Geomechanics and Engineering, *Vol. 30, No. 2 (2022) 187-199*
https://doi.org/10.12989/gae.2022.30.2.187

Adaptively selected autocorrelation structure-based Kriging metamodel for slope reliability analysis

Jing-Ze Li[1a], Shao-He Zhang[1b], Lei-Lei Liu*[1], Jing-Jing Wu[2c] and Yung-Ming Cheng[3d]

[1]*Key Laboratory of Metallogenic Prediction of Nonferrous Metals and Geological Environment Monitoring, Ministry of Education, School of Geosciences and Info-Physics, Central South University, Changsha 410083, P.R. China*
[2]*College of Civil Engineering, Hunan University of Technology, Zhuzhou 412007, P.R. China*
[3]*School of Civil Engineering, Qingdao University of Technology, Qingdao 266033, P.R. China*

(Received November 20, 2021, Revised June 10, 2022, Accepted June 11, 2022)

Abstract. Kriging metamodel, as a flexible machine learning method for approximating deterministic analysis models of an engineering system, has been widely used for efficiently estimating slope reliability in recent years. However, the autocorrelation function (ACF), a key input to Kriging that affects the accuracy of reliability estimation, is usually selected based on empiricism. This paper proposes an adaption of the Kriging method, named as Genetic Algorithm optimized Whittle-Matérn Kriging (GAWMK), for addressing this issue. The non-classical two-parameter Whittle-Matérn (WM) function, which can represent different ACFs in the Matérn family by controlling a smoothness parameter, is adopted in GAWMK to avoid subjectively selecting ACFs. The genetic algorithm is used to optimize the WM model to adaptively select the optimal autocorrelation structure of the GAWMK model. Monte Carlo simulation is then performed based on GAWMK for a subsequent slope reliability analysis. Applications to one explicit analytical example and two slope examples are presented to illustrate and validate the proposed method. It is found that reliability results estimated by the Kriging models using randomly chosen ACFs might be biased. The proposed method performs reasonably well in slope reliability estimation.

Keywords: genetic algorithm; Kriging; reliability analysis; slope stability; Whittle-Matérn

1. Introduction

Probabilistic methods have received considerable attention to deal with uncertainties in geotechnical stability analysis (e.g., slope stability analysis) in recent years (Phoon 2017, Liu *et al.* 2020, Bai *et al.* 2020, Qi and Liu 2019a, Li *et al.* 2014). For example, Monte Carlo Simulation (MCS) is one of the most popular probabilistic methods for estimating slope reliability because of its robustness and conceptual simplicity (Au *et al.* 2010, Wang *et al.* 2010). However, an obvious deficiency of direct MCS is that it generally suffers from a lack of efficiency, especially when the deterministic numerical model is computationally demanding. To alleviate the computational burden, several methods have been proposed to boost the efficiency of the reliability analysis (Jiang *et al.* 2015, Wang *et al.* 2010, Liu *et al.* 2020). Among the optimization methods, metamodels, which often approximate the time-consuming numerical models with closed-form expressions or "black box", have been proposed and used as a suitable alternative for reliability analysis (El Haj and Soubra 2020).

Among the available metamodels in the literature (Zhang *et al.* 2013, Zhang *et al.* 2011, Gaspar *et al.* 2014, Li *et al.* 2016), Kriging has demonstrated its superiority over others. For example, Zhang *et al.* (2013) showed that, compared with the classical response surface method, Kriging produces a more accurate result for slope reliability analysis at the expense of more calibration points. Gaspar *et al.* (2014) presented that Kriging can more accurately evaluate the probability of failure (P_f) than the polynomial regression models do. In addition, Kriging provides not only the estimates at unsampled locations but also the uncertainties associated with them. This additional piece of information allows Kriging to reasonably quantify the prediction error, which further advances its applications in slope reliability analysis. In this regard, Liu and Cheng (2018) proposed an adaptive Kriging model with selecting training samples at the most uncertain points, which overcomes the difficulty of randomly choosing training samples in other models. This manipulation effectively reduces the size of training data or equally enhances the computation efficiency. Furthermore, several global optimization algorithms are proposed to improve the performance and feasibility of Kriging models for an accurate slope reliability analysis, e.g., artificial bee colony algorithm (Luo *et al.* 2012), genetic algorithm (Liu *et al.* 2017), and particle swarm optimization algorithm (Yi *et al.*

*Corresponding author, Associate Professor
E-mail: csulll@foxmail.com
[a]Ph.D. Student
E-mail: jingze_li@csu.edu.cn
[b]Professor
E-mail: zhangshaohe@163.com
[c]Lecturer
E-mail: wujingjing0208@163.com
[d]Professor
E-mail: ceymchen@polyu.edu.hk

ISSN: 2005-307X (Print), 2092-6219 (Online)

2015). Overall, the above literature analyses show that Kriging has been gaining increasing popularity in slope reliability analysis.

However, as a key input to the Kriging model, the autocorrelation function (ACF) is often chosen as the squared exponential function (also known as Gaussian function) by default or selected from a group of candidate ACFs (e.g., the single exponential ACF, squared exponential ACF and second-order Markov ACF) based on empirical knowledge (Lophaven *et al.* 2002, Liu *et al.* 2017). This shows that there might be some uncertainties in the selection of ACF for the Kriging model construction. Such uncertainties can subsequently lead to biased reliability results because autocorrelation structures play a key role in the representation of the spatial correlation underlying the Kriging model (Minasny and McBratney 2005). It has also been suggested that the major problem in the Kriging method is both the type of ACF and the associated parameters are unknown (Kleijnen 2009). Meanwhile, ACF has a significant effect on the estimation of P_f in slope reliability analysis considering spatial variability (Ching and Phoon 2019, Zhang *et al.* 2020, Qi and Li 2018, Qi and Liu 2019b). Therefore, it is of great significance to improve the performance of Kriging metamodel to obtain more reliable reliability results by reducing the uncertainty in ACF selection in the Kriging model. To the best of our knowledge, this work has not been well dealt with or fully investigated.

According to the literature, a non-classical ACF, Whittle-Matérn (WM), can replace the commonly used single-parameter ACFs (e.g., the single exponential ACF, the squared exponential ACF, and second-order Markov ACF) within Matérn family by controlling the smoothness parameter, ν (Minasny and McBratney 2005, Marchant and Lark 2007). For example, when ν takes values of 0.5 and 1.5, the corresponding WM ACF evolves as the single exponential ACF and second-order Markov ACF, respectively. Therefore, the best ACF for the Kriging metamodel can be selected alternatively by optimizing the smoothness parameter in the WM function based on the given data.

This paper develops an adaptive Kriging metamodel named GAWMK for slope reliability analysis, which makes use of the genetic algorithm (GA) to simultaneously optimize the correlation parameter and smoothness parameter of WM function to find the optimal autocorrelation structure in the Kriging model (Minasny and McBratney 2005, Ching *et al.* 2019). The proposed GAWMK overcomes the deficiency of the traditional Kriging method in randomly choosing ACF and thus improves, to a certain extent, the accuracy of the Kriging method. The paper starts with the formulation of the proposed GAWMK method, which is followed by the slope reliability analysis using GAWMK-based MCS. Then, a Matlab toolbox is developed based on the commonly used Kriging Matlab toolbox, Design and Analysis of Computer Experiments (DACE) (Lophaven *et al.* 2002), to implement the proposed method. Thereafter, applications to three illustrative examples, including one explicit analytical example and two slope examples, are presented to illustrate and validate the proposed method.

2. The proposed GAWMK method

The novelty of the proposed method lies in that the non-classical WM function is incorporated into the Kriging model, with GA optimizing the correlation parameter and smoothness parameter, to adaptively select the optimal ACF for Kriging-based metamodeling. Hence, there are three key modules in the proposed method: (1) traditional Kriging method; (2) WM autocorrelation structure; and (3) GA for the optimization of the correlation parameter and smoothness parameter. A detailed introduction of the three modules is provided in this section as follows.

