Journal of Rock Mechanics and Geotechnical Engineering 14 (2022) 1188–1199

Contents lists available at ScienceDirect

Journal of Rock Mechanics and Geotechnical Engineering

journal homepage: www.jrmge.cn

Full Length Article

Bayesian machine learning-based method for prediction of slope failure time

Jie Zhang [a,b], Zipeng Wang [a,b], Jinzheng Hu [a,b,*], Shihao Xiao [a,b], Wenyu Shang [c]

[a] Key Laboratory of Geotechnical and Underground Engineering of the Ministry of Education, Tongji University, Shanghai, 200092, China
[b] Department of Geotechnical Engineering, Tongji University, Shanghai, 200092, China
[c] Natural Science College, Michigan State University, MI, 48825, USA

ARTICLE INFO

Article history:
Received 9 May 2021
Received in revised form
14 August 2021
Accepted 9 September 2021
Available online 20 November 2021

Keywords:
Slope failure time (SFT)
Bayesian machine learning (BML)
Inverse velocity method (INVM)

ABSTRACT

The data-driven phenomenological models based on deformation measurements have been widely utilized to predict the slope failure time (SFT). The observational and model uncertainties could lead the predicted SFT calculated from the phenomenological models to deviate from the actual SFT. Currently, very limited study has been conducted on how to evaluate the effect of such uncertainties on SFT prediction. In this paper, a comprehensive slope failure database was compiled. A Bayesian machine learning (BML)-based method was developed to learn the model and observational uncertainties involved in SFT prediction, through which the probabilistic distribution of the SFT can be obtained. This method was illustrated in detail with an example. Verification studies show that the BML-based method is superior to the traditional inverse velocity method (INVM) and the maximum likelihood method for predicting SFT. The proposed method in this study provides an effective tool for SFT prediction.

1. Introduction

Due to the difficulty in precisely obtain the physical models for revealing the effects of complex factors such as external environments, geological conditions and human activities on the slope failure (Crosta and Agliardi, 2002; Intrieri and Gigli, 2016; Kothari and Momayez, 2018; Kardani et al., 2021), data-driven phenomenological models based on deformation measurements have been widely used to predict the slope failure time (SFT) (Petley et al., 2005; Mufundirwa et al., 2010; Federico et al., 2012, 2015; Xue et al., 2014). Two categories of uncertainties may exist in the phenomenological models (Zhang et al., 2020a), i.e. the observational uncertainty caused by factors including measurement error and external disturbance (Mazzanti et al., 2015; Intrieri and Gigli, 2016; Carlà et al., 2017), and model uncertainty caused by assumptions associated with the phenomenological models (Carlà et al., 2018; Kothari and Momayez, 2018). Due to the existence of the above uncertainties, the predicted SFT calculated from a data-driven model could deviate from the actual SFT (e.g. Venter et al., 2013; Federico et al., 2015).

Recently, efforts have been made to consider the effect of uncertainties on SFT prediction. (Manconi and Giorden, 2015, 2016) assessed the confidence interval (CI) and forecast reliability of the estimated SFT through a bootstrapping resampling strategy considering the influence of the observational uncertainty. Intrieri and Gigli (2016) evaluated the reliability of the SFT prediction by comparing the predictions from several competing phenomenological models at the same time, through which the influence of the model uncertainty on time-to-failure analysis was considered. Zhang et al. (2020a) suggested a maximum likelihood method in which both the effects of observational and model uncertainties on the SFT were considered based on several simplified assumptions. When such assumptions are not valid, how to explicitly consider both model and observational uncertainties when predicting the SFT remains a challenging task.

The Bayesian machine learning (BML) refers to a data-driven method, which can improve modeling capability and prediction performance based on Bayes' theorem. As it is flexible in representing uncertainties from different sources and powerful in dealing with complex real-world data, it has been popularly adopted in different fields. The usefulness of the BML has been demonstrated in many studies (Shirzadi et al., 2017; Ching and

* Corresponding author. Key Laboratory of Geotechnical and Underground Engineering of the Ministry of Education, Tongji University, Shanghai, 200092, China.
E-mail address: tjcce_hujz@tongji.edu.cn (J. Hu).

Peer review under responsibility of Institute of Rock and Soil Mechanics, Chinese Academy of Sciences.

https://doi.org/10.1016/j.jrmge.2021.09.010

Phoon, 2019; Contreras and Brown, 2019; Wang, 2020; Ma et al., 2021).

The objective of this paper is to develop a BML-based method for probabilistic prediction of SFT, which can effectively overcome the limitations of the maximum likelihood method as suggested in Zhang et al. (2020a). The proposed method will not only provide a chance to examine the effect of the simplified assumptions in the maximum likelihood method for SFT prediction, but also offer a tool to predict the SFT when the maximum likelihood method is not applicable.

