



## The quasi-fiducial model selection for Kriging model

Chen Fan, Shuqin Zhang & Xinmin Li

**To cite this article:** Chen Fan, Shuqin Zhang & Xinmin Li (2025) The quasi-fiducial model selection for Kriging model, Statistical Theory and Related Fields, 9:3, 285-296, DOI: [10.1080/24754269.2025.2537484](https://doi.org/10.1080/24754269.2025.2537484)

**To link to this article:** <https://doi.org/10.1080/24754269.2025.2537484>



© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 05 Aug 2025.



Submit your article to this journal [↗](#)



Article views: 70



View related articles [↗](#)



View Crossmark data [↗](#)



# The quasi-fiducial model selection for Kriging model

Chen Fan, Shuqin Zhang and Xinmin Li

School of Mathematics and Statistics, Qingdao University, Qingdao, People's Republic of China

## ABSTRACT

Kriging models are widely employed due to their simplicity and flexibility in a variety of fields. To gain more distributional information about the unknown parameters, we focus on constructing the fiducial distribution of Kriging model parameters. To solve the challenge of constructing the fiducial marginal distribution for the spatially related parameter, we substitute the Bayesian posterior distribution for the fiducial distribution of this spatially related parameter and present a quasi-fiducial distribution for Kriging model parameters. A Gibbs sampling algorithm is given to get the samples of the quasi-fiducial distribution. Then a model selection criterion based on the quasi-fiducial distribution is proposed. Numerical studies demonstrate that the proposed method is superior to the Lasso and Elastic Net.

## ARTICLE HISTORY

Received 27 August 2024

Accepted 11 July 2025

## KEYWORDS

Kriging model; fiducial inference; slice sampling; model selection

## 1. Introduction

With the rapid development of science and technology, computer experiments have emerged as a prominent method, increasingly replacing physical experiments due to their cost-effectiveness and significant impact on scientific research. To better simulate physical phenomena, Sacks et al. (1989) proposed modelling the response of computer experiments as a stochastic process model—Kriging model, which allows for uncertainty estimation in deterministic computer simulations. This model was also utilized by Welch et al. (1992) in their research to avoid overfitting and by Simpson et al. (2001) to fit higher-order and lower-dimensional functions. Besides, it is attractive for Kriging model to model the computer simulation response as a spatially related Gaussian process and suitable for many kinds of modelling.

The Kriging model generally consists of two components: a mean function and a Gaussian process (typically a stationary process). Kriging model is called ordinary Kriging (OK) model when the mean function  $m(x)$  has only one constant, and it is called universal Kriging (UK) model when the  $m(x)$  is partially assumed to have some known variables. The former is parametric while the latter is semiparametric. The UK model allows the mean function to include known basis functions, making the UK model more flexible in handling complex problems. The number of input variables can be large, but not all variables are useful. In other words, not all input variables significantly affect the response variable. It is meaningful to identify variables with a significant impact on the response as it simplifies the model to

**CONTACT** Xinmin Li ✉ [xmli@qdu.edu.cn](mailto:xmli@qdu.edu.cn) School of Mathematics and Statistics, Qingdao University, No. 308 Ningxia Road, Fushan Campus, Qingdao 266071, People's Republic of China

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

avoid overfitting, and select significant variables to enhance model interpretability and save experimental resources.

The identifiability problem has been addressed by many researchers (Hodges & Reich, 2010; Paciorek, 2010; Tuo & Wu, 2015). For example, in model calibration, due to the existence of parameter identification problems, we cannot correctly interpret the actual meaning of each variable from the coefficient of the mean function, and sometimes we even misrecognize the part of the mean function, so as to get the opposite conclusion. There are some relevant research results on the model selection of mean function. Joseph et al. (2008) provided the blind Kriging method which integrates a Bayesian forward selection procedure into the Kriging model. Hung (2011) proposed a penalized blind Kriging approach. Recently, Park (2021) and Zhao et al. (2023) proposed the penalty likelihood method.

Distinguished from the frequentist and Bayesian approaches, fiducial inference has become a popular and highly regarded leading method, which extends fiducial inference by using the inverse function of the data-generating equation to define an estimation of distribution of parameters. Hannig (2009) provided a detailed exposition and analysis of generalized fiducial inference (GFI). Furthermore, Hannig (2013) demonstrated the asymptotic exactness of fiducial confidence sets under general conditions. Cui et al. (2024) proposed a semiparametric fiducial approach recently which was designed for survival analysis models. As fiducial theory develops, model selection methods based on fiducial inference have also emerged, providing probability guarantees for each candidate model being the optimal model. Williams and Hannig (2019) and Zhao et al. (2023) proposed some fiducial model selection criteria to deal with high-dimensional linear regression, respectively.

The remainder of this paper is organized as follows. Section 2 presents the quasi-fiducial distribution for universal Kriging models. In Section 3, we introduce the theoretical framework and algorithm steps of quasi-fiducial model selection for Kriging models. Section 4 compares the performance of the new method with traditional model selection methods: Lasso and Elastic Net (EN) through numerical simulations and case studies. Section 5 concludes this paper.

## 2. Methodology

### 2.1. The universal Kriging model

Suppose inputs  $x_i$  is a design point on the  $d$ -dimensional test regions and  $y_i = y(x_i)$  is the corresponding output response,  $i = 1, \dots, n$ . Let  $x = (x_1, \dots, x_n)$ , and then the universal Kriging (UK) model is defined as

$$y_i = m(x_i) + z(x_i) = f(x_i)^\top \beta + z(x_i), \quad i = 1, \dots, n, \quad (1)$$

where  $m(\cdot) = f(\cdot)^\top \beta$  is the mean function,  $f(\cdot) = (f_1(\cdot), \dots, f_p(\cdot))^\top$  is a known basis function vectors,  $\beta = (\beta_1, \dots, \beta_p)^\top$  is an unknown regression coefficients, and  $z(x_i)$  is a zero-mean Gaussian process with covariance  $\sigma^2 r_\phi(x_i, x_j)$  between  $x_i$  and  $x_j$ . There are many different kinds of basis functions. For example, we could set  $f(x_i) = x_i = (x_{i1}, \dots, x_{id})^\top$  for simplicity.