2.1 Kriging method

Kriging method is a spatial optimal linear prediction using observed data taken at known nearby locations (Cressie 1990). This method is built with the model of stochastic spatial variation which fits well with the reality, such as the borehole or sensor data (Oliver and Webster 1990). In Kriging framework, the response $y(x)$ (or factor of safety (FS) for slope stability analysis) is divided into two parts: the trend part and the stochastic error part (Zhang *et al.* 2011). Suppose x denotes a vector of the input variables, the Kriging model can be expressed as

$$y(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})^T\boldsymbol{\beta} + z(\boldsymbol{x}) \tag{1}$$

where $\boldsymbol{f}(\boldsymbol{x})^T\boldsymbol{\beta}$ is the deterministic trend part, with $\boldsymbol{f}(\boldsymbol{x}) = [f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \dots, f_p(\boldsymbol{x})]^T$ being a vector of p basis functions and $\boldsymbol{\beta} = [\beta_1, \beta_2, \dots, \beta_p]^T$ the vector of the associated p regression parameters; and $z(\boldsymbol{x})$ is the stochastic error part, which is assumed to be a Gaussian process with

$$\begin{cases} E\big(z(\boldsymbol{x})\big) = 0 \\ Var\big(z(\boldsymbol{x})\big) = \sigma^2 \\ Cov\Big(z(\boldsymbol{x}_i), z\big(\boldsymbol{x}_j\big)\Big) = \sigma^2\rho(\boldsymbol{x}_i, \boldsymbol{x}_j) \end{cases} \tag{2}$$

where $E\big(z(\boldsymbol{x})\big)$ denotes the expectation of the process, with a value of zero; $Var\big(z(\boldsymbol{x})\big)$ denotes the variance of the process, with a value of σ^2; and $Cov\Big(z(\boldsymbol{x}_i), z\big(\boldsymbol{x}_j\big)\Big)$ is the covariance between the random variables at points $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, with $\rho(\boldsymbol{x}_i, \boldsymbol{x}_j)$ being the correlation between the two variables. $\rho(\boldsymbol{x}_i, \boldsymbol{x}_j)$ is usually calculated from the autocorrelation structure of the underlying Kriging model, as will be introduced in the following subsection.

Given m sets of input data $\mathbf{S} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_m]^T$ and responses $\boldsymbol{y} = [y_1, y_2, \dots, y_m]^T$, the unknown parameters $\boldsymbol{\beta}$ and σ^2 can be estimated by using the maximum likelihood method (MLE) as

$$\widehat{\boldsymbol{\beta}} = (\mathbf{F}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{F})^{-1}\mathbf{F}^{\mathrm{T}}\mathbf{R}^{-1}\boldsymbol{y} \tag{3}$$

$$\hat{\sigma}^2 = \frac{1}{m}\big(\boldsymbol{y} - \mathbf{F}\hat{\beta}\big)^T\boldsymbol{R}^{-1}(\boldsymbol{y} - \mathbf{F}\hat{\beta}) \tag{4}$$

Table 1 Commonly used classical one-dimensional ACF models

Function type	Function expression	Corresponding WM model
Single exponential	$\rho = \exp(-\theta \cdot \lvert d \rvert)$	$\nu = 1/2$
Whittle	$\rho = \theta \cdot \lvert d \rvert \cdot K_1(\theta \cdot \lvert d \rvert)$	$\nu = 1$
Second-order Markov	$\rho = \exp(-\theta \cdot d) \cdot (1 + \theta \cdot d)$	$\nu = 3/2$
Squared exponential	$\rho = \exp(-\theta \cdot d^2)$	$\nu = \infty$

Note: K_1 means modified Bessel function of the second kind with the first order; and d is the separation distance between two points

where $\mathbf{F} = [\boldsymbol{f}_1(\boldsymbol{x}_1), \boldsymbol{f}_2(\boldsymbol{x}_2), \dots, \boldsymbol{f}_m(\boldsymbol{x}_m)]^T$ is a vector of $\boldsymbol{f}(\boldsymbol{x})$, with a dimension of $m \times p$; and $\mathbf{R}$ denotes the autocorrelation matrix and is expressed as

$$\mathbf{R} = \begin{pmatrix} \rho(\boldsymbol{x}_1, \boldsymbol{x}_1) & \cdots & \rho(\boldsymbol{x}_1, \boldsymbol{x}_m) \\ \vdots & \ddots & \vdots \\ \rho(\boldsymbol{x}_m, \boldsymbol{x}_1) & \cdots & \rho(\boldsymbol{x}_m, \boldsymbol{x}_m) \end{pmatrix} \tag{5}$$

It is noted from Eqs. (3) and (4) that the determination of $\boldsymbol{\beta}$ and σ^2 depend highly on the autocorrelation structure, and according to some further deduction, the MLE problem above is finally changed as minimizing the following objective function as

$$\psi(\boldsymbol{\theta}) = \frac{1}{2}(m \ln \sigma^2 + \ln|\mathbf{R}|) \tag{6}$$

After obtaining the parameters $\boldsymbol{\theta}$, $\boldsymbol{\beta}$ and σ^2, the best linear unbiased prediction $\hat{y}(\boldsymbol{x}')$ and the Kriging variance $\hat{\sigma}^2(\boldsymbol{x}')$ for the unknown point $\boldsymbol{x}'$ can be estimated as

$$\hat{y}(\boldsymbol{x}') = \boldsymbol{f}(\boldsymbol{x}')^{\boldsymbol{T}}\widehat{\boldsymbol{\beta}} + \boldsymbol{r}(\boldsymbol{x}')^T\mathbf{R}^{-1}(\boldsymbol{y} - \mathbf{F}\widehat{\boldsymbol{\beta}}) \tag{7}$$

$$\hat{\sigma}^2(\boldsymbol{x}') = \sigma^2(1 + \boldsymbol{u}^T\left(\mathbf{F}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{F}\right)^{-1}\boldsymbol{u} - \boldsymbol{r}(\boldsymbol{x}')^T\mathbf{R}^{-1}\boldsymbol{r}(\boldsymbol{x}')) \tag{8}$$

where $\boldsymbol{r}(\boldsymbol{x}') = [\rho(\boldsymbol{x}', \boldsymbol{x}_1), \rho(\boldsymbol{x}', \boldsymbol{x}_1), \dots, \rho(\boldsymbol{x}', \boldsymbol{x}_m)]$ is the correlation vector between the unknown point $\boldsymbol{x}'$ and all known training data $(\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_m)$; and $\boldsymbol{u} = \mathbf{F}^{\mathrm{T}}\mathbf{R}^{-1}\boldsymbol{r}(\boldsymbol{x}') - \boldsymbol{f}(\boldsymbol{x}')$.

2.2 Whittle-Matérn autocorrelation structure

As mentioned above, the autocorrelation structure is an important part for formulating the Kriging method. It governs the spatial correlation between two points, which decreases as the relative distance increases. There are several choices of ACF models in Kriging, and Kriging is usually used by randomly selecting one ACF from some candidate ACFs to fit the given data. Consider, for example, the one-dimensional classical ACFs listed in Table. 1 (Ching *et al.* 2019, Li *et al.* 2015, Minasny and McBratney 2005). It is seen from the table that the function forms for different ACFs are different. The autocorrelation structure underlying the given data is determined once the ACF type and the associated correlation parameter, θ, are known. Note that the variable form θ and vector form $\boldsymbol{\theta}$ of the correlation parameter are used interchangeably hereafter, as they are representing problems with different dimensions.

Obviously, different ACFs would lead to different autocorrelation structures and thus resulting in different Kriging predictions. Therefore, traditionally Kriging method with randomly selected ACF may introduce extra model uncertainty to slope reliability analysis. In addition, within the classical ACF models in Table 1, only the correlation parameter θ can be optimized to fit the underlying data. However, as proposed by Ching *et al.* (2019), the spatial structure underlying the data may not be fully depicted by only one correlation parameter. For this situation, the non-classical ACF models with more than one correlation parameter may perform reasonably well, which, however, is rarely seen in the current Kriging metamodeling.

For example, it is advantageous to use the two-parameter ACF since this kind of model can be extended to fit the observed data with not only the correlation parameter but also the smoothness parameter (Ching *et al.* 2019). The smoothness parameter gives the two-parameter ACF greater flexibility for modeling the spatial correlation (Gneiting *et al.* 2010). The two-parameter models usually adopted in the literature include the WM model and the powered exponential model. Given WM model exhibits a more flexible ability for fitting the smoothness feature (Ching *et al.* 2019), it is adopted in this study for constructing the Kriging metamodel, and the expression taken from *Ching et al.* (2019) is written as

$$\rho(d) = \frac{2}{\Gamma(\nu)} \cdot \left(\frac{\sqrt{\pi} \cdot \Gamma(\nu + 0.5) \cdot |d| \cdot \theta}{\Gamma(\nu)}\right) K_\nu \left(\frac{2\sqrt{\pi} \cdot \Gamma(\nu + 0.5) \cdot |d| \cdot \theta}{\Gamma(\nu)}\right) \tag{9}$$

where ν is the smoothness parameter with a range of $0 < \nu \leq \infty$; θ is the correlation parameter as the same in Table 1; $\Gamma(\cdot)$ denotes the Gamma function; and $K_\nu(\cdot)$ is the modified Bessel function of the second kind with order ν. The parameter ν controls the smoothness of the spatial data. The spatial process is smooth when ν is large (e.g., $\nu \to \infty$); otherwise, it is relatively rough when ν is small (e.g., $\nu \to 0$). Within the Matérn correlation family, different ν values represent different classical single-parameter ACF models. As shown in the last column in Table 1, the commonly used single exponential (SExp), Whittle, second-order Markov (SMK), and squared exponential (QExp) ACFs can be replaced by the WM model with ν= 1/2, 1, 3/2, and ∞, respectively. This observation shows that the optimal ACF for Kriging metamodeling can be achieved easily by simultaneously optimizing the two parameters of WM function. Likewise, the uncertainty in ACF selection in traditional Kriging

method is reduced. With the WM function, the objective function in Eq. (6) is, therefore, rewritten as

$$\psi(\boldsymbol{\theta},\boldsymbol{\nu}) = \frac{1}{2}(m\ln\sigma^2 + \ln|\boldsymbol{R}|) \tag{10}$$

where ψ is a function of $\boldsymbol{\theta}$ and $\boldsymbol{\nu}$, which, however, is only the function of $\boldsymbol{\theta}$ in the traditional Kriging method.

