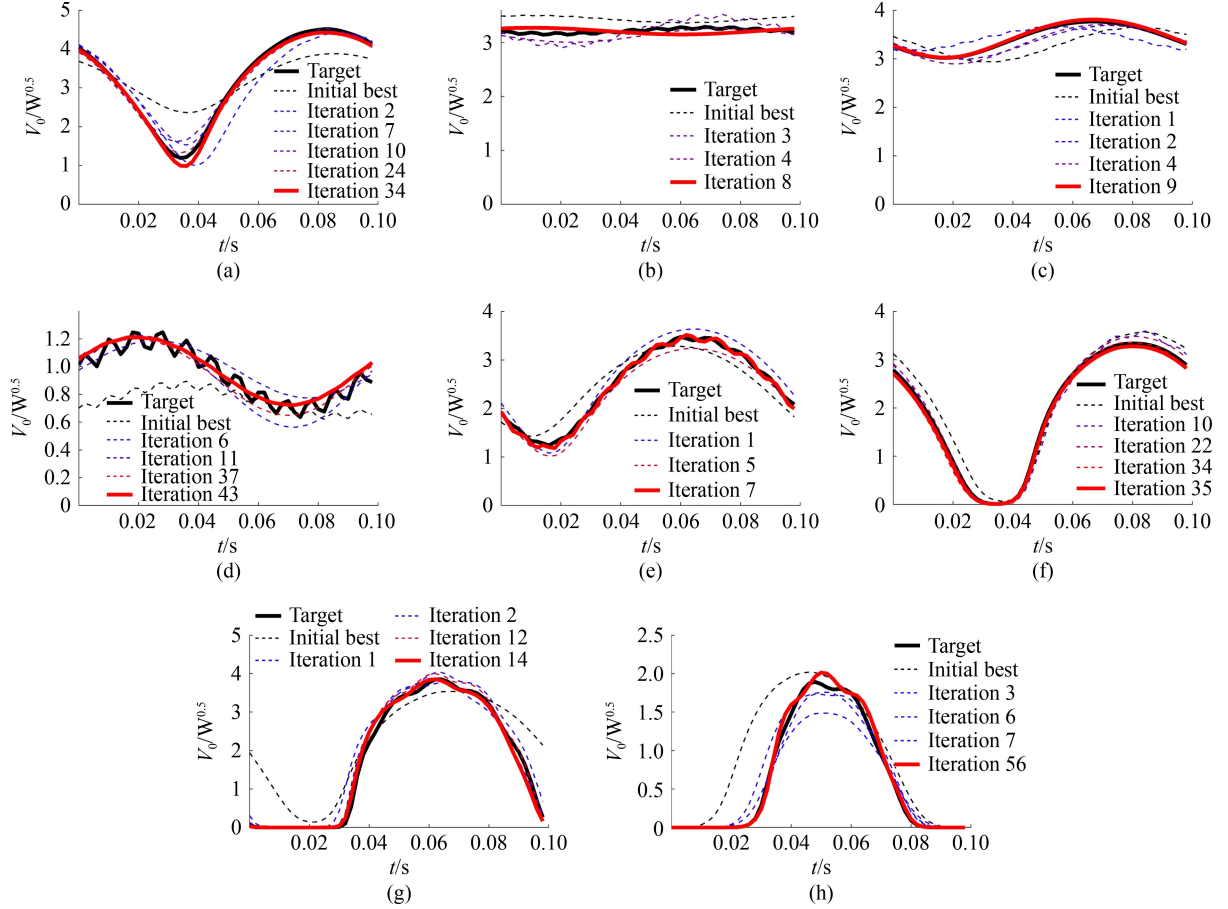


Fig. 7 Iteration process toward the targeted y' for the simulated cases. The dashed lines denote the best in the initial database and the improving steps. The black and red solid lines denote the searched target and the result found, respectively. (a)–(h) correspond to cases 1–8.

study [4], acoustic emission signals were obtained from a gas face seal running at 600 r/min and 0.4 MPa (see test rig in Fig. 8(a)). The signals were processed with a band-pass filter (centered at 160 kHz with 40 kHz width, which successfully filtered out the noises), and the RMS of each short-time wave was calculated.