Let  $Y = (y_1, \dots, y_n)^\top$ ,  $F = (f(x_1)^\top, \dots, f(x_n)^\top)^\top$  and  $Z(x) = (z(x_1)^\top, \dots, z(x_n)^\top)^\top$ . UK model (1) could be rewritten as

$$Y = F\beta + Z(x),$$

where  $Z(x) \sim N(0, \sigma^2 R(\phi))$  with  $R(\phi) = [r_\phi(x_i, x_j)]$ . The correlation function  $r_\phi(x_i, x_j)$  is related to the spatially relevant parameter  $\phi$ , and the commonly used one includes the power exponential correlation function, the Matérn correlation function and the Gaussian correlation function. Although several correlation functions are available, we consider the power exponential correlation function in this paper, which is perhaps the most commonly used correlation function,

$$R(\phi, x_i, x_j) = \text{corr}(Z(x_i), Z(x_j)) = \prod_{k=1}^d \phi_k^{|x_{ik} - x_{jk}|^2},$$

where  $\phi = (\phi_1, \dots, \phi_d) \in (0, 1)^d$ .

To effectively capture the overall trend, important regression coefficients  $\beta$  should be selected for a more accurate mean. Let  $\mathcal{M} = \{M \mid M \subset \{1, 2, \dots, p\}\}$  be the set of all models and  $|M|$  be the cardinality of the set  $M$ , and  $\theta_M = (\beta_M^\top, \sigma_M^2, \phi_M^\top)^\top$ . For the candidate model  $M$ , we have

$$Y = F_M \beta_M + Z_M, \quad (2)$$

where  $Z_M \sim N(0, \sigma_M^2 R_M)$  with  $R_M = R(\phi_M) = [r_{\phi_M}(x_i, x_j)]$ , the design matrix  $F_M$  is defined as the matrix composed of only those columns of  $F$  corresponding to the index set  $M$ , and  $\beta_M$  is a  $|M|$ -dimensional vector indexed by the subset  $M$ .

The likelihood of the Kriging model (2) is

$$L(\theta_M \mid y) = (2\pi\sigma_M^2)^{-\frac{n}{2}} |R_M|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2\sigma_M^2} (y - F_M \beta_M)^\top R_M^{-1} (y - F_M \beta_M) \right\}. \quad (3)$$

Thus the MLE of  $\theta_M$  is

$$\hat{\beta}_M = (F_M^\top R_M^{-1} F_M)^{-1} F_M^\top R_M^{-1} Y, \quad (4)$$

$$\hat{\sigma}_M^2 = \frac{(Y - F_M \hat{\beta}_M)^\top R_M^{-1} (Y - F_M \hat{\beta}_M)}{n}, \quad (5)$$

$$\hat{\phi}_M = \underset{\phi}{\text{argmin}} \{ \log |R_M| + n \log(\hat{\sigma}_M^2) \}. \quad (6)$$

The MLE algorithm of parameters is shown in Algorithm 1.

---

**Algorithm 1** (The MLE algorithm of parameters in universal Kriging model)

---

**Input:** Data  $(x_i, y_i)_{i=1}^n$ , randomly given initial values  $\hat{\phi}_0$ ,  $t = 0$ ; **Step 1:** Substitute  $\hat{\phi}_t$  into Equations (4) and (5) to get  $\hat{\beta}_{t+1}$  and  $\hat{\sigma}_{t+1}^2$ . **Step 2:** Substitute last obtained  $\hat{\beta}_{t+1}$  and  $\hat{\sigma}_{t+1}^2$  into Equation (6) to get  $\hat{\phi}_{t+1}$ , and then we get a group of estimator  $(\hat{\beta}_{t+1}, \hat{\sigma}_{t+1}^2, \hat{\phi}_{t+1})$ . **Step 3:**  $t = t + 1$ . **Step 4:** Repeat Steps 1–3 until  $\hat{\beta}_M$ ,  $\hat{\sigma}_M^2$  and  $\hat{\phi}_M$  all have converged.

---

Based on the result, we can get the best linear unbiased prediction (BLUP) of the new sample  $x_*$

$$\hat{y}(x_*) = f(x_*) \hat{\beta}_M + \hat{r}_M^\top(x_*) \hat{R}_M^{-1} (y - F_M \hat{\beta}_M),$$

where  $\hat{R}_M = R(\hat{\phi}_M)$  and  $\hat{r}_M(x_*) = (r_{\hat{\phi}_M}(x_*, x_1), \dots, r_{\hat{\phi}_M}(x_*, x_n))^\top$ .

## 2.2. The quasi-fiducial distribution for Kriging model

The fiducial inference was originally introduced by Fisher, which aims to construct a statistical distribution for the parameter space without the prior information. The key point is that any observation  $Y$  can be regarded as the outcome of the structural equation

$$Y = G(M, U_M, \theta_M), \quad M \in \mathcal{M}, \theta_M \in \Theta_M, \quad (7)$$

where the distribution of  $U_M$  is completely known and independent of  $\theta_M$ . If  $Y = y$  is observed and the inverse  $G^{-1}$  for  $\theta_M$  exists, a distribution of  $\theta_M$  can be defined by inverting the structural Equation (7) as follows:

$$\theta_M = G^{-1}(M, y, u_M).$$

A set of random samples called the fiducial samples of  $\theta_M$  is obtained if  $U_M$  is repeatedly sampled.