The paper is organized as follows. First, the uncertainties relevant to SFT prediction are explained, and the key assumptions involved in the maximum likelihood method are discussed. Then, Bayesian methods are suggested to learn the model and observational uncertainties, through which the probability distribution of the SFT can be determined. Thereafter, with an example, the proposed method is illustrated in detail. Finally, the proposed method is compared with the traditional method for SFT prediction, i.e. the deterministic method and the maximum likelihood method. The suggested method in this paper provides a versatile tool for SFT prediction considering both the slope-specific information and the information from other slopes.

2. Uncertainties involved in SFT prediction

Many phenomenological models have been suggested for SFT prediction, e.g. Saito (1969)'s method based on the plot of time vs. strain or displacement, the inverse velocity method (INVM) based on the plot of time vs. reciprocal of the velocity (R) (Fukuzono, 1985), the slope gradient method via the plot of velocity vs. the value of velocity multiplied by time (Mufundirwa et al., 2010), and the tangential angle method using the transformed plot of displacement vs. time (Xu et al., 2011). A comprehensive review and comparison of methods for SFT prediction have been conducted in Federico et al. (2015). Among these methods, the INVM has been widely used because it is easy to use and the interpretation of results is intuitive. In addition, as it can be expressed in a linear form, efficient machine learning algorithm can be developed based on this method. In this paper, the INVM is considered. The suggested method may also be potentially applicable to other methods such as Saito's method and the slope gradient method, in which the phenomenological models can also be expressed in a linear form. In the following, the INVM will be briefly described.

It is empirically shown that the plot of the reciprocal of the velocity vs. time often approaches linearity, especially in the critical stage of failure (Rose and Hungr, 2007). Assuming that the reciprocal of the velocity is a linear function of the time at the pre-failure stage of the slope, the relationship between the reciprocal of the velocity and the SFT can be expressed as follows (e.g. Fukuzono, 1985; Voight, 1988; Rose and Hungr, 2007; Carlà et al., 2017).

$$R = A(t_c - t) \tag{1}$$

where R and t represent the reciprocal of the velocity and the time, respectively; and A and t_c are the parameters to be calibrated.

The pre-failure stage is often defined as the stage after the point of onset of acceleration (OOA) (e.g. Dick et al., 2014; Carlà et al., 2017). Fig. 1 illustrates how to determine the SFT by the INVM. Comparing Eq. (1) with Fig. 1, it is shown that t_c is indeed the intercept of the fitted R-t linear relationship with the time axis. Assuming that the velocity of the slope movement is infinite when the slope failure occurs, t_c can be interpreted as the SFT predicted from the INVM. To derive an efficient algorithm for model uncertainty characterization, Eq. (1) can be rearranged to a linear relationship of unknown parameters as follows:

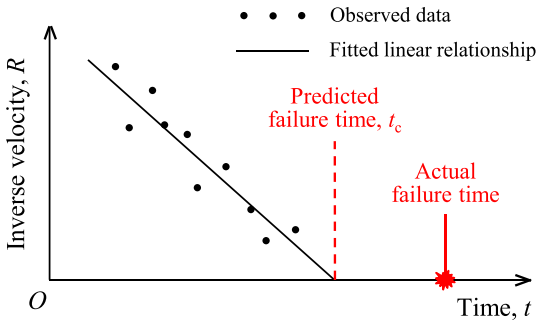

Fig. 1. Schematic of the observational and model uncertainties in the SFT prediction.

$$t = t_c - BR \tag{2}$$

where B is a coefficient associated with the slope of the R-t curve.

As shown in Fig. 1, the scattered observational data points around the straight line represent the observational uncertainty. The observational uncertainty can be modeled through a normal random variable ε_o with the mean and standard deviation (SD) of 0 and σ_o, respectively:

$$t = t_c - BR + \varepsilon_o \tag{3}$$

Due to the simplified modeling assumptions, the intercept of the R-t curve may not be exactly the actual SFT, as shown in Fig. 1. The disparity between the calculated SFT and the actual one is called the model uncertainty. Let t_a represent the actual SFT. In order to analyze the effect of model uncertainty, the t_a related to the calculated SFT t_c is as follows (Zhang et al., 2020a):

$$t_a = t_c + \varepsilon_m \tag{4}$$

where ε_m is a normal random variable with the mean and SD of μ_m and σ_m, respectively.

To realistically predict the SFT, both the model and observational uncertainties should be taken into consideration. The maximum likelihood method suggested by Zhang et al. (2020a) was used to calibrate the above uncertainties. In the following text, the maximum likelihood method will be briefly reviewed, and the key assumptions involved will be discussed.

3. Review of maximum likelihood method

In Zhang et al. (2020a), the observational uncertainty modeled based on Eq. (1) can be written as follows:

$$R = A(t_c - t) + \varepsilon_{oML} \tag{5}$$

where ε_{oML} is a normal random variable.