Unlike validation on simulated cases, one cannot rule out unintentional errors in experimental processes. The strongest starting process data (found at 60 r/min, see Fig. 8(b)) were compared with the simulated low-speed operation results to determine K in Eq. (9) because this moment is when the contact is the strongest (the RMS is not that high because of the low speed). Thus, it is the point that is least affected by unintentional errors. The simulated V_0 here is $1.28 W^{0.5}$. Thus, K is estimated to be $21.1 \text{ mV} \cdot W^{-0.5}$.

A loading mechanism appended to the test rig exerted a force at an eccentric point on the stator, as shown in Fig. 8(a). In cases 1, 2, and 3, the axial forces of 29.7, 54.6, and 79.5 N were exerted, respectively, at a point 51 mm eccentric from the axis, thereby inducing a moment. These are the only known components. The unintentional components of F and M , as well as the value of Q_s and γ_r ,

are inaccessible. However, they are assumed to be minor factors here.

The results are presented in Table 5 and Figs. 6 and 9. From these results, we can state the following:

(1) The trend of the identified F and M is consistent with the exerted values. The systematic bias in M may be caused by unintentional factors in the experiments that aggravate the contact.

(2) The spire in the experimental observation of case 1 fails to patch with physical model results, represented by high discrepancy (0.195). This phenomenon probably comes from the face deformation and no longer appears explicitly in cases 2 and 3, where the intentionally controlled factors have great dominance.

(3) The values of $Q_s \gamma_r$ are approximate across the three cases (compared with its range of 0 to 70.7 mrad). This finding demonstrates the effectiveness of the identification by obtaining consistent results.

The proposed method was proved effective in real-life experimental cases on the basis of the above results. Given the advantage of using the physical model as a basis, one can reconstruct the detailed transient motion and lubrication states of the seal, as shown in Fig. 10.

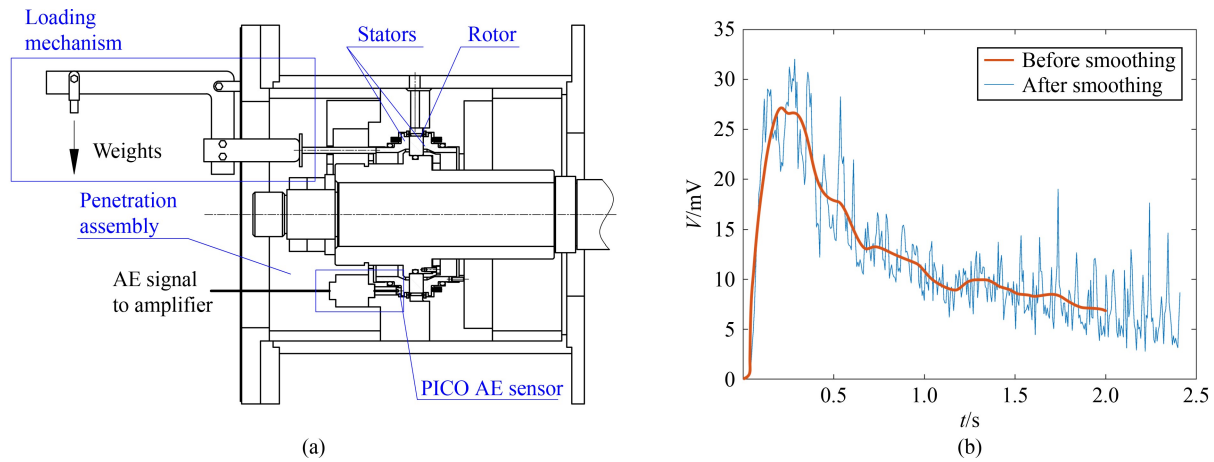


Fig. 8 (a) Test rig for exerting abnormalities and online monitoring and (b) filtered acoustic emission RMS during startup.