Let  $\tilde{Y} = R_M^{-\frac{1}{2}} Y$ ,  $\tilde{F}_M = R_M^{-\frac{1}{2}} F_M$  and  $\varepsilon_M = R_M^{-\frac{1}{2}} Z_M$ . Equation (2) could be rewritten as

$$\tilde{Y}_M = \tilde{F}_M \beta_M + \varepsilon_M, \quad (8)$$

where  $\varepsilon_M \sim N(0, \sigma^2 I_n)$ .

When  $\phi_M$  is known, the Kriging model (8) is indeed a linear regression model. Zhao et al. (2023) proposed a fiducial density of parameters based on the sufficient statistic. Similar to the results in Zhao et al. (2023), we have

$$\begin{aligned} r(\beta_M | \sigma_M^2, \phi_M, y) &= (2\pi\sigma_M^2)^{-\frac{|M|}{2}} |F_M^\top R_M^{-1} F_M|^{\frac{1}{2}} \\ &\times \exp\left(-\frac{(\beta_M - \hat{\beta}_M)^\top (F_M^\top R_M^{-1} F_M)(\beta_M - \hat{\beta}_M)}{2\sigma_M^2}\right) \end{aligned} \quad (9)$$

and

$$r(\sigma_M^2 | \phi_M, y) = \frac{(\sigma_M^2)^{-\left(\frac{n-|M|}{2}+1\right)}}{2^{\frac{n-|M|}{2}} \Gamma\left(\frac{n-|M|}{2}\right)} \text{RSS}_M^{\frac{n-|M|}{2}} \exp\left(-\frac{\text{RSS}_M}{2\sigma_M^2}\right), \quad (10)$$

where  $\text{RSS}_M = n\hat{\sigma}_M^2$ ,  $\hat{\beta}_M$  and  $\hat{\sigma}_M^2$  are given in (4) and (5), respectively. Thus, the joint conditional fiducial density of  $(\beta_M^\top, \sigma_M^2)^\top$  for known  $\phi_M$  is

$$\begin{aligned} r(\beta_M, \sigma_M^2 | \phi_M, y) &= r(\sigma_M^2 | \phi_M, y) r(\beta_M | \sigma_M^2, \phi_M, y) \\ &= \frac{\pi^{-\frac{|M|}{2}}}{2^{\frac{n}{2}} \Gamma\left(\frac{n-|M|}{2}\right)} \text{RSS}_M^{\frac{n-|M|}{2}} |F_M^\top R_M^{-1} F_M|^{\frac{1}{2}} (\sigma_M^2)^{-\frac{n}{2}-1} \\ &\times \exp\left\{-\frac{(y - F_M \beta_M)^\top R_M^{-1} (y - F_M \beta_M)}{2\sigma_M^2}\right\}. \end{aligned} \quad (11)$$

When  $\phi_M$  is unknown and the fiducial distribution  $r(\phi_M | y)$  is given, we could obtain the joint fiducial density of  $\theta_M$  as follows:

$$r(\theta_M | y) = r(\phi_M | y) r(\sigma_M^2 | \phi_M, y) r(\beta_M | \sigma_M^2, \phi_M, y).$$

However, it is infeasible to directly obtain the fiducial distribution  $r(\phi_M | y)$  since there is no analytical solution for the partial derivatives of the structural equation with respect to  $\phi_M$ . It

is worth noting that the fiducial distribution and the posterior distribution are distributions of parameter for given observed value.

In this paper, we suggest to use the conditional Bayesian posterior distribution of  $\phi_M$  instead of its conditional fiducial distribution to obtain the Fiducial distribution of  $\theta_M$ , which is called the quasi-fiducial distribution of  $\theta_M$ . Then the following results can be obtained.

**Theorem 2.1:** We assume that the prior of  $\phi_M$  is  $\pi(\phi_M)$ , and it is independent of  $(\beta_M^\top, \sigma_M^2)^\top$ . The full conditional quasi-fiducial distribution of  $\theta$  can be expressed as follows:

$$(\beta_M | \sigma_M^2, \phi_M, y) \sim N(\hat{\beta}_M, \sigma_M^2 F_M^\top R^{-1} F_M), \quad (12)$$

$$(\sigma_M^2 | \phi_M, y) \sim IG\left(\frac{n - |M|}{2}, \frac{\text{RSS}_M}{2}\right), \quad (13)$$

$$r(\phi_M | \beta_M, \sigma_M^2, y) \propto |R_M|^{-\frac{1}{2}} \exp\left\{-\frac{(y - F_M \beta_M)^\top R_M^{-1} (y - F_M \beta_M)}{2\sigma_M^2}\right\} \pi(\phi_M). \quad (14)$$

**Proof:** For given  $\phi_M$ , (12) and (13) could be obtained by (9) and (10), respectively.

For given  $(\beta_M^\top, \sigma_M^2)^\top$ , since  $\pi(\phi_M)$  is independent of  $(\beta_M^\top, \sigma_M^2)^\top$ , the Bayesian posterior distribution of  $\phi_M$  for given  $(\beta_M^\top, \sigma_M^2)^\top$  is

$$\begin{aligned} r(\phi_M | \beta_M, \sigma_M^2, y) &\propto L(\theta_M | y) \pi(\phi_M) \\ &= (2\pi\sigma_M^2)^{-\frac{n}{2}} |R_M|^{-\frac{1}{2}} \exp\left\{-\frac{(y - F_M \beta_M)^\top R_M^{-1} (y - F_M \beta_M)}{2\sigma_M^2}\right\} \pi(\phi_M) \\ &\propto |R_M|^{-\frac{1}{2}} \exp\left\{-\frac{(y - F_M \beta_M)^\top R_M^{-1} (y - F_M \beta_M)}{2\sigma_M^2}\right\} \pi(\phi_M), \end{aligned}$$

where  $L(\theta_M | y)$  is the likelihood function in (3). We use  $r(\phi_M | \beta_M, \sigma_M^2, y)$  instead of its conditional fiducial distribution of  $\phi_M$ . Then the full conditional quasi-fiducial distribution of  $\theta$  could be obtained. ■