The mean of ε_{oML} is 0 and its SD is σ_{oML}. For ease of illustration, let $\boldsymbol{\theta} = \{A, t_c, \sigma_{oML}\}$. Suppose n data points are obtained to calibrate Eq. (5). Assuming that the observations are statistically independent when the value of $\boldsymbol{\theta}$ is known, the likelihood function of $\boldsymbol{\theta}$ is expressed as

$$l(\boldsymbol{\theta}) = \prod_{i=1}^{n} \phi\left(\frac{\rho_i - A(t_c - t_i)}{\sigma_{oML}}\right) \tag{6}$$

where $\phi(\cdot)$ represents the standard normal probability density function (PDF); and ρ_i represents the reciprocal of the observed velocity at time t_i, i.e. the observed value of R.

Let $\boldsymbol{\theta}^*$ represent the maximum value of $\boldsymbol{\theta}$. According to the principle of maximum likelihood, when the number of observations is large, $\boldsymbol{\theta}$ can be approximated as a multivariate normal vector with a mean $\boldsymbol{\mu}_{\boldsymbol{\theta}} = \boldsymbol{\theta}^*$ and a covariance matrix $\boldsymbol{C}_{\boldsymbol{\theta}}$, where $\boldsymbol{C}_{\boldsymbol{\theta}}$ is associated with the Hessen matrix of the logarithm of the likelihood function of $\boldsymbol{\theta}$ at point $\boldsymbol{\theta}^*$. Note that t_c is one of the elements of $\boldsymbol{\theta}$. Provided that the PDF of $\boldsymbol{\theta}$ is obtained, the PDF of t_c, which is the marginal PDF of $\boldsymbol{\theta}$, can also be obtained.

In Eq. (4), the mean and the SD of ε_m, i.e. μ_m and σ_m, characterize the model uncertainty. Let $\boldsymbol{\gamma} = \{\mu_m, \sigma_m\}$. Suppose r slopes are collected to calibrate the model uncertainty. Let μ_{cj} and σ_{ci} represent the mean and the SD of t_c of the jth slope. Note that the values of μ_{cj} and σ_{ci} are calculated by means of the maximum likelihood procedure introduced previously. Let δ_j represent the observed SFT of the jth slope. Let $\boldsymbol{\delta} = \{\delta_1, \delta_2, \ldots, \delta_r\}$. Supposing the SFTs of all the slopes are statistically independent, the likelihood function of μ_m and σ_m is written as follows:

$$l(\boldsymbol{\gamma} \mid \boldsymbol{\delta}) = \prod_{j=1}^{r} \phi\left(\frac{\delta_j - \mu_{cj} - \mu_m}{\sqrt{\sigma_{cj}^2 + \sigma_m^2}}\right) \tag{7}$$

According to the principle of maximum likelihood, when the number of slopes is large, the optimal values of μ_m and σ_m can be derived by maximizing Eq. (7).

After the model and observational uncertainties are characterized, the mean and the SD of the SFT can be obtained. Assuming the SFT follows the normal distribution, the cumulative distribution function (CDF) of SFT is computed as follows (Zhang et al., 2020a):

$$P(t_a < t) = \Phi\left(\frac{t - \mu_c - \mu_m}{\sqrt{\sigma_c^2 + \sigma_m^2}}\right) \tag{8}$$

where $\Phi(\cdot)$ represents the standard normal CDF.

As mentioned above, the maximum likelihood method is founded on the following assumptions:

(1) Both the model and the observational uncertainties are estimated by means of the maximum likelihood procedure, in which it is assumed that the amount of observed data is far larger than the number of model parameters.
(2) When estimating the observational uncertainty, the distribution of t_c is assumed to be normal.
(3) When estimating the model uncertainty, only the best estimate values of the model uncertainty parameters are figured out. The uncertainties associated with μ_m and σ_m are not considered.
(4) The distribution of SFT is assumed to be normal.

Therefore, although the above maximum likelihood method suggested in Zhang et al. (2020a) is easy to use, it is necessary to develop methods which can bypass the above assumptions. The availability of such methods will not only provide a chance to examine the effect of the simplified assumptions in the maximum likelihood method for SFT prediction, but also provide a tool to predict the SFT while the maximum likelihood method is not applicable. In the following, we will introduce such a method through BML. As will be shown, the proposed method in the paper requires less assumptions and can provide predictions which accord better with the observed SFT.

4. Machine learning of model uncertainty

4.1. Bayesian network

As the model uncertainty refers to the disparity between model predictions and observed SFT, it can be studied through a systematic analogy between the predicted SFT and the observed one for large amounts of slopes, which involves large amounts of uncertain variables. As the Bayesian network is capable of modeling the complex dependence relationships among a large set of uncertain variables (e.g. Aguilera et al., 2011; Bartlett and Cussens, 2017), the machine learning method proposed in this study will be developed through a Bayesian network.