Table 5 Validation of the experimental cases by comparing the exerted and identified parameters and checking the consistency of parameters not changed through the experiments

No.	Parameter type	F/N	$M/(N \cdot m)$	$Q_s\gamma_r/mrad$	Iteration count	Discrepancy
1	Exerted	-29.7	1.51	—	88/100	0.195
	Identified	-37.5	2.16	6.0		
	Error	-7.8	0.65	—		
2	Exerted	-54.6	2.78	—	7/100	0.088
	Identified	-85.7	3.72	12.1		
	Error	-31.1	0.94	—		
3	Exerted	-79.5	4.05	—	65/100	0.047
	Identified	-87.5	5.01	5.3		
	Error	-8.0	0.96	—		

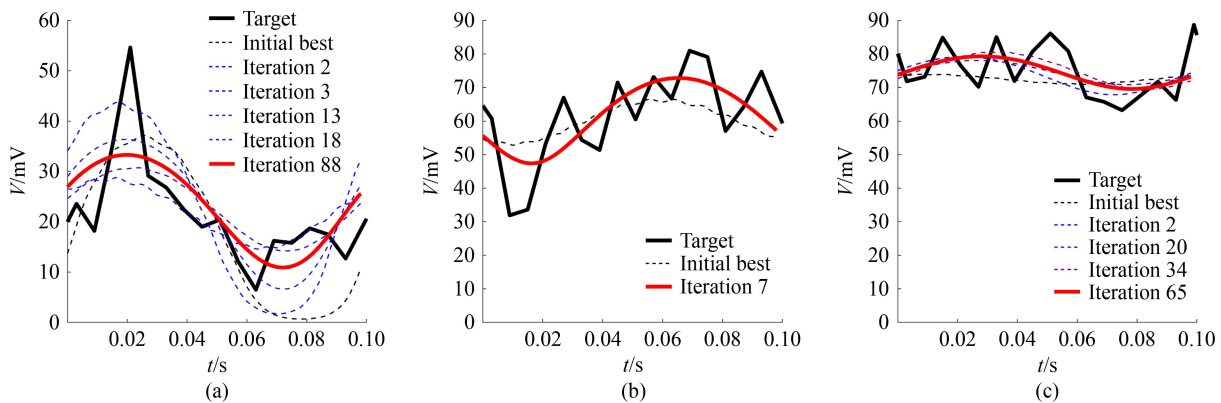


Fig. 9 Experimental data and results found by the method for each case. The dashed lines denote the initial and improving steps; the black and red solid lines denote the searched target and the result found, respectively. (a)–(c) correspond to cases 1–3.

This approach is useful for comprehensively explaining the result or establishing a digital twin system.

6 Conclusions

(1) A surrogate-model-assisted method is developed to solve the computational difficulty of the inverse problem abstracted from physics-based diagnosis tasks of gas face

seals by efficiently inferring the input for a targeted output. The method starts with a small initial database. It performs iterations that take advantage of the rapidity of the surrogate models and the reliability of the physical model.

(2) For the establishment of the surrogate models of the signal output, the Kriging model is validated as capable of using small multivariable databases to provide rapid estimations of the output distribution instead of merely an