Since  $\phi_k \in (0, 1)$ ,  $U(0, 1)$  is chosen as the prior of  $\phi_k$  to ensure computational efficiency (Huang et al., 2020). Therefore, we could adopt this prior in our study. Once the conditional quasi-fiducial distribution of  $\phi_M$  is determined, fiducial samples can be generated in a manner similar to the Bayesian approach. With an initial value, the Gibbs sampler (Shao, 2010) can be used to generate the quasi-fiducial distribution of  $\theta_M$  by repeated successive sampling from (12) through (14), which is detailed in Algorithm 2, and the obtained samples are called quasi-fiducial samples. The initial value  $\phi_0$  is given in the next algorithm which is the maximum likelihood estimation algorithm for the parameters. For Algorithm 2, sampling from  $r(\beta_M | \sigma_M^2, \phi_M, y)$  and  $r(\sigma_M^2 | \phi_M, y)$  is relatively straightforward, nevertheless sampling from  $r(\phi_M | \beta_M, \sigma_M^2, y)$  remains challenging because it does not have a precise analytical form and is only known to be proportional to a specific known function. So we can address the issue by employing the slice sampling method from Neal (2003). The slice sampling method for the conditional fiducial distribution of  $\phi$  is shown in Algorithm 3.

**Algorithm 2** (A Gibbs sampler of the quasi-fiducial samples of  $\theta_M$ )**Input:** Data  $(x_i, y_i)_{i=1}^n$ , initial values  $\phi_0$ , number of samples  $T$ , threshold parameter  $n_1, t = 0$ :**Step 1:** Draw the fiducial sample  $\sigma_{M;t}^2$  of  $\sigma_M^2$  according to Equation (12).**Step 2:** Draw the fiducial sample  $\beta_{M;t}$  of  $\beta_M$  according to Equation (13). **Step 3:** Draw the fiducial sample  $\phi_{M;t}$  of  $\phi_M$  according to Equation (14) and obtain the new sample  $(\beta_{M;t+1}, \sigma_{M;t+1}^2, \phi_{M;t+1})$ . Let  $t = t + 1$ . **Step 4:** Repeat Steps 1–3 until  $t = n_1 + T$ . **Output:** The quasi-fiducial samples  $(\beta_{M;n_1+1}, \sigma_{M;n_1+1}^2, \phi_{M;n_1+1}), \dots, (\beta_{M;n_1+T}, \sigma_{M;n_1+T}^2, \phi_{M;n_1+T})$ .

Let  $r(\phi_M | y)$  and  $(\phi_{M;1}, \dots, \phi_{M;T})$  be the quasi-fiducial distribution and a sample of  $\phi_M$ , respectively. The joint quasi-fiducial density of  $(\beta_M^\top, \sigma_M^2)$  is

$$r(\beta_M, \sigma_M^2 | y) = \int r(\beta_M, \sigma_M^2 | \phi_M, y) r(\phi_M | y) d\phi_M, \quad (15)$$

which could be computed by  $\frac{1}{T} \sum_{t=1}^T r(\beta_M, \sigma_M^2 | \phi_{M;t}, y)$ .

**Algorithm 3** (The slice sampling method for samples of  $\phi_M$ )**Input:** The function  $r(\phi | \beta, \sigma^2, y)$ , the initial values  $\phi_0 = (\phi_{0,1}, \dots, \phi_{0,d}), H = (L_1, R_1) \times \dots \times (L_d, R_d)$ . **Step 1:** Draw a random value  $z$  from  $U(0, r(\phi_0 | \beta, \sigma^2, y))$  and define a set  $S = \{\phi \in \mathbb{R}^d : z < r(\phi | \beta, \sigma^2, y)\}$ . **Step 2:** Draw sample  $\phi_*$  from the set  $S$  and insert the  $\phi_*$  into the following While loop :**While**  $r(\phi_0 | \beta, \sigma^2, y) < z$  :    **For**  $i = 1, \dots, d$  :        if  $\phi_{*,i} < \phi_{0,i}$ , then  $L_i \leftarrow \phi_{*,i}$  ;        or  $R_i \leftarrow \phi_{*,i}$ .    **End;**    Resample  $\phi_*$  from the updated  $H$ ;**End.**

After completing this loop iteration, the updated  $\phi_*$  is our new sample  $\phi_1$ . **Output:** The new sample  $\phi_1 = (\phi_{1,1}, \dots, \phi_{1,d})$ .

**3. The quasi-fiducial model selection**

A fiducial model selection method for high-dimensional regression was studied by Zhao et al. (2023) and proposed the following definition of fiducial marginal likelihood function.

**Definition 3.1:** Let  $L(\theta_M | y)$  be the likelihood function of the observed sample under the  $M$ th model and  $r(\theta_M | y)$  be the fiducial distribution of  $\theta_M$ . Define

$$P_r(y | M) \triangleq \int_{\Theta_M} L(\theta_M | y) \cdot r(\theta_M | y) d\theta_M.$$

We call  $P_r(y | M)$  the fiducial marginal likelihood function (FML).

Assuming that the priors  $\pi(M)$  of all models are uniformly distributed, then

$$P_r(M|y) = \frac{P_r(y|M)\pi(M)}{\sum_{M' \in \mathcal{M}} P_r(y|M')\pi(M')} \propto P_r(y|M).$$

$P_r(M|y)$  can be used for model selection and inference because  $P_r(M|y)$  reflects which model the observed sample more likely comes from. In the following, we give the FML for the UK model (2).

**Theorem 3.1:** Assuming the quasi-fiducial distribution of  $\phi_M$  is  $r(\phi_M|y)$ , the FML of model (2) can be expressed as follows:

$$P_r(y|M) \propto E_{\phi_M} \left\{ E_{\beta_M|\phi_M} \left\{ \left[ (y - F_M\beta_M)^\top R_M^{-1} (y - F_M\beta_M) \right]^{-\frac{n}{2}} \right\} |R_M|^{-\frac{1}{2}} \right\}, \quad (16)$$

where the expectation  $E_{\phi_M}$  is taken with respect to the quasi-fiducial distribution of  $\phi_M$ , and the expectation  $E_{\beta_M|\phi_M}$  is taken with respect to the fiducial distribution of  $\beta_M$  condition on  $\phi_M$ .