Like the maximum likelihood method, suppose that r slopes are collected to learn the model uncertainty. Fig. 2 depicts the structure of the Bayesian network. In this figure, B_j, t_{cj} and σ_{oj} denote the unknown parameters when analyzing the jth slope with the INVM as given by Eq. (3). For the jth slope ($j = 1, 2, \ldots, r$), suppose there are n_j data points for calibration of the INVM. Let t_{ji} ($i = 1, 2, \ldots, n_j$) denote the ith observation time corresponding to the observed R_{ji} of the jth slope. According to Eq. (3) and the normal assumption on ε_o, the conditional PDF of t_{ji} given B_j, t_{cj} and σ_{oj}^2 can be written as follows:

$$f\left(t_{ji} \,\middle|\, B_j, t_{cj}, \sigma_{oj}^2\right) = \phi\left(\frac{t_{ji} - (t_{cj} - B_j R_{ji})}{\sigma_{oj}}\right) \tag{9}$$

Let t_{aj} represent the actual SFT of the jth slope. According to Eq. (4), when the values of μ_m, σ_m^2 and t_{cj} are known, the conditional PDF of t_{aj} is as follows:

$$f\left(t_{aj} \,\middle|\, \mu_m, \sigma_m^2, t_{cj}\right) = \phi\left(\frac{t_{aj} - (t_{cj} + \mu_m)}{\sigma_m}\right) \tag{10}$$

In a Bayesian network, the nodes at the head and the tail of an arrow are called the child node and the parent node, respectively. The nodes without parents are called the root nodes. In Fig. 2, the

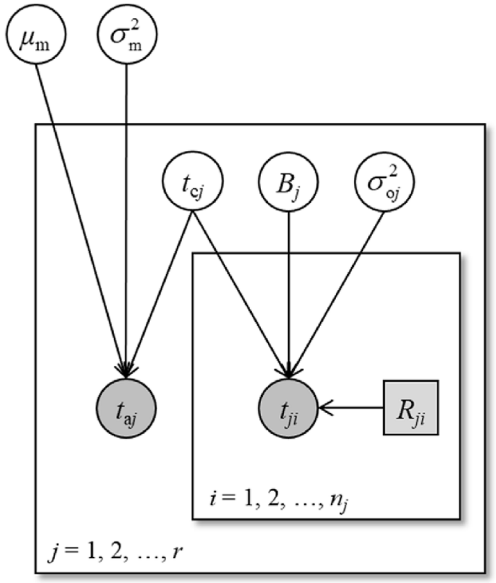

Fig. 2. Bayesian network for model uncertainty calibration.

root nodes include μ_m, σ_m^2, B_j, t_{cj} and σ_{oj}^2. To perform the Bayesian learning, the prior PDFs of the root nodes should be specified.

4.2. Prior PDF for the Bayesian network

In principle, the prior PDF of the random variables of the root nodes should be settled according to the prior knowledge so that the prior PDF of all variables can be obtained. When the prior knowledge about the root nodes is lacking, non-informative priors can be used (Del Castillo, 2007). On the other hand, when conjugate priors are adopted, the involved computational work can be significantly simplified (Del Castillo, 2007). Due to the above considerations, conjugate prior distributions are adopted for the root nodes in this study, which facilitate the development of an efficient Gibbs sampling algorithm to learn the database of SFT. Based on the Bayesian network and the normality of Eqs. (9) and (10), the conjugate prior PDFs for μ_m, B_j, and t_{cj} are normal (e.g. Gelman et al., 2013):

$$f(\mu_m) = \phi\left(\frac{\mu_m - \mu_\mu}{\sigma_\mu}\right) \tag{11}$$

$$f(B_j) = \phi\left(\frac{B_j - \mu_{Bj}}{\sigma_{Bj}}\right) \tag{12}$$

$$f(t_{cj}) = \phi\left(\frac{t_{cj} - \mu_{tcj}}{\sigma_{tcj}}\right) \tag{13}$$

where μ_μ and σ_μ are the mean and the SD of μ_m, respectively; μ_{Bj} and σ_{Bj} are the mean and the SD of B_j, respectively; and μ_{tcj} and σ_{tcj} are the mean and the SD of t_{cj}, respectively. When μ_μ, μ_{Bj} and μ_{tcj} take finite numbers and σ_μ, σ_{Bj} and σ_{tcj} take large numbers, i.e. $\sigma_\mu \to +\infty$, $\sigma_{Bj} \to +\infty$ and $\sigma_{tcj} \to +\infty$, the above prior PDFs of μ_m, B_j and t_{cj} may be considered to be non-informative.

Based on the Bayesian network and the normality of Eqs. (9) and (10), the conjugate prior PDFs for σ_m^2 and σ_{oj}^2 are scaled inverse chi-squared distributions (e.g. Del Castillo, 2007; Gelman et al., 2013):

$$f\left(\sigma_m^2\right) = \text{Scale} - \text{Inv} - \chi^2\left(\nu_m, s_m^2\right) \tag{14}$$

$$f\left(\sigma_{oj}^2\right) = \text{Scale} - \text{Inv} - \chi^2\left(\nu_{oj}, s_{oj}^2\right) \tag{15}$$

where ν_m and ν_{oj} represent the degrees of freedom; s_m^2 and s_{oj}^2 are the scale parameters; and Scale-Inv-$\chi^2(\nu, s^2)$ denotes the scaled inverse chi-squared distribution whose degree of freedom and scale parameter are ν and s^2, respectively. When the degrees of freedom are small and the scale parameters are finite, i.e. $\nu_m \to 0^+$, $\nu_{oj} \to 0^+$ and $0 < s_m^2 < +\infty$, $0 < s_{oj}^2 < +\infty$, the scaled inverse chi-squared distributions may be considered to be non-informative.