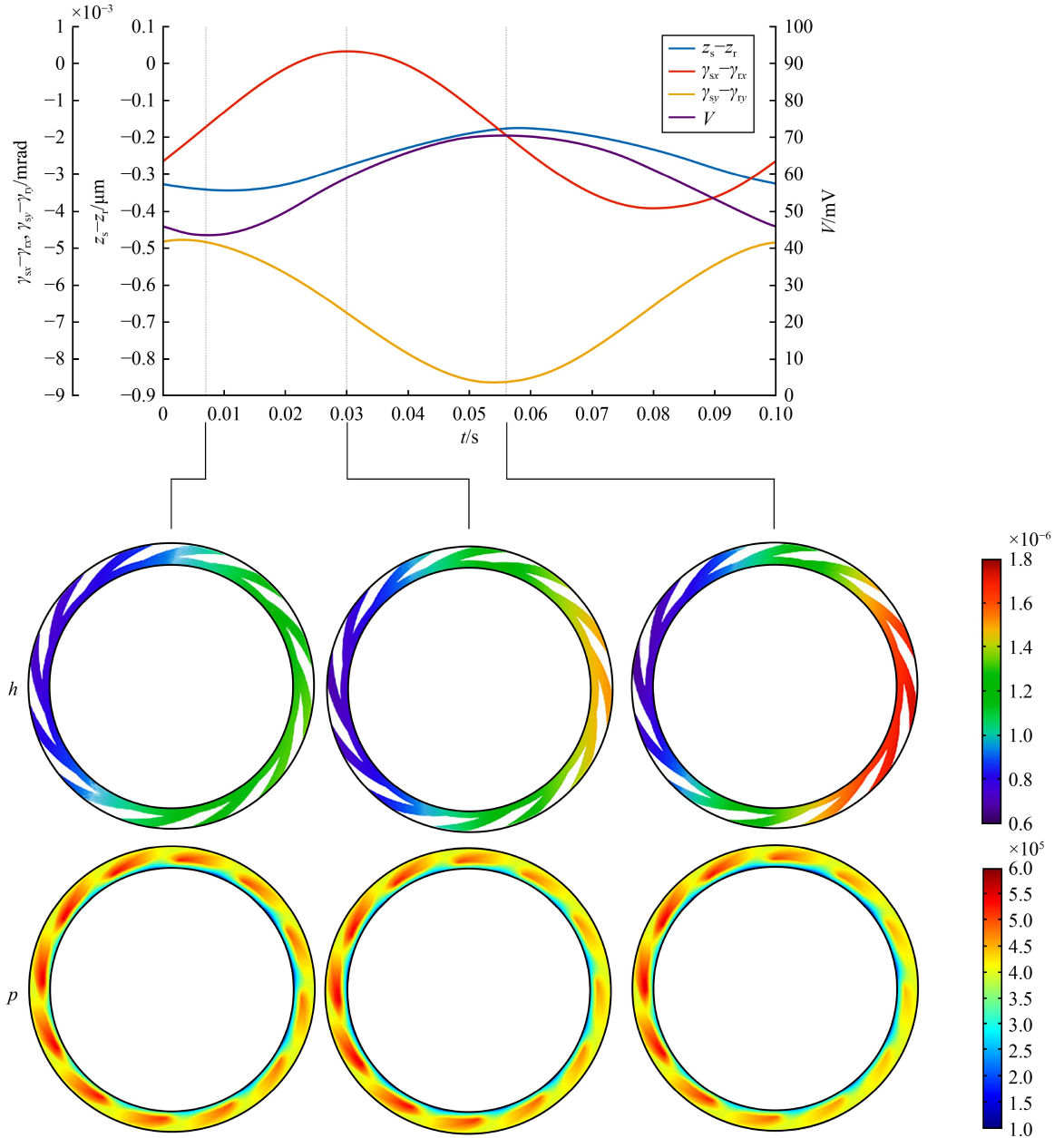


Fig. 10 Reconstructing the detailed relative motion and the separation and pressure distributions by taking experimental case 2 as an example.

expectation. Thus, it provides good support in determining the optimal evaluation sites of the physical model while an extensive database is inaccessible.

(3) The rapid surrogate models are employed for the construction of objective functions for large-scale iterations to find optimized sites for physical model evaluations, thus overcoming the infeasibility of applying massively iterative search algorithms to the physical model. The physical model is evaluated at the selected sites to provide detailed feedback to the surrogate models and overcome the limited accuracy of the surrogate models on small databases. The physical model is evaluated only a limited number of times using such an

iteration routine. However, the expert knowledge in the physical model is delivered by the surrogate models, and the reliability of the final result is ensured.

(4) The proposed method is validated on simulated and experimental cases of gas face seal dynamics. The results demonstrate that the method can effectively identify the fault parameters inducing the signal output in limited physical model evaluations.

The work presented in this paper provides a quantitative, explainable, and feasible approach for identifying the cause of gas face seal contact. The method is also suitable for diagnosis tasks in mechanical devices that face similar difficulties.

Appendix

The physical model of the gas face seal uses the parameters listed in Table A1.