It should be noted that the WM function in Eq. (9) is one-dimensional. However, according to Satria Palar et al. (2020), it is straightforward to extend it to multi-dimension by using the following composite function as

$$\rho(\boldsymbol{\theta},\boldsymbol{\nu},\boldsymbol{\omega}) = \sum_{i=1}^{K}\omega_i\prod_{j=1}^{m}(\rho_i(\theta_j,\upsilon_j)) \tag{11}$$

where K is the number of individual ACF types (e.g., a WM function with a specific ν value is considered an individual ACF) used in the composite function; m denotes the input dimension; $\boldsymbol{\omega} = (\omega_1,\omega_2,\dots,\omega_K)^T$ is the weight vector for the individual ACF types that should satisfy the equality constraint of $\sum_{i=1}^{K}\omega_i = 1$. In the current study, only one kind of individual ACF is adopted, which indicates that K is equal to one here. Therefore, ρ for multidimensional problem in Eq. (11) reduces to the product of m one-dimensional WM functions.

2.3 Genetic algorithm for parameter optimization

In general, the aforementioned DACE toolbox is often used to construct the Kriging metamodel. However, the built-in pattern search method (PSM) in DACE for optimizing the correlation parameters of the objective function of Eq. (6) is a kind of local optimization method, which is prone to fall into local minima. As suggested by Liu *et al.* (2017), the GA well outperforms the PSM in locating the minimum of the objective function of Eq. (6). Therefore, the global optimization method, GA, is used to search for the optimal correlation and smoothness parameters of the WM function of Eq. (10) to calibrate the Kriging metamodel of slope stability analysis in this study.

GA is an optimization tool inspired by natural selection (Gomes *et al.* 2011). It searches for the optimum solution based on the evolution of the generations. Compared with the local optimization methods, GA searches for the optimal solution on a global scale and avoids falling into local optimization. GA process starts with the candidate solution from the initial generation and evaluates the quality of the chromosomes by the objective function. The genes from the parent generation are recombined in the mating pool to produce the next generation (Kavzoglu *et al.* 2015). The iteration terminates when the generation reached the predefined maximum breeding generation. For the realization of the GA, the parameters within the algorithm are critical to the optimization result, such as the initial population (NIP), the maximum breeding generation (MBD), the generation gap (GGAP), the crossover probability (CP), and the mutation probability (MP). In general, these parameters are determined by extensive trials and errors with the consideration of the balance between accuracy and computational effort. Herein, the parameter setting of GA is selected based on numerical experiments and Liu's work (Liu *et al.* 2017).

Table 2 Parameter setting of GA in this study

Parameter	NIP	MBD	GGAP	CP	MP
Value	50	100	0.99	0.7	0.01

It is worthwhile to point out that, for convenience, DACE is also used as the major tool for constructing the Kriging metamodel in this study, but with some modifications. The modifications are two-fold: (1) Eq. (10) is used to replace Eq. (6) as the final objective function for a general representation of the autocorrelation structure that is governed by the correlation and smoothness parameters; and (2) GA is developed to replace the PSM to adaptively optimize the correlation and smoothness parameters to obtain the optimal autocorrelation structure. More details about implementing these modifications on DACE are introduced in the "Implementation procedure of the proposed method" section.

3. Monte Carlo simulation based on Kriging metamodel for slope reliability analysis

Monte Carlo simulation is considered as a robust method for slope reliability analysis due to its conceptual simplicity (Wang *et al.* 2020). Therefore, after obtaining the Kriging metamodel of a slope stability analysis model, MCS is performed on the Kriging metamodel, rather than on the original slope stability analysis model, to calculate the P_f of the slope to improve the computation efficiency (Zhang *et al.* 2020). The P_f is defined as the probability that the factor of safety (FS) is less than unity, and can be calculated as

$$P_f = \frac{1}{N_{mcs}}\sum_{i=1}^{N_{mcs}} I(FS_i < 1) \tag{12}$$

where FS_i denotes the FS calculated from the Kriging metamodel (i.e., Eq. (7)) for the ith MCS sample; N_{mcs} denotes the number of the MCS samples, which should be reasonably large (e.g., in the order of magnitude of 10^{-4}) to ensure the accuracy of the estimation of P_f; and $I(\cdot)$ is an indicator function, which is equal to unity when FS is less than one and zero otherwise.

4. Implementation procedure of the proposed method for slope reliability analysis

To facilitate the understanding and application of the proposed method in slope reliability analysis, this section introduces the implementation procedure of the method in detail. As stated before, the DACE toolbox is used as the major tool for realizing the proposed method. For convenience, the implementation procedure is schematically plotted in Fig. 1, which is composed of five steps. Each of the five steps is described as follows.

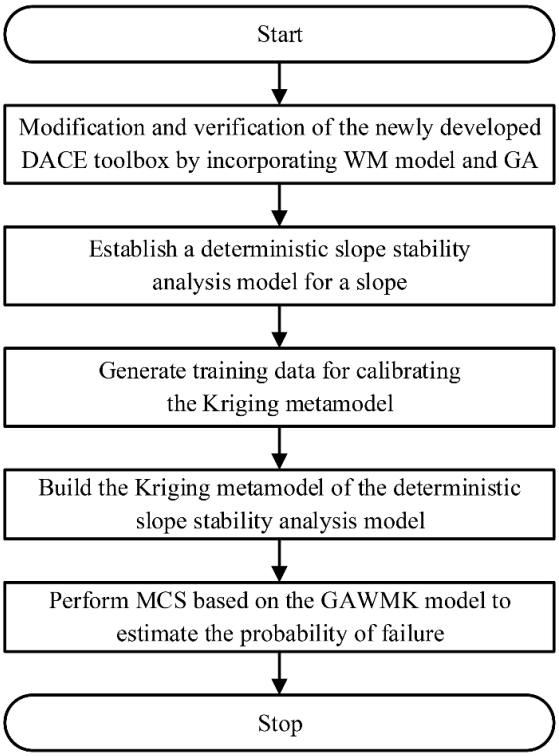


Fig. 1 Flow chart of the GAWMK model

Step 1: Modification and verification of the newly developed DACE toolbox by incorporating WM model and GA. At this step, a new sub-function $@corr\_WM$ is first developed to describe the WM structure in the DACE toolbox. Then, GA is coded to replace the PSM that is embedded in DACE for parameter optimization. Finally, the whole GAWMK method is realized by invoking a newly developed main function based on DACE as $[dmodel, perf] = dacegafit\_WM(\boldsymbol{S}, \boldsymbol{y}, @regr, @corr\_W$, where $\boldsymbol{S}$ and $\boldsymbol{y}$ keep the same meaning as those described in the second section; and $@regr$ is denoted as the trend function described in Eq. (1), which is selected as the constant function herein, i.e., $f(x) = 1$ ($\boldsymbol{f}(\boldsymbol{x})$ in Eq. (1) is degenerated to a variable as p = 1 for constant function), representing the ordinary Kriging model. Besides, $dmodel$ denote the calibrated Kriging model $perf$ saves the information about the optimization process. Thereafter, the adapted DACE toolbox is subsequently validated by two simple analytical examples, as will be illustrated in the following section, before its application to slope reliability analysis.

Step 2: Establish a deterministic slope stability analysis model for a slope. This step mainly includes collecting geometrical and geotechnical parameters to perform conventional slope stability analysis for a slope. Both the commonly used finite element method and limit equilibrium method (LEM) can be used to establish the slope stability analysis model. Since LEM can improve the computational efficiency and practicality of slope system reliability analysis by implementing relatively simple stability analysis methods (Duncan 1996, Liu *et al*. 2020), the LEM is adopted in this study.

Step 3: Generate training data for calibrating the Kriging metamodel. This requires to first generate a number of training samples to constitute the matrix $\boldsymbol{S} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_m]^T$. Latin hypercube sampling (LHS) based on the statistics of soil properties is used for this purpose in this study, where the number of training samples is determined based on the rule of thumb proposed by Kang *et al.* (2015) and Silvestrini *et al.* (2013). Then, the training samples are repeatedly substituted into the slope stability analysis model established in Step (3) to calculate the corresponding FS to obtain the response vector $\boldsymbol{y} = [y_1, y_2, \ldots y_m]^T$. The training samples $\boldsymbol{S}$ and associated responses $\boldsymbol{y}$ finally constitute the training data for calibrating the Kriging models in the next step.