**Proof:** Note that  $(y - F_M\beta_M)^\top R_M^{-1} (y - F_M\beta_M) = \text{RSS}_M + (\beta_M - \hat{\beta}_M)^\top F_M^\top R_M^{-1} F_M (\beta_M - \hat{\beta}_M)$ . By Equations (10) and (11), the fiducial distribution of  $\beta_M$  condition on  $\phi_M$  is

$$\begin{aligned} r(\beta_M|\phi_M, y) &= \int_0^\infty r(\beta_M, \sigma_M^2|\phi_M, y) d\sigma_M^2 \\ &= \int_0^\infty \frac{\pi^{-\frac{|M|}{2}}}{2^{\frac{n}{2}} \Gamma\left(\frac{n-|M|}{2}\right)} \text{RSS}_M^{\frac{n-|M|}{2}} |F_M^\top R_M^{-1} F_M|^{\frac{1}{2}} (\sigma_M^2)^{-(\frac{n}{2}+1)} \\ &\quad \times \exp \left\{ -\frac{(y - F_M\beta_M)^\top R_M^{-1} (y - F_M\beta_M)}{2\sigma_M^2} \right\} d\sigma_M^2 \\ &= \frac{\Gamma\left(\frac{n}{2}\right)}{\pi^{\frac{|M|}{2}} \Gamma\left(\frac{n-|M|}{2}\right)} \text{RSS}_M^{\frac{n-|M|}{2}} |F_M^\top R_M^{-1} F_M|^{\frac{1}{2}} \left[ (y - F_M\beta_M)^\top R_M^{-1} (y - F_M\beta_M) \right]^{-\frac{n}{2}} \\ &= \frac{\Gamma\left(\frac{n}{2}\right) \left| \frac{\text{RSS}_M}{n-|M|} (F_M^\top R_M^{-1} F_M)^{-1} \right|^{-\frac{1}{2}}}{(n-|M|)^{\frac{|M|}{2}} \pi^{\frac{|M|}{2}} \Gamma\left(\frac{n-|M|}{2}\right)} \\ &\quad \times \left[ 1 + \frac{1}{n-|M|} (\beta_M - \hat{\beta}_M)^\top \frac{n-|M|}{\text{RSS}_M} F_M^\top R_M^{-1} F_M (\beta_M - \hat{\beta}_M) \right]^{-\frac{n}{2}}, \end{aligned}$$

so  $\beta_M$  conditional on  $\phi_M$  follows multivariate  $t$ -distribution

$$t_{n-|M|} \left( \hat{\beta}_M, \frac{\text{RSS}_M}{n-|M|} (F_M^\top R_M^{-1} F_M)^{-1} \right).$$

By Equations (3) and (11),

$$\begin{aligned} L(\theta_M | y) r(\theta_M | y) &= L(\theta_M | y) r(\beta_M, \sigma_M^2 | \phi_M, y) r(\phi_M | y) \\ &= \frac{\pi^{-\frac{n+|M|}{2}} (\sigma_M^2)^{-(n+1)}}{2^n \Gamma\left(\frac{n-|M|}{2}\right) |R_M|^{\frac{1}{2}}} \text{RSS}_M^{\frac{n-|M|}{2}} |F_M^\top R_M^{-1} F_M|^{\frac{1}{2}} \\ &\quad \times \exp \left\{ -\frac{(y - F_M \beta_M)^\top R_M^{-1} (y - F_M \beta_M)}{\sigma_M^2} \right\} r(\phi_M | y), \end{aligned}$$

and then the FML of model  $M$  is

$$\begin{aligned} P_r(y | M) &= \iiint L(\theta_M | y) r(\theta_M | y) d\sigma^2 d\beta_M d\phi_M \\ &= \iint \frac{\pi^{-\frac{n+|M|}{2}} \Gamma(n)}{2^n \Gamma\left(\frac{n-|M|}{2}\right) |R_M|^{\frac{1}{2}}} \text{RSS}_M^{\frac{n-|M|}{2}} |F_M^\top R_M^{-1} F_M|^{\frac{1}{2}} \\ &\quad \times \left[ (y - F_M \beta_M)^\top R_M^{-1} (y - F_M \beta_M) \right]^{-n} r(\phi_M | y) d\beta_M d\phi_M \\ &= \iint \frac{\Gamma(n)}{2^n \pi^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right) |R_M|^{\frac{1}{2}}} \left[ (y - F_M \beta_M)^\top R_M^{-1} (y - F_M \beta_M) \right]^{-\frac{n}{2}} \\ &\quad \times r(\beta_M | \phi_M, y) r(\phi_M | y) d\beta_M d\phi_M \\ &\propto \int E_{\beta_M | \phi_M} \left\{ \left[ (y - F_M \beta_M)^\top R_M^{-1} (y - F_M \beta_M) \right]^{-\frac{n}{2}} \right\} |R_M|^{-\frac{1}{2}} r(\phi_M | y) d\phi_M \\ &= E_{\phi_M} \left\{ E_{\beta_M | \phi_M} \left\{ \left[ (y - F_M \beta_M)^\top R_M^{-1} (y - F_M \beta_M) \right]^{-\frac{n}{2}} \right\} |R_M|^{-\frac{1}{2}} \right\}, \end{aligned}$$

where the expectation  $E_{\phi_M}$  is taken with respect to the quasi-fiducial distribution of  $\phi_M$ , and the expectation  $E_{\beta_M | \phi_M}$  is taken with respect to the conditional fiducial distribution of  $\beta_M$ . ■