Nomenclature

c_{sz}, c_{sy}	Axial and angular damping, respectively	g	Probability density of consistent output
c_{sz0}, c_{sy0}	Baseline axial and angular damping, respectively	h	Film thickness field
\mathbf{C}_s	Damping matrix of the stator	h_0	Film thickness on the land at $\mathbf{U}_s - \mathbf{U}_t = \mathbf{0}$
d	Discrepancy measurement	h_{grv}	Groove contribution to film thickness field
d_T	Terminating threshold for discrepancy measurement	h_{spr}	Depth of the spiral grooves
\mathbf{D}, d_{ij}	Random permutation matrix for Latin hypercube design ($i = 1, 2, \dots, M_0; j = 1, 2, \dots, P$)	J	Moment of inertia of the stator
E	Elastic modulus of the equivalent contact pair	k_{sz}, k_{sy}	Axial and angular stiffness, respectively
F	Exerted axial force modeling seal faults	k_{sz0}, k_{sy0}	Baseline axial and angular stiffness, respectively
F_b	Closing force at $\mathbf{U}_s - \mathbf{U}_t = \mathbf{0}$	k_T	Maximum iteration count
$\mathbf{f}_q(\mathbf{x})$	Different forms of products of the components of \mathbf{x} (of order not higher than q)	K	Coefficient of the acoustic emission transmission
		\mathbf{K}_s	Stiffness matrix of the stator
		l_j, u_j	Lower bound and upper bound of design variables ($j = 1, 2, \dots, P$), respectively
		$L(\boldsymbol{\theta}, \mathbf{p})$	Object function for training the parameters $(\boldsymbol{\theta}, \mathbf{p})$ of the correlation function
		m	Mass of the stator
		M	Moment exerted modeling seal faults
		M_0	Target count of sites for Latin hypercube design

Table A1 Fixed parameters of the gas face seal

Classification	Parameter	Value
Geometric parameters of the seal face	Outer radius, r_o/mm	61.6
	Inner radius, r_i/mm	51.6
	Spiral groove inner radius, r_{spr}/mm	55.5
	Spiral groove angle, $\beta_{spr}/(^{\circ})$	15
	Groove depth, $h_{spr}/\mu\text{m}$	6
	Groove count, N_{spr}	12
	Stator mass, m/kg	0.2
Mechanical parameters	Stator moment of inertia, $J/(\text{kg} \cdot \text{m}^2)$	3.5×10^{-4}
	Baseline axial stiffness, $k_{sz0}/(\text{N} \cdot \text{m}^{-1})$	1.64×10^4
	Baseline axial damping, $c_{sz0}/(\text{N} \cdot \text{s} \cdot \text{m}^{-1})$	1×10^3
	Baseline angular stiffness, $k_{sy0}/(\text{N} \cdot \text{m} \cdot \text{rad}^{-1})$	31
	Baseline angular damping, $c_{sy0}/(\text{N} \cdot \text{m} \cdot \text{s} \cdot \text{rad}^{-1})$	1.4
	Standard deviation of surface height, $\sigma_a/\mu\text{m}$	0.177
	Standard deviation of asperity height, $\sigma_s/\mu\text{m}$	0.170
Equivalent contact pair parameters	Mean of asperity height, $\bar{w}_s/\mu\text{m}$	0.114
	Asperity density, $\eta_s/\mu\text{m}^{-2}$	0.133
	Average radius of asperity tips, $R_s/\mu\text{m}$	3.02
	Elastic modulus, E/GPa	151
	Gas dynamic viscosity, $\mu/(\text{Pa} \cdot \text{s})$	1.8×10^{-5}
	Gas density under 0.1 MPa pressure, $\rho/(\text{kg} \cdot \text{m}^{-3})$	1.2
	Gas pressure at inner radius, p_i/MPa	0.1
Operating condition parameters	Gas pressure at outer radius, p_o/MPa	0.4
	Closing force, P_b/N	1320.1
	Rotation speed, $n_{\omega}/(\text{r} \cdot \text{min}^{-1})$	600