Step 4: Build and validate the Kriging metamodel of the deterministic slope stability analysis model. At this step, the adapted DACE toolbox in Step (1) is first used to construct the Kriging model by substituting the training data obtained in the last step into the $dacegafit\_WM$ function to calibrate the unknown parameters. Then, the calibrated Kriging model is validated by applications to a certain number of testing data that can be formulated in the same way as the training data through the coefficient of determination (R^2). In the adapted DACE toolbox, the testing process is realized through the prediction function embedded in DACE as $[\hat{\boldsymbol{y}}, \hat{\boldsymbol{\sigma}}^2] = predictor(\boldsymbol{x}', dmodel)$, where $\hat{\boldsymbol{y}}$, $\hat{\boldsymbol{\sigma}}^2$ and $\boldsymbol{x}'$ keep the same meaning as those described in the second section and $dmodel$ saves the model information generated from the function $dacegafit\_WM$, as described in Step (1).

Step 5: Perform MCS based on the GAWMK model to estimate the probability of failure. This includes generating

a large number of samples based on predefined statistics of soil properties and substituting them into the established GAWMK metamodel in Step (4) to obtain the corresponding FS values. The probability of failure is then evaluated based on Eq. (10).

5. Validation of the proposed method and the modified DACE toolbox

Before directly applying the proposed method for slope reliability analysis, this section provides an analytical example to validate the effectiveness of the proposed GAWMK method and the improved DACE toolbox that is used to implement the method. The example is an explicit one-dimensional cubic function, which has been studied by Liu *et al*. (2017). The function is expressed as

$$y = x^3 + x^2 - 6 \tag{13}$$

where x is subjected to standard normal distribution. Following Liu *et al.* (2017), ten training samples $\boldsymbol{S} = [x_1, x_2, \ldots x_{10}]^T$ based on LHS and the statistics of x are also generated and substituted to Eq. (13) to obtain the corresponding responses $\boldsymbol{y} = [y_1, y_2, \ldots y_{10}]^T$, which consitute the training data here to calibrate the Kriging metamodel of the function. First, to intuitively illustrate the effect of ACF on Kriging metamodeling, the commonly used ACFs in Table 1 are compared in terms of the resulting Kriging metamodels of Eq. (13) with the same training data. Fig. 2 plots the Kriging metamodels based on different ACFs as well as the true function shape of Eq. (13). Note that, unless otherwise specified, all results presented from here on are obtained by using the improved DACE toolbox. Therefore, GA is also used to optimize the correlation parameters for these ACFs. As can be seen from the figure, different ACF models generally produce different Kriging results. The QExp model exhibits the best prediction among the four ACFs since the value predicted by the QExp model almost coincides with the actual y value. By contrast, the SExp, Whittle, and SMK models show relatively poor prediction performance. This observation indicates that the performance of Kriging models would be significantly affected by ACFs.

Besides, Fig. 2 also shows that the classical single parameter ACFs can be uniformly characterized by the WM function through varying the smoothness parameter ν. Therefore, the influence of ACF on Kriging metamodeling can be indirectly reflected by the smoothness parameter in WM function. Fig. 3 shows the variations of the minimum value of the objective function ψ, $\psi_{\min}$, and the optimal θ against various ν values for the same number of training samples. The result shows that both the $\psi_{\min}$ and θ change significantly when the smoothness parameter ν changes. The ψ reaches its minimum (about 15) neither at a small ν value nor at a large value, while the optimal θ varies rapidly for different ν values. This implies that the WM function can represent different autocorrelation structures with both the smoothness parameter and correlation parameter, which validates the feasibility of the proposed method.

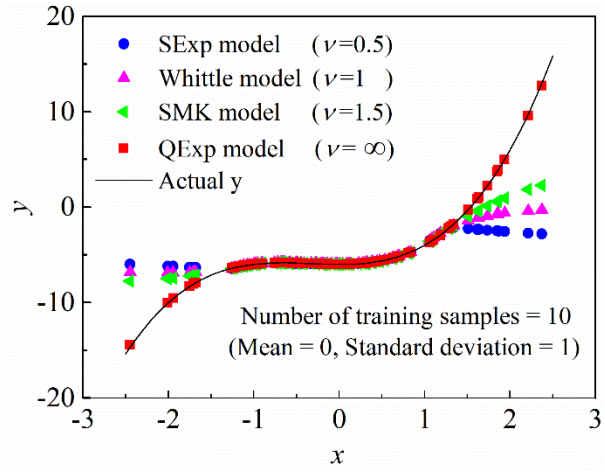


Fig. 2 Performance of Kriging metamodels based on different ACFs with ten training samples

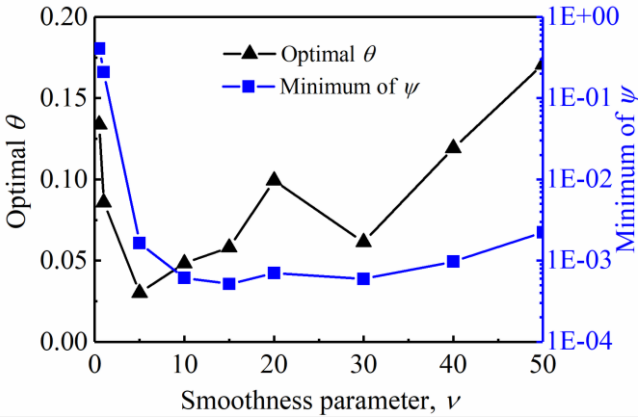
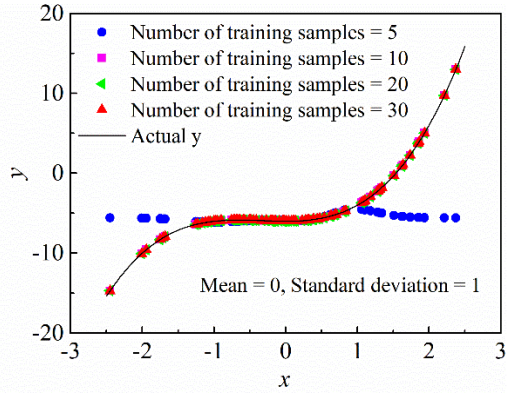


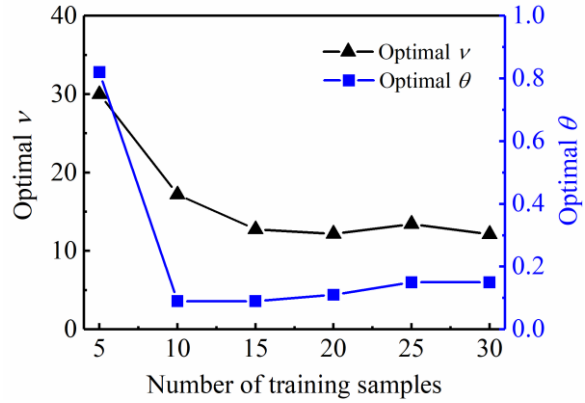
Fig. 3 Variation of ψ and θ with smoothness parameter in WM model with ten training samples

Then, the proposed GAWMK method is used to construct the Kriging metamodel of Eq. (13), and the results are schematically plotted in Fig. 4. Fig. 4(a) presents the prediction results by the proposed method when different numbers of training samples are used. The actual y function is also shown in the subfigure. It is generally seen from the figure that the proposed GAWMK method fits the actual y function well when the number of training samples is larger than 10. This observation can be easily reasoned from the optimization results of the objective function in Eq. (10), as shown in Fig. 4(b). The results in Fig. 4(b) illustrate that the optimal correlation parameter θ and smoothness parameter ν remain nearly unchanged when the number of training samples is larger than 10.