With the above two types of prior PDFs, one can obtain the closed-form solutions of conditional distributions of random variables involved in the Bayesian network, which are summarized in the Appendix. As the purpose of the Bayesian network is to learn the model uncertainty, which are characterized by μ_m and σ_m^2, the key step is how to determine the posterior distributions of μ_m and σ_m^2. In the following, an efficient algorithm is described to draw samples for the random variables in the Bayesian network, through which the samples of μ_m and σ_m^2 can also be obtained.

4.3. Algorithm for machine learning

Let $\boldsymbol{t}_j = \{t_{j1}, t_{j2}, \ldots, t_{jn_j}\}$ $(j = 1, 2, \ldots, r)$ denote the time when the velocities are measured for the jth slope. For ease of presentation, let $\boldsymbol{T} = \{\boldsymbol{t}_1, \boldsymbol{t}_2, \ldots, \boldsymbol{t}_r\}$, $\boldsymbol{t}_a = \{t_{a1}, t_{a2}, \ldots, t_{ar}\}$, $\boldsymbol{B} = \{B_1, B_2, \ldots, B_r\}$, $\boldsymbol{t}_c = \{t_{c1}, t_{c2}, \ldots, t_{cr}\}$, and $\boldsymbol{\Sigma}_o = \{\sigma_{o1}^2, \sigma_{o2}^2, \ldots, \sigma_{or}^2\}$. In this way, all the random variables can be expressed as $\{\mu_m, \sigma_m^2, \boldsymbol{B}, \boldsymbol{t}_c, \boldsymbol{\Sigma}_o, \boldsymbol{T}, \boldsymbol{t}_a\}$. According to the chain rule and the Markov property of the Bayesian network (e.g. Pearl, 1988; Neapolitan, 2004; Darwiche, 2009; Koller and Friedman, 2009), the joint PDF of the variables involved in the Bayesian network is expressed as follows:

$$f\left(\mu_m, \sigma_m^2, \boldsymbol{B}, \boldsymbol{t}_c, \boldsymbol{\Sigma}_o, \boldsymbol{T}, \boldsymbol{t}_a\right) = f(\mu_m) f\left(\sigma_m^2\right)$$

$$\cdot \prod_{j=1}^{r}\left[f(B_j) f(t_{cj}) f\left(\sigma_{oj}^2\right) f\left(t_{aj} \middle| \mu_m, \sigma_m^2, t_{cj}\right) \prod_{i=1}^{n_j} f\left(t_{ji} \middle| B_j, t_{cj}, \sigma_{oj}^2\right)\right] \tag{16}$$

where $f(\mu_m)$, $f(\sigma_m^2)$, $f(B_j)$, $f(t_{cj})$ and $f(\sigma_{oj}^2)$ represent the prior PDFs of μ_m, σ_m^2, B_j, t_{cj} and σ_{oj}^2, respectively; $f(t_{aj} \mid \mu_m, \sigma_m^2, t_{cj})$ represents the PDF of t_{aj} conditional on μ_m and σ_m^2; and $f(t_{ji} \mid B_j, t_{cj}, \sigma_{oj}^2)$ represents the PDF of t_{ji} conditional on B_j, t_{cj} and σ_{oj}^2.

Let d_{ji} represent the observational value of t_{ji} $(j = 1, 2, \ldots, r; i = 1, 2, \ldots, n_j)$, and let $\boldsymbol{d}_j = \{d_{j1}, d_{j2}, \ldots, d_{jn_j}\}$ denote the data of the observation time at the jth slope. Let τ_{aj} denote the observational value of t_{aj} $(j = 1, 2, \ldots, r)$, i.e. the observed SFT of the jth slope. For ease of presentation, let $\boldsymbol{\tau}_a = \{\tau_{a1}, \tau_{a2}, \ldots, \tau_{ar}\}$, and $\boldsymbol{D} = \{\boldsymbol{d}_1, \boldsymbol{d}_2, \ldots, \boldsymbol{d}_r\}$. When some of the variables are observed, the distribution of other variables in the network can be updated (Pearl, 1988). Based on Eq. (16), the conditional distribution of $\{\mu_m, \sigma_m^2, \boldsymbol{B}, \boldsymbol{t}_c, \boldsymbol{\Sigma}_o\}$, given $\boldsymbol{T} = \boldsymbol{D}$ and $\boldsymbol{t}_a = \boldsymbol{\tau}_a$, is expressed as follows:

$$f\left(\mu_m, \sigma_m^2, \boldsymbol{B}, \boldsymbol{t}_c, \boldsymbol{\Sigma}_o \middle| \boldsymbol{T} = \boldsymbol{D}, \boldsymbol{t}_a = \boldsymbol{\tau}_a\right) = \frac{f\left(\mu_m, \sigma_m^2, \boldsymbol{B}, \boldsymbol{t}_c, \boldsymbol{\Sigma}_o, \boldsymbol{T} = \boldsymbol{D}, \boldsymbol{t}_a = \boldsymbol{\tau}_a\right)}{f\left(\boldsymbol{T} = \boldsymbol{D}, \boldsymbol{t}_a = \boldsymbol{\tau}_a\right)} \tag{17}$$