M_x, M_y	Moment around X and Y axes exerted modeling seal faults, respectively	$z_j(\mathbf{x})$	Random variables submitting to the Gaussian process of \mathbf{x} ($j = 1, 2, \dots, Q$)
M_s	Inertia matrix of the stator	β_j	Coefficient toward dimension j of the output vector
N	Nugget effect matrix	γ_r	Tilt angle of the rotor
P	Film pressure field (absolute pressure)	γ_{rx}, γ_{ry}	Tilt angle around X and Y axes of the rotor, respectively
P_c	Solid contact pressure field	$\gamma_{rx0}, \gamma_{ry0}$	Tilt angle around X and Y axes of the rotor at $t = 0$, respectively
p_i, p_o	Boundary pressure on inner and outer radii, respectively	γ_{sx}, γ_{sy}	Tilt angle around X and Y axes of the stator, respectively
\mathbf{p}, p_i	Exponential parameters as vector and components ($i = 1, 2, \dots, P$) of the correlation function, respectively	δ_{ij}	Random variables for Latin hypercube design ($i = 1, 2, \dots, M_0; j = 1, 2, \dots, P$)
P	Count of design variables	η_s	Asperity density of equivalent contact pair
\mathbf{P}	Generalized force exerted modeling seal faults	ϕ	Probability density function of the standard normal distribution
\mathbf{P}_b	Generalized closing force at $\mathbf{U}_s - \mathbf{U}_r = \mathbf{0}$	φ	Flow factor for the average flow Reynolds equation
\mathbf{P}_c	Generalized force of solid contact	μ	Dynamic viscosity
\mathbf{P}_f	Generalized force of the fluid film	σ_a	Standard deviation of the surface height of the equivalent contact pair
Q	Count of time steps	σ_j	Standard variance of the Gaussian process ($j = 1, 2, \dots, Q$)
Q_s	Scaling coefficient of the support stiffness and damping	σ_s	Standard deviation of the asperity height of the equivalent contact pair
r, θ	Polar coordinates	Ω	Area of the tribol-pair
r_i, r_o	Inner and outer radii of the tribol-pair, respectively	ω	Rotation speed
R	Correlation function	θ, θ_i	Scaling parameters as vector/components ($i = 1, 2, \dots, P$) of the correlation function
R_s	Average radius of asperity tips of the equivalent contact pair		
\mathbf{R}	Correlation matrix of the training samples		
\hat{s}_j	Estimated standard variance of the dimension j of the Kriging model output ($j = 1, 2, \dots, Q$)		
t	Time		
\mathbf{U}_r	Generalized displacement of the rotor		
\mathbf{U}_s	Generalized displacement of the stator		
v_{ij}	Design results for the P variables on M_0 sites by Latin hypercube design		
V	Root-mean-square value of the monitored acoustic emission signal		
V_0	Root-mean-square value of the generated acoustic emission signal		
w	Surface height of the equivalent contact pair		
\bar{w}_s	Mean of asperity height of the equivalent contact pair		
W	Dynamically updated database for surrogate models		
x, y	Orthogonal coordinates		
\mathbf{x}, x_i	Input vector and components ($i = 1, 2, \dots, P$) to the surrogate models, respectively		
$\mathbf{X}, \mathbf{x}^{(k)}$	Dataset of input to the surrogate models as matrix and samples ($k = 1, 2, \dots, M$), respectively		
\mathbf{y}, y_j	Output vector and components ($j = 1, 2, \dots, Q$) to the surrogate models, respectively		
$\hat{\mathbf{y}}, \hat{y}_j$	Estimated output vector and components ($j = 1, 2, \dots, Q$) to the surrogate models, respectively		
$\mathbf{Y}, \mathbf{y}^{(k)}$	Dataset of output to the surrogate models as matrix and samples ($k = 1, 2, \dots, M$), respectively		
z_r	Axial displacement of the rotor		
z_s	Axial displacement of the stator		

Acknowledgements This work was supported by the National Key R&D Program of China (Grant No. 2020YFB2010000) and the National Natural Science Foundation of China (Grant No. U1737209). None of the funding bodies influenced the study at any stage.