However, it might be argued that the prediction results obtained by the Kriging models based on the classical ACFs, e.g., the QExp function, are also compared with the proposed method, as presented in the previous work by Liu *et al.* (2017). Hence, to further illustrate the superiority of the proposed method over the available study, Fig. 5 compares the optimized minimum values of ψ for different methods under different numbers of training samples. For simplicity, only the commonly used QExp and SExp functions are compared with the proposed method. As can



(a) Prediction results of the GAWMK for different number of training samples



(b) Parameters optimized by GAWMK for different number of training samples

Fig. 4 Modelling results of GAWMK for different number of training samples

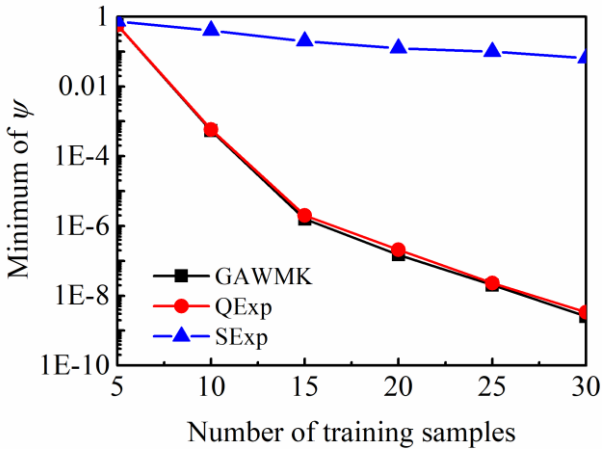
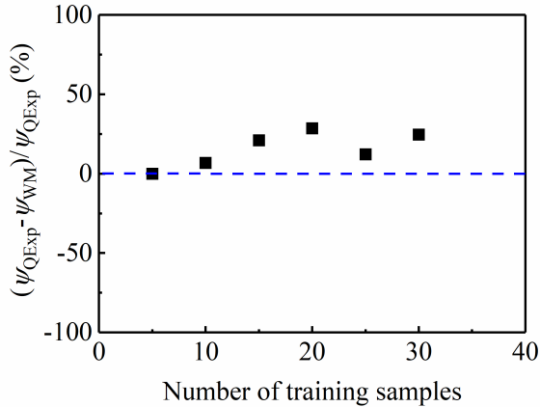
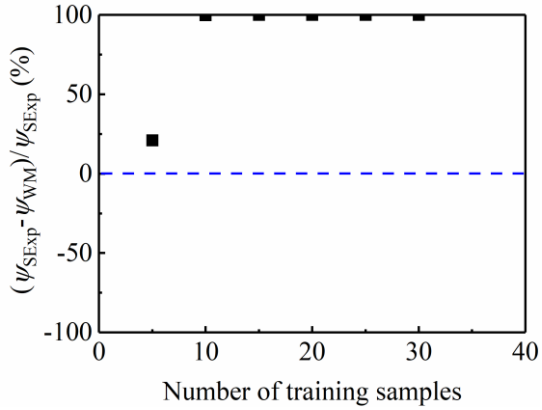


Fig. 5 Minimum ψ values optimized based on different ACFs against the number of training samples

be seen from Fig. 4, $\psi_{\min}$ for all three ACFs decreases with increasing the number of training samples. This observation is expected, as more data would yield more accurate Kriging models. In addition, it is observed that the results obtained by GAWMK and QExp are much smaller than those by SExp, suggesting that the Kriging model based on QExp is less accurate than that based on the other two ACFs. Although the difference between QExp and



(a) Relative differences of $\psi_{\min}$ between QExp and GAWMK



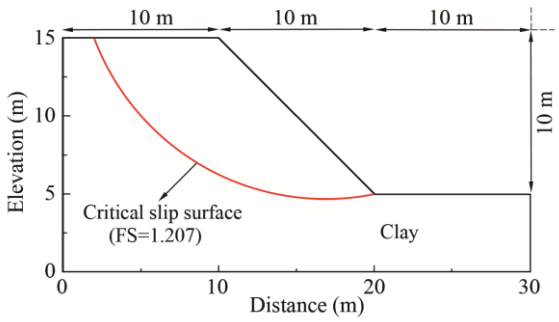
(b) Relative differences of $\psi_{\min}$ between SExp and GAWMK

Fig. 6 Relative differences of $\psi_{\min}$ between GAWMK and two commonly used single parameter ACFs for different numbers of training samples

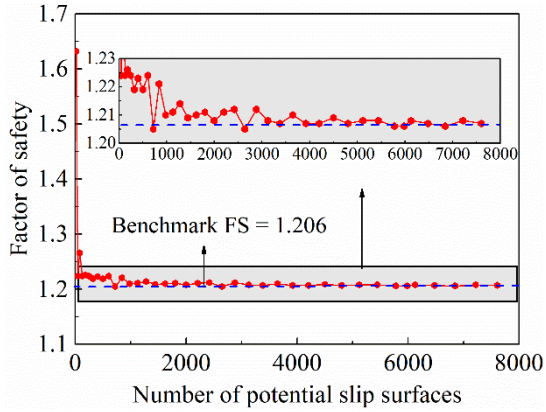
GAWMK is small, it is still distinguishable from the figure that GAWMK is more accurate than QExp. Such differences among different ACFs can be further quantified by the relative difference of $\psi_{\min}$. The results are plotted in Fig. 6 by solid squares, where the ψ_{SExp}, ψ_{QExp} and ψ_{WM} denotes the minimum ψ based on SExp, QExp and GAWMK, respectively. It is seen from the figure that the relative difference results between GAWMK and SExp (Fig. 6(a)) are relatively small but can also be as large as about 30% when the number of training samples is 20. By contrast, the relative difference results between GAWMK and QExp (Fig. 6(b)) are significantly large e.g., nearly 100% when the number of training samples is larger than 10. These results are consistent with the examination of Fig. 5 by naked eyes. Overall, GAWMK performs better than SExp and QExp from the perspective of accuracy and robustness.

6. Applications to slope reliability analysis

To illustrate the proposed method for slope reliability analysis, two slope examples are taken into account in this section. The SExp (Cho 2007, Tabarroki *et al.* 2013,



(a) Slope geometry and deterministic slope stability analysis results based on mean soil properties



(b) Factor of safety versus Number of potential slip surfaces

Fig. 7 Determination of slope geometry and number of potential slip surface for Example 1

Suchomel and Mašín 2010) and QExp (Luo *et al.* 2012, Zhang *et al.* 2011, Yi *et al.* 2015), which are the most commonly used ACFs in slope reliability analysis, are adopted to compare with the GAWMK method.

6.1 Example 1: A homogeneous c-φ slope

6.1.1 Deterministic slope stability analysis

The first example concerns a homogenous *c*-φ slope, which has been analyzed by Cho (2010), Jiang *et al.* (2015), and Liu *et al.* (2017) in the literature. The slope geometry is schematically plotted in Fig. 7(a). The statistics of the soil properties are listed in Table 3. The critical deterministic FS obtained from Bishop's simplified method (BSM) with the mean values of soil properties is 1.207, which is consistent with the FS of 1.204, 1.206, and 1.206 by Cho (2010), Jiang *et al.* (2015), and Liu *et al.* (2017), respectively. The corresponding critical slip surface passes through the slope toe, as shown in Fig. 7, which also matches well with the literature. To determine the number of potential slip surfaces (N_s), Fig. 7(b) plots the influence of N_s on the FS. The benchmark FS with the value of 1.206 is plotted by the blue dashed line. In order to express the convergence of the data more clearly, the local zoomed-in view is placed in the upper part of the figure. It can be seen from Fig. 7(b) that the FS gradually converges to the benchmark and remain stable when N_s reach around 4000. Therefore, the N_s is determined as 5985 for the slope reliability analysis to keep the balance between the accuracy and efficiency.

Table 3 Statistics of soil properties for Example 1

Parameter	Mean	COV	Distribution	Correlation coefficient
Cohesion, c	10 kPa	0.3	Lognormal	−0.7
Friction angle, φ	30 °	0.2	Lognormal	
Unit weight, γ	20 kN/m^3	–	–	–

6.1.2 Probabilistic slope stability analysis

Based on the above deterministic slope stability analysis model, three Kriging metamodels based on GAWMK, QExp and SExp are subsequently constructed with 20 LHS training samples. Note that, there is currently no general guideline on determining the number of training samples for calibrating the Kriging models in the literature. This study follows the suggestion by Kang *et al.* (2015) and Silvestrini *et al.* (2013) that, as a rule of thumb, a number 10D-15D of samples are sufficient to calibrate a surrogate model, where D is the number of independent variables; and 10D is determined by considering the balance between computation accuracy and efficiency. The influence of training sample size will be discussed later. Then, the Kriging model accuracy is examined by comparing the FS values calculated from BSM and predicted by the Kriging metamodels for 100 randomly generated testing samples. The results are plotted in Fig. 8, with Figs. 8(a)-8(c) for GAWMK, QExp, and SExp, respectively. For reference, the 45-degree (or 1:1) line and coefficient of determination R^2 are also shown in each subfigure. It is generally observed that the 100 FS samples are consistent with the 45-degree line for GAWMK and QExp, as shown in Figs. 8(a) and 8(b), respectively. The FS samples for SExp in Fig. 8(c), however, present greater dispersion around the 45-degree line, compared with their counterparts in Figs. 8(a) and 8(b). These results are also coincident with the R^2 values of the three models, which are 0.9917, 0.9388, and 0.8881 for GAWMK, QExp, and SExp, respectively. Therefore, the GAWMK has a better prediction ability in slope stability FS evaluation than QExp and SExp.