Based on this model uncertainty measure (16), a new fiducial model selection criterion is proposed for model (2). The optimal model selected from  $\mathcal{M}$  should satisfy the largest fiducial marginal likelihood, i.e.,

$$M^* = \underset{M \in \mathcal{M}}{\operatorname{argmax}} P_r(y | M).$$

Because it is hard to obtain a closed form of the FML  $P_r(y | M)$ , which could be computed using quasi-fiducial samples of  $\theta$  based on Algorithm 2. Let  $(\phi_{M;1}^\top, \dots, \phi_{M;N_1}^\top)^\top$  be a quasi-fiducial sample of  $\phi_M$  in Algorithm 2, and  $(\beta_{M;1}^\top, \dots, \beta_{M;N_2}^\top)^\top$  be a conditional fiducial sample from  $t_{n-|M|}(\hat{\beta}_M, \frac{\text{RSS}_M}{n-|M|} (F_M^\top R_{M;i}^{-1} F_M)^{-1})$  for given  $R_{M;i} = R(\phi_{M;i})$  ( $i = 1, \dots, T$ ), respectively. Then we have

$$\hat{P}_r(y | M) = \frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \left\{ \left[ (y - F_M \beta_{M;j})^\top R_{M;i}^{-1} (y - F_M \beta_{M;j}) \right]^{-\frac{n}{2}} \right\} |R_{M;i}|^{-\frac{1}{2}}. \quad (17)$$

We call the aforementioned model selection method as the quasi-fiducial Model Selection (QFMS for short) and list its detailed steps in Algorithm 4.

---

**Algorithm 4** (The QFMS on the UK models)
 

---

**Step 1:** According to the Sure Independence Screening (SIS) criterion, we rank the variables by importance and construct a set of nested candidate models  $\mathcal{M} = \{M_1, M_2, \dots, M_K\}$ . Let  $k = 1$ . **Step 2:** For model  $M_k \in \mathcal{M}$ , compute  $(\hat{\beta}_{M_k}^\top, \hat{\sigma}_{M_k}^2, \hat{\phi}_{M_k}^\top)^\top$  and generate a quasi-fiducial sample  $(\phi_{M_k;1}^\top, \dots, \phi_{M_k;N_1}^\top)^\top$  via Algorithm 2. **Step 3:** For each  $\phi_{M_k;i}$ , generate a fiducial sample  $(\beta_{M_k;1}^\top, \dots, \beta_{M_k;N_2}^\top)^\top$  from  $t_{n-|M|} \left( \hat{\beta}_{M_k}, \frac{\text{RSS}_{M_k}}{n-|M|} (F_{M_k}^\top R_{M_k;i}^{-1} F_{M_k})^{-1} \right)$ . **Step 4:** Substitute the quasi-fiducial samples into (17) and calculate the  $\hat{P}_r(y|M_k)$ .  $k = k + 1$ . **Step 5:** Repeat Step 2 to Step 4 and calculate  $\hat{P}_r(y|M_{k+1})$  until  $k = K$ . **Step 6:** Find out the optimal model  $M^*$  such that  $M^* = \text{argmax}_{M \in \mathcal{M}} \hat{P}_r(y|M)$ .

---

#### 4. Numerical simulations and case analysis

In this section, we primarily evaluate the numerical performance of the proposed method through two numerical simulations and a case analysis, and compare it with the well-studied methods such as Lasso and EN with respect to estimation accuracy.

The quasi-fiducial distribution of  $\phi_M$  depends on its prior  $\pi(\phi_M)$ ,  $\phi_M \in (0, 1)^d$ . In this study, we choose its prior is  $\phi_i \sim \beta(0.5, 0.5)$  and  $\phi_i \sim \beta(1, 1)$  ( $i = 1, \dots, d$ ), and two corresponding methods are denoted QFMS<sub>0.5</sub> and QFMS<sub>1</sub>, respectively. To study the selection performance of QFMS, we use the root mean square error of prediction (RMSEP) to measure prediction accuracy, and it is defined as

$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^N \left( y(x_{\text{test}}^{(i)}) - \hat{y}(x_{\text{test}}^{(i)}) \right)^2}{N}},$$

where  $N$  is the number of testing samples. The mean of the root mean square prediction error (MRMSEP) and the standard deviance of the root mean square prediction error (sd(RMSEP)) are computed respectively. Furthermore, we need identifications to measure the accuracy of variable selection for linear regression models, which use the average number of inactive effect identified rate (IEIR), the average number of active effect identified rate (AEIR) and the average size of the identified model (MEAN). The accuracy of variable identification is measured by

$$\text{AEIR} = \frac{1}{N} \sum_{i=1}^N \frac{|M_i \cap M_o|}{|M_o|}, \quad \text{IEIR} = \frac{1}{N} \sum_{i=1}^N \frac{|M_i \cap M_o^c|}{|M_o^c|}, \quad \text{MEAN} = \frac{1}{N} \sum_{i=1}^N |M_i|,$$

where  $M_o$  and  $M_o^c$  are active effect and inactive effect of the true model respectively, and  $M_i$  is the set of variables of the optimal model in the  $i$ th experiment.

In simulation study,  $N = 1000$  samples are randomly selected as testing data, and we implement the proposed method with 1500 MCMC steps for each data set, and discard the first  $N_{\text{burnin}} = 500$ .