References

1. Fan Y E, Gu F S, Ball A. A review of the condition monitoring of mechanical seals. In: Proceedings of ASME the 7th Biennial Conference on Engineering Systems Design & Analysis. Manchester, 2004, 3: 179–184
2. Phillips R L, Jacobs L E, Merati P. Experimental determination of the thermal characteristics of a mechanical seal and its operating environment. Tribology Transactions, 1997, 40(4): 559–568
3. DiRusso E. Film Thickness Measurement for Spiral Groove and Rayleigh Step Lift Pad Self-Acting Face Seals. NASA TP-2058, 1982
4. Huang W F, Lin Y B, Liu Y, Liu X F, Gao Z, Wang Y M. Face rub-impact monitoring of a dry gas seal using acoustic emission. Tribology Letters, 2013, 52(2): 253–259
5. Towsyfyhan H, Gu F S, Ball A D, Liang B. Tribological behaviour diagnostic and fault detection of mechanical seals based on acoustic emission measurements. Friction, 2019, 7(6): 572–586
6. Daraz A, Alabied S, Zhen D, Gu F S, Ball A D. Detection and diagnosis of mechanical seal faults in centrifugal pumps based on acoustic measurement. In: Ball A, Gelman L, Rao B K N, eds.

- Advances in Asset Management and Condition Monitoring. Cham: Springer, 2020, 963–975
7. Medina-Arenas M, Sopp F, Stolle J, Schley M, Kamieth R, Wassermann F. Measurement and analysis of inadequate friction mechanisms in liquid-buffered mechanical seals utilizing acoustic emission technique. *Vibration*, 2021, 4(1): 263–283
 8. Feldman Y, Kligerman Y, Etsion I. Stiffness and efficiency optimization of a hydrostatic laser surface textured gas seal. *Journal of Tribology*, 2007, 129(2): 407–410
 9. Blasiak S, Zahorulko A V. A parametric and dynamic analysis of non-contacting gas face seals with modified surfaces. *Tribology International*, 2016, 94: 126–137
 10. Jiang J B, Zhao W J, Peng X D, Li J Y. A novel design for discrete surface texture on gas face seals based on a superposed groove model. *Tribology International*, 2020, 147: 106269
 11. Zhang Z, Li X H. Acoustic emission monitoring for film thickness of mechanical seals based on feature dimension reduction and cascaded decision. In: *Proceedings of 2014 the 6th International Conference on Measuring Technology and Mechatronics Automation*. Zhangjiajie: IEEE, 2014, 64–70
 12. Yin Y, Liu X F, Huang W F, Liu Y, Hu S T. Gas face seal status estimation based on acoustic emission monitoring and support vector machine regression. *Advances in Mechanical Engineering*, 2020, 12(5): 168781402092132
 13. Zhang Z H, Min F, Chen G S, Shen S P, Wen Z C, Zhou X B. Tripartition state alphabet-based sequential pattern for multivariate time series. *Cognitive Computation*, 2021 (in press)
 14. Li T Y, Qian Z J, Deng W, Zhang D Z, Lu H H, Wang S H. Forecasting crude oil prices based on variational mode decomposition and random sparse Bayesian learning. *Applied Soft Computing*, 2021, 113: 108032
 15. Ran X J, Zhou X B, Lei M, Tepsan W, Deng W. A novel *k*-means clustering algorithm with a noise algorithm for capturing urban hotspots. *Applied Sciences*, 2021, 11(23): 11202
 16. Miller B A, Green I. Numerical formulation for the dynamic analysis of spiral-grooved gas face seal. *Journal of Tribology*, 2001, 123(2): 395–403
 17. Green I. A transient dynamic analysis of mechanical seals including asperity contact and face deformation. *Tribology Transactions*, 2002, 45(3): 284–293
 18. Zirkelback N, San Andre's L. Effect of frequency excitation on force coefficients of spiral groove gas seals. *Journal of Tribology*, 1999, 121(4): 853–861
 19. Miller B, Green I. Constitutive equations and the correspondence principle for the dynamics of gas lubricated triboelements. *Journal of Tribology*, 1998, 120(2): 345–352
 20. Simpson T W, Peplinski J D, Koch P H, Allen J K. Metamodels for computer-based engineering design: survey and recommendations. *Engineering with Computers*, 2001, 17(2): 129–150