MCS is subsequently conducted based on the three calibrated models above with a number of 10,000 samples to evaluate the reliability of the slope. The results for the three models are listed in Table 4. For validation, the P_f values calculated by the direct MCS based on BSM and reported in the literature are also tabulated in the table. It is seen that the proposed method provides almost the same result as the underlying true result by MCS (0.0481 vs. 0.0488). Meanwhile, the results in this study are also comparable with those reported in the literature, except for the result by SExp. Such agreement thus validates the accuracy of the proposed method.

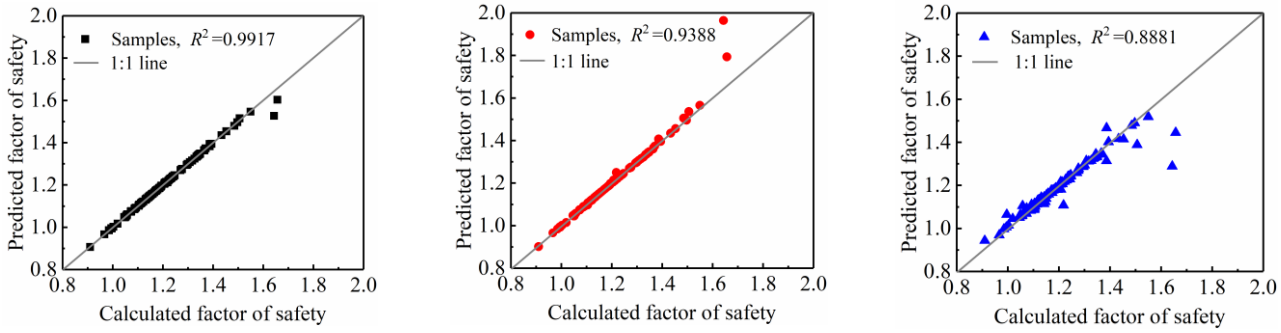
6.1.3 Influence of number of training samples on slope reliability evaluation

This subsection investigates the influence of training sample size on the slope reliability evaluation to examine the sensitivity of the method against sample size and the effectiveness of the 10D-15D rule of thumb. The variation

Table 4 Reliability results obtained from different methods for Example 1

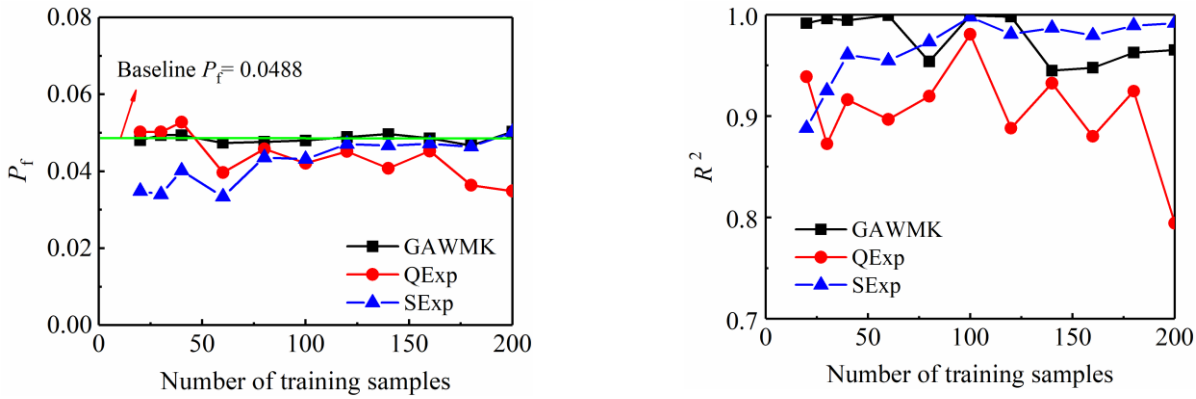
Method	Probability of failure	COV	Source
MCS with 10000 samples	0.0488	4.41%	This study
GAWMK	0.0481	1.41%	This study
QExp	0.0517	1.35%	This study
SExp	0.0349	1.66%	This study
LHS with1000 samples	0.0480	1.40%	Liu *et al.* (2017)
OK with 400 training samples	0.0366	1.62%	Liu *et al.* (2017)
TK4 with 400 training samples	0.0478	1.41%	Liu *et al.* (2017)
SQRSM	0.0420	0.50%	Li *et al.* (2016)
MQRSM	0.0430	0.50%	Li *et al.* (2016)

Note: OK means ordinary Kriging; TK4 means the fourth-order Taylor Kriging; SQRSM means single quadratic response surface method; and MQRSM means multiple quadratic response surface method



(a) GAWMK (b) QExp (c) SExp

Fig. 8 Comparison between FSs calculated by BSM and predicted by Kriging models for Example 1 given 20 training samples



(a) Variation of P_f with the number of training samples (b) Variation of R^2 with the number of training samples

Fig. 9 Effects of training samples size on probability of failure P_f and coefficient of determination R^2 of Example 1

of P_f estimated based on SExp, QExp, and GAWMK with the training sample size is plotted in Fig. 9(a). The baseline result, which is calculated by direct MCS with 1×10^4 samples, is also shown in the figure for reference. It can be seen from the figure that the P_f estimated by GAWMK remains nearly unchanged and is almost overlapped with the baseline result when the number of training samples increases. Such agreement shows that the proposed method is robust against the sample size. Meanwhile, the method provides accurate reliability estimation when $10D$ training samples are used. The results obtained from the other two models, however, seem to vary with the number of training samples. The implication is that the underlying spatial correlation structure of the FS data might not be accurately represented by the classical single parameter ACF, even though the sample size is increased to a large value. Generally, the accuracy of the model by SExp increases with the increase of the sample size, but the $10D$-$15D$ rule of thumb may not work well for SExp. For QExp, the $10D$-$15D$ rule of thumb is applicable, but the increase in sample

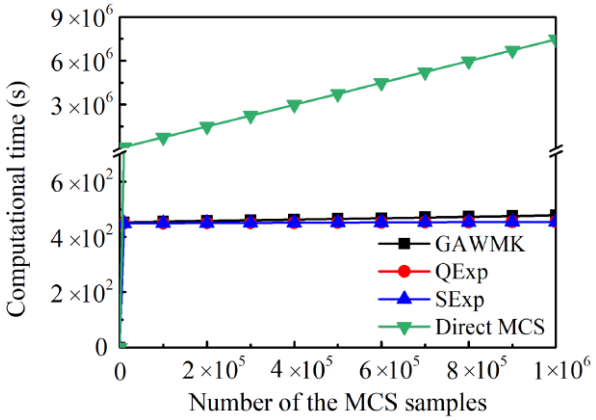


Fig. 10 Comparison of computation efficiency between the Kriging-based MCS and direct MCS for Example 1

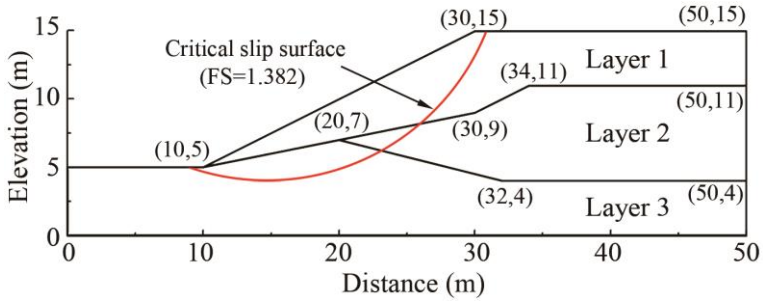


Fig. 11 Slope geometry and slope stability analysis results based on mean soil properties for Example 2

Table 5 Statistics of soil properties for Example 2

Slope layer	Unit weight, γ (kN/m^3)	Cohesion, c (kPa)		Friction angle, φ (°)	
		Mean	COV	Mean	COV
Layer 1	19.5	0	–	38	–
Layer 2	19.5	5.3	0.3	23	0.2
Layer 3	19.5	7.2	0.3	20	0.2

size may not ensure a high accuracy. These observations are further validated by the R^2 values of the three models for different numbers of training samples, which present a similar variation pattern as the P_f, as plotted in Fig. 9(b). Therefore, the proposed GAWMK model can be well established with only $10D$ training samples and is robust against different numbers of training samples.

6.1.4 Computation efficiency of the proposed method

In order to explore the computation efficiency of the proposed method, the time cost by Kriging-based MCS and direct MCS with different sample sizes are compared. The computation experiments are performed on a desktop with Intel Core i7-8700, 3.20 GHz processor, and 16 GB RAM. It can be observed in Fig. 10 that the time cost by the Kriging-based method is far less than the direct MCS, which is expected. Besides, GAWMK is less efficient than QExp and SExp because the number of optimization parameters increases in the WM model. However, compared with QExp and SExp, the additional time spent by GAWMK is not significant and can be negligible.