##### 4.1. A linear regression function

The known function is defined on a 12-dimensional ( $d = 12$ ) input space  $[0, 1]^{12}$ , where the first 6 variables  $x_1, \dots, x_6$  have decreasing effects on the computer experiment output results,

**Table 1.** The results of the model selection of linear models.

| $n$ | Method              | AEIR (%) | IEIR (%) | MEAN | MRMSEP | sd(RMSEP) |
|-----|---------------------|----------|----------|------|--------|-----------|
| 50  | QFMS <sub>1</sub>   | 99.65    | 23.26    | 7.38 | 0.1243 | 0.0254    |
|     | QFMS <sub>0.5</sub> | 99.30    | 22.57    | 7.31 | 0.1241 | 0.0252    |
|     | Lasso               | 97.40    | 37.50    | 8.09 | 0.1396 | 0.0276    |
|     | EN                  | 98.44    | 37.33    | 8.15 | 0.1366 | 0.0246    |
| 80  | QFMS <sub>1</sub>   | 99.48    | 38.02    | 8.25 | 0.1196 | 0.0229    |
|     | QFMS <sub>0.5</sub> | 99.65    | 40.28    | 8.40 | 0.1202 | 0.0224    |
|     | Lasso               | 98.09    | 51.04    | 8.95 | 0.1408 | 0.0357    |
|     | EN                  | 98.96    | 51.39    | 9.02 | 0.1355 | 0.0233    |
| 100 | QFMS <sub>1</sub>   | 98.61    | 46.18    | 8.69 | 0.1221 | 0.0241    |
|     | QFMS <sub>0.5</sub> | 98.96    | 48.96    | 8.88 | 0.1227 | 0.0239    |
|     | Lasso               | 99.13    | 55.90    | 9.30 | 0.1385 | 0.0238    |
|     | EN                  | 99.65    | 56.25    | 9.35 | 0.1396 | 0.0243    |

and the coefficients of the remaining variables  $x_7, \dots, x_{12}$  are zero. The mean function is

$$y(x) = 0.4x_1 + 0.3x_2 + 0.2x_3 + 0.1x_4 + 0.05x_5 + 0.01x_6 + z(x), \quad (18)$$

where  $z(x) \sim N(0, \sigma^2 R)$ ,  $\sigma^2 = 0.01$  and  $R$  is the covariance matrix based on the power exponential correlation coefficients from (2) with  $\phi_i = 0.5$ ,  $i = 1, \dots, 12$ , respectively.

This simulation generates samples with sample sizes  $n = 50, 80, 100$  and dimensions  $d = 12$  using Latin hypercube sampling from Chen et al. (2016). We rank the importance of the 12 input variables by the criterion SIS to construct the 12 nested UK models as a candidate model set. As shown in Table 1, both QFMS<sub>1</sub> and QFMS<sub>0.5</sub> demonstrate superior performance to LASSO and EN, with lower values for IEIR, MEAN and MRMSEP, while exhibiting higher AEIR values.

#### 4.2. A non-linear model

Borehole function of Morris et al. (1993)

$$y(x) = 2\pi x_3(x_4 - x_6) \left\{ \log\left(\frac{x_2}{x_1}\right) \left(1 + 2\frac{x_3 x_7}{\log(x_2/x_1 x_1^2 x_8)} + \frac{x_3}{x_5}\right) \right\}^{-1}. \quad (19)$$

The input space is a rectangular interval  $[0.05, 0.015] \times [100, 5000] \times [63070, 115600] \times [990, 1110] \times [63.1, 116] \times [700, 820] \times [1120, 1680] \times [9855, 12045]$ , and the samples were generated through Latin hypercube sampling with dimension  $p = 8$  and sample size  $N = 100$ . We use the UK model to fit the borehole function and set its mean function as

$$m(x) = \beta_0 + \beta_1 x_1 + \dots + \beta_8 x_8.$$

The performances of Lasso, EN, QFMS<sub>1</sub> and QFMS<sub>0.5</sub> are shown in Table 2, where MEAN, MRMSEP and sd(RMSEP) are calculated by the selected optimal model.

Results in Table 2 show that MRMSEP and sd(RMSEP) of QFMS<sub>1</sub> and QFMS<sub>0.5</sub> are lower than that of Lasso and EN, which means that the prediction accuracy and stability of QFMS are better. QFMS<sub>1</sub> and QFMS<sub>0.5</sub> perform similarly.

#### 4.3. Case analysis

We validate all methods by an engineering problem called piston slap noise. Piston slap noise, caused by the secondary motion of pistons within cylinders, affects engine performance and

**Table 2.** The results of the model selection of non-linear models.

| $n$ | Method              | MEAN | MRMSEP | sd( RMSEP) |
|-----|---------------------|------|--------|------------|
| 50  | QFMS <sub>1</sub>   | 7.00 | 5.0196 | 1.6701     |
|     | QFMS <sub>0.5</sub> | 7.00 | 4.9926 | 1.7421     |
|     | Lasso               | 5.00 | 6.5298 | 3.7572     |
|     | EN                  | 8.00 | 8.2972 | 2.4832     |
| 80  | QFMS <sub>1</sub>   | 6.52 | 2.6178 | 0.9748     |
|     | QFMS <sub>0.5</sub> | 6.49 | 2.6399 | 1.0660     |
|     | Lasso               | 5.32 | 3.5016 | 2.8685     |
|     | EN                  | 7.82 | 5.7124 | 2.6871     |
| 100 | QFMS <sub>1</sub>   | 6.49 | 1.9897 | 1.0526     |
|     | QFMS <sub>0.5</sub> | 6.46 | 1.9724 | 1.0502     |
|     | Lasso               | 5.21 | 1.9001 | 1.6537     |
|     | EN                  | 7.76 | 4.7061 | 2.3499     |

contributes to environmental noise pollution. To minimize the noise, our study analyses six significant factors: piston-to-cylinder clearance ( $x_1$ ), peak pressure position ( $x_2$ ), skirt length ( $x_3$ ), skirt profile shape ( $x_4$ ), skirt ovality ( $x_5$ ) and piston pin offset ( $x_6$ ). Data used in this study were sourced from Huang et al. (2020), comprising 100 observations with 6 input variables per sample. Model construction considered all possible linear main effects.

To validate the model's performances, the 100 samples were divided into 2 parts: 80 samples were used as the training set for model training and parameter estimations, and the remaining 20 samples were used as the test set for evaluating the predictive performance. The RMSEPs of QFMS<sub>0.5</sub>, QFMS<sub>1</sub>, Lasso and EN are 0.2711, 0.6224, 1.1817 and 1.2047, respectively, which indicate that prediction accuracy of optimal model selected by QFMS<sub>1</sub> is better than the two penalized methods and QFMS<sub>0.5</sub>.