 21. Bhosekar A, Ierapetritou M. Advances in surrogate based modeling, feasibility analysis, and optimization: a review. *Computers & Chemical Engineering*, 2018, 108: 250–267
 22. Han Z H. Kriging surrogate model and its application to design optimization: a review of recent progress. *Acta Aeronautica et Astronautica Sinica*, 2016, 37(11): 3197–3225
 23. Kontogiannis S G, Savill M A. A generalized methodology for multidisciplinary design optimization using surrogate modelling and multifidelity analysis. *Optimization and Engineering*, 2020, 21(3): 723–759
 24. Garg A, Liu C, Jishnu A K, Gao L, Le Phung M L, Tran V M. A Thompson sampling efficient multi-objective optimization algorithm (TSEMO) for lithium-ion battery liquid-cooled thermal management system: study of hydrodynamic, thermodynamic, and structural performance. *Journal of Electrochemical Energy Conversion and Storage*, 2021, 18(2): 021009
 25. Tang H S, Ren Y, Kumar A. Optimization tool based on multi-objective adaptive surrogate modeling for surface texture design of slipper bearing in axial piston pump. *Alexandria Engineering Journal*, 2021, 60(5): 4483–4503
 26. Liu Y, Liu Q X, Yin M, Yin G F. Dynamic analysis and structure optimization of a floating ring system in dry gas seal. *Journal of Advanced Mechanical Design, Systems and Manufacturing*, 2018, 12(7): JAMDSM0128
 27. Patir N, Cheng H S. An average flow model for determining effects of three-dimensional roughness on partial hydrodynamic lubrication. *Journal of Tribology*, 1978, 100(1): 12–17
 28. Fan Y B, Gu F S, Ball A. Modeling acoustic emissions generated by sliding friction. *Wear*, 2010, 268(5–6): 811–815
 29. Greenwood J A, Williamson J B P. Contact of nominally flat surfaces. *Proceedings of the Royal Society of London*, 1966, 295(1442): 300–319
 30. Yin Y, Hu S T, Huang W F, Liu X F, Liu Y, Wang Z X. A bi-Gaussian acoustic emission model for sliding friction. *IOP Conference Series: Material Science and Engineering*, 2019, 686(1): 012026
 31. Mckay M D, Beckman R J, Conover W J. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 1979, 21: 239–245
 32. Morris M D, Mitchell T J. Exploratory designs for computational experiments. *Journal of Statistical Planning and Inference*, 1995, 43(3): 381–402
 33. Lee B W, Peterson J J, Yin K H, Stockdale G S, Liu Y C, O'Brien A. System model development and computer experiments for continuous API manufacturing. *Chemical Engineering Research & Design*, 2020, 156: 495–506
 34. Sacks J, Welch W J, Mitchell T J, Wynn H P. Design and analysis of computer experiments. *Statistical Science*, 1989, 4(4): 409–423
 35. Yin J, Ng S H, Ng K M. Kriging metamodel with modified nugget-effect: the heteroscedastic variance case. *Computers & Industrial Engineering*, 2011, 61(3): 760–777
 36. Toal D J J, Bressloff N W, Keane A J. Kriging hyperparameter tuning strategies. *AIAA Journal*, 2008, 46(5): 1240–1252
 37. Byrd R H, Gilbert J C, Nocedal J. A trust region method based on interior point techniques for nonlinear programming. *Mathematical Programming*, 2000, 89(1): 149–185
 38. Herrera F, Lozano M, Verdegay J L. Tackling real-coded genetic algorithms: operators and tools for behavioural analysis. *Artificial Intelligence Review*, 1998, 12(4): 265–319
 39. Deng W, Zhang X X, Zhou Y Q, Liu Y, Zhou X B, Chen H L, Zhao H M. An enhanced fast non-dominated solution sorting genetic algorithm for multi-objective problems. *Information*

- Sciences, 2022, 585: 441–453
40. Jones D R, Schonlau M, Welch W J. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 1998, 13(4): 455–492
41. Hu S T, Huang W F, Liu X F, Wang Y M. Influence analysis of secondary O-ring seals in dynamic behavior of spiral groove gas face seals. *Chinese Journal of Mechanical Engineering*, 2016, 29(3): 507–514