6.2 Example 2: A three-layered c-φ slope

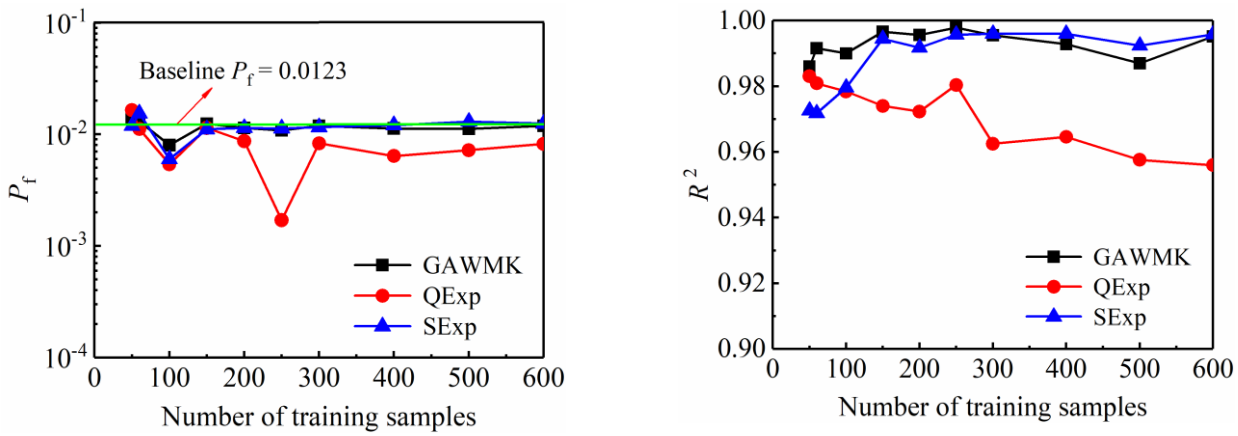
This example is a complex three-layered c-φ slope, which has been examined by Ji and Low Bak (2012) and Liu and Cheng (2016). The geometry of the slope is shown in Fig. 11. The statistics of soil properties are given in Table 5, where only the shear strengths of Layer 2 and 3 are considered as random variables. The factor of the safety for slope is calculated based on 11067 potential slip surfaces. Based on the mean values of the shear strengths, the FS is calculated as 1.382 using Spencer method, and the critical slip surface is plotted in Fig. 11. Based on the deterministic slope stability model, probabilistic slope stability analysis of the slope is conducted subsequently and the results are presented as follows.

Similarly, three Kriging models are first built for this slope based on 60 (i.e., $15D$) training samples. MCS with 1×10^5 samples is then conducted based on the three models to obtain the P_f of the slope. For reference, the results along with those from direct MCS with 1×10^4 samples and literature are displayed in Table 6. It should be noted that

Table 6 Reliability results obtained from different methods for Example 4

Methodology	Probability of failure	COV	Source
MCS	0.0123	8.96%	This study
GAWMK	0.0131	2.75%	This study
QExp	0.0112	2.97%	This study
SExp	0.0154	2.53%	This study
GPR-based RSM	0.0159	2.50%	Kang *et al.* (2015)
QRSM	0.0138	2.67%	Liu and Cheng (2016)
QRSM with cross terms	0.0148	2.58%	Liu and Cheng (2016)
SRSM	0.0155	2.52%	Liu and Cheng (2016)

Note: QRSM means quadratic response surface method; SRSM means stochastic response surface method



(a) Variation of P_f with the number of training samples (b) Variation of R^2 with the number of training samples

Fig. 12 Effects of training samples size on probability of failure P_f and coefficient of determination R^2 of Example 4

the samples are drawn between the range of [μ-3σ, μ+3σ] to avoid negative values, where μ and σ denote the mean value and standard deviation of a considered normal variable, respectively. As can be seen from the table, the results of this study are generally comparable with those from direct MCS. This shows that the proposed method can accurately estimate the reliability of the slope based on 60 training samples. To gain more insights into the robustness of the proposed method, Fig. 12(a) plots the variation of P_f against the number of training samples for the three models. It is observed that (1) the accuracy of the model by SExp and GAWMK shows ideal performance; and (2) for QExp, the increase of sample size may not ensure a stable accuracy as the SExp and GAWMK. These observations are expected from the R^2 values of the three models for different numbers of training samples, which present a similar variation pattern as the P_f, as plotted in Fig. 12(b). Overall, the proposed method generally outperforms traditional Kriging methods in accuracy, and the method is more robust against the variation of the training sample size.

7. Conclusions

This paper proposed an adaptive Kriging method for slope reliability analysis. The method employs WM function to describe the spatial correlation structure of observed data and uses a global GA method to optimize the correlation and smoothness parameters of the WM function. Formulations of the proposed method are derived, and an adapted Matlab toolbox based on DACE is developed to implement the proposed method. Finally, the proposed method is illustrated and validated through an analytical example and two slope reliability analysis examples. The following conclusions are made from the study as:

- Unlike the traditional Kriging method where the QExp or SExp is randomly selected, the proposed method can fit the data much better with better optimization results. The proposed method, therefore, overcomes the deficiency of the traditional Kriging method in randomly choosing ACF, while improving the accuracy of the Kriging model.
- The proposed method can accurately estimate slope reliability for different slope examples. It is robust against the size of training samples, which is not the case for the commonly used QExp and SExp models. For the commonly used ACFs, from the perspective of robustness, the SExp-based model performs better than QExp-based model in the estimation of the probability of failure for the considered slope examples. In addition, the 10D-15D as a rule of thumb is further validated to be applicable to the proposed method.
- Compared with the traditional direct MCS method, the

efficiency of the proposed method is greatly enhanced. Considering the number of optimization parameters is doubled in the proposed method, the time cost of the method is slightly higher than the normal Kriging models using classical ACFs. Nevertheless, such additional computation resources are negligible and acceptable in practical slope reliability analysis.

Acknowledgments

The work described in this paper was funded by grants from the National Natural Science Foundation of China (Project No. 41902291), the Natural Science Foundation of Hunan Province, China (Project No. 2020JJ5704), and the Open Research Fund Program of Key Laboratory of Metallogenic Prediction of Nonferrous Metals and Geological Environment Monitoring (Central South University), Ministry of Education (Project No. 2020YSJS21), and the Fundamental Research Funds for Central Universities of the Central South University (Project No. 2021zzts0268). The financial support is greatly acknowledged.

References

Au, S.K., Cao, Z.J. and Wang, Y. (2010), "Implementing advanced Monte Carlo simulation under spreadsheet environment", *Struct. Saf.*, **32**(5), 281-292. https://doi.org/10.1016/j.strusafe.2010.03.004.

Bai, T., Yang, H., Chen, S. and Zhang, S. (2020), "In-situ monitoring and reliability analysis of an embankment slope with soil variability", *Geomech. Eng.*, **23**(3), 261-273. https://doi.org/10.12989/gae.2020.23.3.261.

Ching, J. and Phoon, K.K. (2019), "Impact of autocorrelation function model on the probability of failure", *J. Eng. Mech.*, **145**, 04018123. https://doi.org/10.1061/(ASCE)EM.1943-7889.0001549.

Ching, J., Phoon, K.K., Stuedlein, A.W. and Jaksa, M. (2019), "Identification of sample path smoothness in soil spatial variability", *Struct. Saf.*, **81**, 101870. https://doi.org/10.1016/j.strusafe.2019.101870.

Cho, S.E. (2007), "Effects of spatial variability of soil properties on slope stability", *Eng. Geol.*, **92**(3-4), 97-109. https://doi.org/10.1016/j.enggeo.2007.03.006.

Cho, S.E. (2010), "Probabilistic assessment of slope stability that considers the spatial variability of soil properties", *J. Geotech. Geoenviron. Eng.*, **136**(7), 975-984. https://doi.org/10.1061/(ASCE)GT.1943-5606.0000309.

Cressie, N. (1990), "The origins of kriging", *Math. Geol.*, **22**(3), 239-252. https://doi.org/10.1007/BF00889887.

Duncan James, M. (1996), "State of the art: Limit equilibrium and finite-element analysis of slopes", *J. Geotech. Eng.*, **122**(7), 577-596. https://doi.org/10.1061/(ASCE)0733-9410(1996)122:7(577).

El Haj, A.K. and Soubra, A.H. (2020), "Efficient estimation of the failure probability of a monopile foundation using a Kriging-based approach with multi-point enrichment", *Comput. Geotech.*, **121**(121), 103451. https://doi.org/10.1016/j.compgeo.2020.103451.

Gaspar, B., Teixeira, A.P. and Soares, C.G. (2014), "Assessment of the efficiency of Kriging surrogate models for structural reliability analysis", *Probab. Eng. Mech.*, **37**, 24-34. https://doi.org/10.1016/j.probengmech.2014.03.011.