## 5. Conclusion

Fiducial inference is applied to obtain fiducial distributions for different parameters in the Kriging model. To solve the challenge of constructing the fiducial marginal distribution for space related parameter, we substitute the Bayesian posterior distribution of spatially related parameter for the fiducial marginal distribution and present a quasi-fiducial distribution for Kriging model parameters. Additionally, a model uncertainty measure based on the fiducial marginal likelihood function and its approximation method is proposed, leading to the development of a new model selection method, QFMS. Numerical studies demonstrated that the QFMS method outperforms Lasso and Elastic-Net in terms of prediction precision and stability. And in many situations, the identification accuracy of QFMS performs best when we concentrate on linear models.

There are several model selection issues for further investigation. First, we do not provide the joint fiducial distribution of Kriging model parameters, and statistical inference is conducted through the marginal distribution of the parameters instead. Exploring the joint fiducial distribution of the parameters remains a future research direction. Second, the explicit expression for the marginal distribution of the spatial parameter  $\phi$  is not provided. If this issue can be resolved, we could derive a more precise double expectation expression for the uncertainty probability measure and further obtain a more accurate Monte Carlo estimation for it. At last, the relationship between model complexity and prediction performance has always been an important aspect of model selection methods, which deserves further in-depth exploration.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

]This research is supported by the National Social Science Found of China [Grant number 23BTJ064].

## References

- Chen, H., Huang, H., Lin, D. K., & Liu, M. Q. (2016). Uniform sliced Latin hypercube designs. *Applied Stochastic Models in Business and Industry*, 32(5), 574–584. <https://doi.org/10.1002/asmb.2192>
- Cui, Y., Hannig, J., & Edlefsen, P. (2024). Semiparametric fiducial inference. <https://arxiv.org/abs/2404.18779>
- Hannig, J. (2009). On generalized fiducial inference. *Statistica Sinica*, 19(2), 491–544. <https://doi.org/10.5705/ss.2011.102>
- Hannig, J. (2013). Generalized fiducial inference via discretization. *Statistica Sinica*, 23(2), 489–514.
- Hodges, J. S., & Reich, B. J. (2010). Adding spatially-correlated errors can mess up the fixed effect you love. *The American Statistician*, 64(4), 325–334. <https://doi.org/10.1198/tast.2010.10052>
- Huang, H., Lin, D. K., Liu, M. Q., & Zhang, Q. (2020). Variable selection for Kriging in computer experiments. *Journal of Quality Technology*, 52(1), 40–53. <https://doi.org/10.1080/00224065.2019.1569959>
- Hung, Y. (2011). Penalized blind Kriging in computer experiments. *Statistica Sinica*, 21(3), 1171–1190. <https://doi.org/10.5705/ss.2011.v21n3a>
- Joseph, V. R., Hung, Y., & Sudjianto, A. (2008). Blind Kriging: A new method for developing metamod-els. *Journal of Mechanical Design*, 130(3), 350–353. <https://doi.org/10.1115/1.2829873>
- Morris, M. D., Mitchell, T. J., & Ylvisaker, D. (1993). Bayesian design and analysis of com-puter experiments: Use of derivatives in surface prediction. *Technometrics*, 35(3), 243–255. <https://doi.org/10.1080/00401706.1993.10485320>
- Neal, R. M. (2003). Slice sampling. *Annals of Statistics*, 31(3), 705–767. <https://doi.org/10.1214/aos/1056562461>
- Paciorek, C. J. (2010). The importance of scale for spatial-confounding bias and precision of spatial regression estimators. *Statistical Science*, 25(1), 107–125. <https://doi.org/10.1214/10-STS326>
- Park, I. (2021). Lasso Kriging for efficiently selecting a global trend model. *Structural and Multidisci-iplinary Optimization*, 64(3), 1527–1543. <https://doi.org/10.1007/s00158-021-02939-7>
- Sacks, J., Schiller, S. B., & Welch, W. J. (1989). Designs for computer experiments. *Technometrics*, 31(1), 41–47. <https://doi.org/10.1080/00401706.1989.10488474>
- Shao, J. (2010). *Mathematical Statistics* (2nd ed.). Springer.
- Simpson, T. W., Mauery, T. M., Korte, J. J., & Mistree, F. (2001). Kriging models for global approxima-tion in simulation-based multidisciplinary design optimization. *AIAA Journal*, 39(12), 2233–2241. <https://doi.org/10.2514/2.1234>
- Tuo, R., & Wu, C. J. (2015). Efficient calibration for imperfect computer models. *Annals of Statistics*, 43(6), 2331–2352. <https://doi.org/10.1214/15-AOS1314>
- Welch, W. J., Buck, R. J., Sacks, J., Wynn, H. P., Mitchell, T. J., & Morris, M. D. (1992). Screening, pre-dicting, and computer experiments. *Technometrics*, 34(1), 15–25. <https://doi.org/10.2307/1269548>
- Williams, J. P., & Hannig, J. (2019). Nonpenalized variable selection in high-dimensional linear model settings via generalized fiducial inference. *Annals of Statistics*, 47(3), 1723–1753. <https://doi.org/10.1214/18-AOS1733>
- Zhao, Y., Feng, Z., Li, M., & Li, X. (2023). Modified penalized blind Kriging for efficiently select-ing a global trend model. *Journal of Statistical Computation and Simulation*, 93(17), 3052–3066. <https://doi.org/10.1080/00949655.2023.2216335>
- Zhao, Y., Liang, H., & Li, X. (2023). Fiducial variable selection for the high-dimensional regression model. *Scientia Sinica Mathematica*, 53(6), 839–858. <https://doi.org/10.1360/SCM-2022-0245>