Eq. (17) sheds light on the posterior PDF of the variables of interest in the Bayesian network, including μ_m and σ_m, i.e. the mean and the SD of model uncertainty. Note the dimension of the random variables to be updated in Eq. (17) is $3r + 2$. If $r = 50$, the number of variables to be updated is 152. The number of variables to be updated will increase as the number of slopes further increases. Such a Bayesian problem with high dimensions is generally difficult to solve. In the suggested method, the Markov-chain Monte Carlo simulation (MCMCS) is conducted to readily study the properties of the posterior distributions. Particularly, the Gibbs sampler (e.g. Geman and Geman, 1984; Ching and Phoon, 2019; Zhang et al., 2020b) is adopted to construct the Markov chains. In a Gibbs sampler, the samples of a variable are generated from the PDF of the random variable conditional on other variables. Let $\{\mu_m^{(0)}, \sigma_m^{2(0)}, \boldsymbol{B}^{(0)}, \boldsymbol{t}_c^{(0)}, \boldsymbol{\Sigma}_o^{(0)}\}$ be an arbitrarily chosen initial point of the Markov chain. The following procedure shows how to conduct the Gibbs sampler. Here k serves as the loop counter, which starts from 1:

(1) Draw a sample $\mu_m^{(k)}$ from its conditional distribution given the current sample points of other variables on the Markov chain, i.e. $f(\mu_m^{(k)} \mid \sigma_m^{2(k-1)}, \boldsymbol{B}^{(k-1)}, \boldsymbol{t}_c^{(k-1)}, \boldsymbol{\Sigma}_o^{(k-1)}, \boldsymbol{T} = \boldsymbol{D}, \boldsymbol{t}_a = \boldsymbol{\tau}_a)$;

(2) Draw a sample $\sigma_m^{2(k)}$ from its conditional distribution given the current sample points of other variables on the Markov chain, i.e. $f(\sigma_m^{2(k)} | \mu_m^{(k)}, \boldsymbol{B}^{(k\text{-}1)}, \boldsymbol{t}_c^{(k-1)}, \boldsymbol{\Sigma}_o^{(k-1)}, \boldsymbol{T} = \boldsymbol{D}, \boldsymbol{t}_a = \tau_a)$;

(3) Loop through $j = 1, 2, \ldots, r$ as follows:

(i) Draw a sample $B_j^{(k)}$ from $f(B_j^{(k)} | \mu_m^{(k)}, \sigma_m^{2(k)}, \boldsymbol{B}_{*j}^{(k\text{-}1)}, \boldsymbol{t}_c^{(k-1)}, \boldsymbol{\Sigma}_o^{(k-1)}, \boldsymbol{T} = \boldsymbol{D}, \boldsymbol{t}_a = \tau_a)$, in which $\boldsymbol{B}_{*j}^{(k\text{-}1)} = \{B_1^{(k)}, B_2^{(k)}, \ldots, B_{j-1}^{(k)}, B_{j+1}^{(k-1)}, \ldots, B_r^{(k-1)}\}$ represents the vector consisting of all the elements of $\boldsymbol{B}$ except for B_j at their current values;

(ii) Draw a sample $t_{cj}^{(k)}$ from i.e. $f(t_{cj}^{(k)} | \mu_m^{(k)}, \sigma_m^{2(k)}, \boldsymbol{B}_j^{(k)}, \boldsymbol{t}_{c*j}^{(k-1)}, \boldsymbol{\Sigma}_o^{(k-1)}, \boldsymbol{T} = \boldsymbol{D}, \boldsymbol{t}_a = \tau_a)$, in which $\boldsymbol{B}_j^{(k)} = \{B_1^{(k)}, B_2^{(k)}, \ldots, B_j^{(k)}, B_{j+1}^{(k-1)}, \ldots, B_r^{(k-1)}\}$ represents the vector of $\boldsymbol{B}$ with the first j elements updated, and $\boldsymbol{t}_{c*j}^{(k-1)} = \{t_{c1}^{(k)}, t_{c2}^{(k)}, \ldots, t_{c(j-1)}^{(k)}, t_{c(j+1)}^{(k)}, \ldots, t_{cr}^{(k)}\}$ represents the vector consisting of all the elements of $\boldsymbol{t}_c$ except for t_{cj} at their current values;