Gneiting, T., Kleiber, W. and Schlather, M. (2010), "Matérn cross-covariance functions for multivariate random fields", *J. Am. Stat. Assoc.*, **105**(491), 1167-1177. https://doi.org/10.1198/jasa.2010.tm09420.

Gomes, H.M., Awruch, A.M. and Lopes, P.A.M. (2011), "Reliability based optimization of laminated composite structures using genetic algorithms and artificial neural networks", *Struct. Saf.*, **33**(3), 186-195. https://doi.org/10.1016/j.strusafe.2011.03.001.

Ji, J. and Low, B.K. (2012), "Stratified response surfaces for system probabilistic evaluation of slopes", *J. Geotech. Geoenviron. Eng.*, **138**(11), 1398-1406. https://doi.org/10.1061/(ASCE)GT.1943-5606.0000711.

Jiang, S.H., Li, D.Q., Cao, Z.J., Zhou, C.B. and Phoon, K.K. (2015), "Efficient system reliability analysis of slope stability in spatially variable soils using Monte Carlo simulation", *J. Geotech. Geoenviron. Eng.*, **141**(2), 04014096. https://doi.org/10.1061/(ASCE)GT.1943-5606.0001227.

Kang, F., Han, S., Salgado, R. and Li, J. (2015), "System probabilistic stability analysis of soil slopes using Gaussian process regression with Latin hypercube sampling", *Comput. Geotech.*, **63**, 13-25. https://doi.org/10.1016/j.compgeo.2014.08.010.

Kavzoglu, T., Kutlug Sahin, E. and Colkesen, I. (2015), "Selecting optimal conditioning factors in shallow translational landslide susceptibility mapping using genetic algorithm", *Eng. Geol.*, **192**, 101-112. https://doi.org/10.1016/j.enggeo.2015.04.004.

Kleijnen, J.P.C. (2009), "Kriging metamodeling in simulation: A review", *Eur. J. Operational Res.*, **192**(3), 707-716. https://doi.org/10.1016/j.ejor.2007.10.013.

Li, D.Q., Qi, X.H., Phoon, K.K., Zhang, L.M. and Zhou, C.B. (2014), "Effect of spatially variable shear strength parameters with linearly increasing mean trend on reliability of infinite slopes", *Struct. Saf.*, **49**, 45-55. https://doi.org/10.1016/j.strusafe.2013.08.005.

Li, D.Q., Jiang, S.H., Cao, Z.J., Zhou, W. and Zhou, C.B. (2015), "A multiple response-surface method for slope reliability analysis considering spatial variability of soil properties", *Eng. Geol.*, **187**, 60-72. https://doi.org/10.1016/j.enggeo.2014.12.003

Li, D.Q., Zheng, D., Cao, Z.J. Tang, X.S. and Phoon, K.K. (2016), "Response surface methods for slope reliability analysis: Review and comparison", *Eng. Geol.*, **203**, 3-14. https://doi.org/10.1016/j.enggeo.2015.09.003.

Liu, L.L. and Cheng, Y.M. (2016), "Efficient system reliability analysis of soil slopes using multivariate adaptive regression splines-based Monte Carlo simulation", *Comput. Geotech.*, **79**, 41-54. https://doi.org/10.1016/j.compgeo.2016.05.001.

Liu, L.L. and Cheng, Y.M. (2018), "System reliability analysis of soil slopes using an advanced kriging metamodel and quasi–Monte Carlo simulation", *Int. J. Geomech.*, **18**(8), 06018019. https://doi.org/10.1061/(ASCE)GM.1943-5622.0001209.

Liu, L., Cheng, Y. and Wang, X. (2017), "Genetic algorithm optimized Taylor Kriging surrogate model for system reliability analysis of soil slopes", *Landslides*, **14**(2), 535-546. https://doi.org/10.1007/s10346-016-0736-0.

Liu, X., Li, D.Q. Cao, Z.J. and Wang, Y. (2020), "Adaptive Monte Carlo simulation method for system reliability analysis of slope stability based on limit equilibrium methods", *Eng. Geol.*, **264**, 105384. https://doi.org/10.1016/j.enggeo.2019.105384/

Lophaven, S., Nielsen, H. and Sndergaard, J. (2002), DACE - A MATLAB Kriging Toolbox.

Luo, X., Li, X., Zhou, J. and Cheng, T. (2012), "A Kriging-based hybrid optimization algorithm for slope reliability analysis", *Struct. Saf.*, **34**(1), 401-406. https://doi.org/10.1016/j.strusafe.2011.09.004.

Marchant, B.P. and Lark, R.M. (2007), "The Matérn variogram model: Implications for uncertainty propagation and sampling in geostatistical surveys", *Geoderma*, **140**(4), 337-345. https://doi.org/10.1016/j.geoderma.2007.04.016.

Minasny, B. and McBratney, A.B. (2005), "The Matérn function as a general model for soil variograms", *Geoderma*, **128**(3-4), 192-207. https://doi.org/10.1016/j.geoderma.2005.04.003.

Oliver, M.A. and Webster, R. (1990), "Kriging: a method of interpolation for geographical information systems", *Int. J. Geograph. Inform. Syst.*, **4**(3), 313-332. https://doi.org/10.1080/02693799008941549.

Phoon, K.K. (2017), "Role of reliability calculations in geotechnical design", *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, **11**(1), 4-21. https://doi.org/10.1080/17499518.2016.1265653.

Qi, X.H. and Li, D.Q. (2018), "Effect of spatial variability of shear strength parameters on critical slip surfaces of slopes", *Eng. Geol.*, **239**, 41-49. https://doi.org/10.1016/j.enggeo.2018.03.007.

Qi, X.H. and Liu, H.X. (2019a), "Estimation of autocorrelation distances for in-situ geotechnical properties using limited data", *Struct. Saf.*, **79**, 26-38. https://doi.org/10.1016/j.strusafe.2019.02.003.

Qi, X.H. and Liu, H.X. (2019b), "An improved global zonation method for geotechnical parameters", *Eng. Geol*., **248**, 185-196. https://doi.org/10.1016/j.enggeo.2018.11.013.

Satria Palar, P., Rizki Zuhal, L. and Shimoyama, K. (2020), "Gaussian process surrogate model with composite kernel learning for engineering design", *AIAA J*., **58**(4), 1864-1880. https://doi.org/10.2514/1.J058807.

Silvestrini, R.T., Montgomery, D.C. and Jones, B. (2013), "Comparing computer experiments for the Gaussian process model using integrated prediction variance", *Quality Eng.*, **25**(2), 164-174. https://doi.org/10.1080/08982112.2012.758284.

Suchomel, R. and Mašín, D. (2010), "Comparison of different probabilistic methods for predicting stability of a slope in spatially variable c–φ soil", *Comput. Geotech.*, **37**(1-2), 132-140. https://doi.org/10.1016/j.compgeo.2009.08.005.

Tabarroki, M., Ahmad, F., Banaki, R., Jha, S.K. and Ching, J. (2013), "Determining the factors of safety of spatially variable slopes modeled by random fields", *J. Geotech. Geoenviron. Eng*., **139**(12), 2082-2095. https://doi.org/10.1061/(ASCE)GT.1943-5606.0000955.

Wang, L., *et al.* (2020), "Probabilistic stability analysis of earth dam slope under transient seepage using multivariate adaptive regression splines", *Bull. Eng. Geol. Environ*., **79**(6), 2763-2775. https://doi.org/10.1007/s10064-020-01730-0.

Wang, Y., Cao, Z. and Au, S.K. (2010), "Efficient Monte Carlo simulation of parameter sensitivity in probabilistic slope stability analysis", *Comput. Geotech*., **37**(7-8), 1015-1022. https://doi.org/10.1016/j.compgeo.2010.08.010.

Yi, P., Wei, K., Kong, X. and Zhu, Z. (2015), "Cumulative PSO-Kriging model for slope reliability analysis", *Probab. Eng. Mech*., **39**, 39-45. https://doi.org/10.1016/j.probengmech.2014.12.001.

Zhang, J., Huang, H. and Phoon, K.K. (2013), "Application of the Kriging-based response surface method to the system reliability of soil slopes", *J. Geotech. Geoenviron. Eng*., **139**(4), 651-655. https://doi.org/10.1061/(ASCE)GT.1943-5606.0000801.

Zhang, J., Zhang, L. and Tang, W. (2011), "Kriging numerical models for geotechnical reliability analysis", *Soils Found.*, **51**(6), 1169-1177. https://doi.org/10.3208/sandf.51.1169.

Zhang, S., Li, Y., Li, J. and Liu, L. (2020), "Reliability analysis of layered soil slopes considering different spatial autocorrelation structures", *Appl. Sci.*, **10**(11), 4029. https://doi.org/10.3390/app10114029**.**

IC