(iii) Draw the sample $\sigma_{oj}^{2(k)}$ from $f(\sigma_{oj}^{2(k)} | \mu_m^{(k)}, \sigma_m^{2(k)}, \boldsymbol{B}_j^{(k)}, \boldsymbol{t}_{cj}^{(k)}, \boldsymbol{\Sigma}_{o*j}^{(k\text{-}1)}, \boldsymbol{T} = \boldsymbol{D}, t_a = \tau_a)$, in which $\boldsymbol{t}_{cj}^{(k)} = \{t_{c1}^{(k)}, t_{c2}^{(k)}, \ldots, t_{cj}^{(k)}, t_{c(j+1)}^{(k-1)}, \ldots, t_{cr}^{(k-1)}\}$ represents the vector of $\boldsymbol{t}_c$ with the first j elements updated, and $\boldsymbol{\Sigma}_{o*j}^{(k-1)} = \{\sigma_{o1}^{2(k)}, \sigma_{o2}^{2(k)}, \ldots, \sigma_{o(j-1)}^{2(k)}, \sigma_{o(j+1)}^{2(k-1)}, \ldots, \sigma_{or}^{2(k-1)}\}$ represents the vector consisting of all the elements of $\boldsymbol{\Sigma}_o$ except for σ_{oj}^2 at their current values.

(4) Let $k = k + 1$. Then return to step (1) unless the sufficient samples are collected.

The implementation of the above algorithm requires the analytical expressions of a series of conditional PDFs as mentioned in Steps (1)–(3), which have been summarized in the Appendix. With the above procedure, the samples of different variables in the Bayesian network can be obtained, including those of μ_m and σ_m^2. When drawing samples with MCMCS, one should judge if the Markov chain converges to the equilibrium state. Assessing the convergence of a Markov chain is one of the most challenging problems in MCMCS, and many methods have been suggested to analyze the convergence of the Markov chain (e.g. Cowles and Carlin, 1996; Brooks and Roberts, 1998; Kass et al., 1998; Sinharay, 2003). However, none of these methods can ensure the convergence of a Markov chain within a finite number of samples. A review and comparison of different techniques for convergence checking can be found in Cowles and Carlin (1996). In practice, the convergence is often checked empirically by observing if the Markov chain generates samples with stable statistics such as the median and the correlation (e.g. Geman and Geman, 1984). If stable statistics can be obtained, the number of samples in the Markov chain can be considered sufficient (e.g. Gelman et al., 2013; Ching and Phoon, 2019). In this paper, such an empirical method is adopted. For ease of illustration, the dataset of these samples of μ_m and σ_m^2 is denoted as S-1.

Note in addition to the MCMCS, approaches such as the Bayesian updating with structural reliability method suggested in Straub and Papaioannou (2015) and adaptive Bayesian updating with subset simulation (e.g. Giovanis et al., 2017; Jiang et al., 2020) are also increasingly used for solving high dimensional Bayesian problems. Although MCMCS is used in this study, other approaches may also be potentially useful to solve the model uncertainty characterization problem as described in this study.

5. Machine learning of the observational uncertainty

In the preceding section, the model uncertainty in Eq. (4) is calibrated through the Bayesian network. Its observational uncertainty should also be calibrated when predicting the failure time of a new slope. Let t_{cN}, B_N and σ_{oN}^2 denote t_c, B and σ_o^2 of the new slope, respectively. When applying Eq. (3) to a new slope, there are three uncertain variables, i.e. t_{cN}, B_N, and σ_{oN}^2. Let $\boldsymbol{\beta} = (t_{cN}, -B_N)^T$. Suppose at time t_{Ni}, the observed reciprocal of velocity is R_{Ni}, and there are n_N pairs of t_{Ni} and R_{Ni}. For ease of presentation, let $\boldsymbol{t}_N = (t_{N1}, t_{N2}, \ldots, t_{NnN})^T$ and $\boldsymbol{R}_N = \{R_{N1}, R_{N2}, \ldots, R_{NnN}\}^T$. As the values of $\boldsymbol{R}_N$, $\boldsymbol{\beta}$ and σ_{oN}^2 are given, the likelihood function is expressed as follows:

$$L\left(\mathbf{t}_N \middle| \boldsymbol{R}_N, \boldsymbol{\beta}, \sigma_{oN}^2\right) = \prod_{i=1}^{n_N} \phi\left(\frac{t_{Ni} - t_{cN} + B_N R_{Ni}}{\sigma_{oN}}\right) \tag{18}$$

The conjugate prior distributions for σ_{oN}^2 and $\boldsymbol{\beta}$ are as follows (e.g. Walter and Augustin, 2010):

$$f\left(\sigma_{oN}^2\right) = \text{Scale} - \text{Inv} - \chi^2\left(\nu_0, s_0^2\right) \tag{19}$$

$$f\left(\boldsymbol{\beta} \middle| \sigma_{oN}^2\right) = \text{MVNormal}\left(\boldsymbol{\mu}_0, \sigma_{oN}^2 \boldsymbol{\Lambda}_0^{-1}\right) \tag{20}$$

where ν_0 represents the degree of freedom; s_0^2 represents the scale parameter; $\boldsymbol{\mu}_0$ represent the mean; and $\sigma_{oN}^2 \boldsymbol{\Lambda}_0^{-1}$ represents the covariance matrix. When $\nu_0 \to 0^+$, $0 < s_0^2 < +\infty$ and $\boldsymbol{\Lambda}_0 \to \mathbf{0}$, the above prior PDFs of σ_{oN}^2 and $\boldsymbol{\beta}$ may be considered to be non-informative (Gelman et al., 2013). When the above conjugate priors and the likelihood function are given, the posterior distribution of σ_{oN}^2 and $\boldsymbol{\beta}$ are as follows (e.g. Korner-Nievergelt et al., 2015):

$$f\left(\sigma_{oN}^2 \middle| \boldsymbol{t}_N, \boldsymbol{R}_N\right) = \text{Scale} - \text{Inv} - \chi^2\left(\nu_N, s_N^2\right) \tag{21}$$

$$f\left(\boldsymbol{\beta} \middle| \sigma_{oN}^2, \boldsymbol{t}_N, \boldsymbol{R}_N\right) = \text{MVNormal}\left(\boldsymbol{\mu}_N, \sigma_{oN}^2 \boldsymbol{\Lambda}_N^{-1}\right) \tag{22}$$

where $\boldsymbol{\mu}_N$, $\boldsymbol{\Lambda}_N$, ν_N, and s_N^2 are the parameters of the updated distribution. These parameters can be calculated analytically as follows:

$$\boldsymbol{\mu}_N = \left(\boldsymbol{X}^T\boldsymbol{X} + \boldsymbol{\Lambda}_0\right)^{-1}\left(\boldsymbol{\Lambda}_0\boldsymbol{\mu}_0 + \boldsymbol{X}^T\boldsymbol{X}\widehat{\boldsymbol{\beta}}\right) \tag{23}$$

$$\boldsymbol{\Lambda}_N = \boldsymbol{X}^T\boldsymbol{X} + \boldsymbol{\Lambda}_0 \tag{24}$$

$$\nu_N = \nu_0 + (n_N - p) \tag{25}$$

$$s_N^2 = \frac{\nu_0}{\nu_N} s_0^2 + \frac{1}{\nu_N}\left(\boldsymbol{t}_N^T\boldsymbol{t}_N + \boldsymbol{\mu}_0^T\boldsymbol{\Lambda}_0\boldsymbol{\mu}_0 - \boldsymbol{\mu}_N^T\boldsymbol{\Lambda}_N\boldsymbol{\mu}_N\right) \tag{26}$$

where $\boldsymbol{X} = (\mathbf{1}_{nN\times 1}\ \boldsymbol{R}_N)$ contains all the observed data; $\mathbf{1}_{nN\times 1} = (1,1, \ldots,1)^T$ is the vector consists of n_N 1's; $\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{t}_N$; and p is the dimension of $\boldsymbol{\beta}$ which equals 2 in this study. In the literature, $\boldsymbol{X}$ is usually called the design matrix (e.g. Castillo et al., 2015). Based on Eqs. (23)-(26), the posterior samples of $\boldsymbol{\beta}$ can be obtained, including the samples of t_{cN}. Note that the large sample assumption and the

normal assumption about t_{cN} are both not required in the suggested method in this study. The above method is also called Bayesian linear regression in the literature (Smith, 1973; Walter and Augustin, 2010). Starting from $k = 1$, the samples of t_{cN} can be drawn based on Eqs. (21) and (22) using the following procedure, i.e. the procedure of the Gibbs sampler.

(1) Draw the sample $\sigma_{oN}^{2(k)}$ from the distribution as given by Eq. (21);
(2) Draw the sample $\beta^{(k)}$ from the distribution as given by Eq. (22) conditional on $\sigma_{oN}^{2(k)}$. Note that t_{cN} is the first element of β. After the sample $\beta^{(k)}$ is drawn, the kth sample of t_{cN}, i.e. $t_{cN}^{(k)}$ can be obtained;
(3) Let $k = k + 1$. Then return to step (1) unless the sufficient samples are collected.

For ease of illustration, the dataset of the samples of t_{cN} is called S-2 in this paper.

6. Failure time prediction of the new slope

Let t_{aN} represent the actual SFT of the new slope to be predicted. Following Eq. (4), t_{aN} is assumed to follow the normal distribution given the values of t_{cN}, μ_m, and σ_m^2, and its mean is $t_{cN} + \mu_m$ and its SD is σ_m. Hence, its PDF conditional on t_{cN}, μ_m, and σ_m^2 can be written as follows:

$$f\left(t_{aN} \mid \mu_m, \sigma_m^2, t_{cN}\right) = \phi\left(\frac{t_{aN} - (t_{cN} + \mu_m)}{\sigma_m}\right) \tag{27}$$

As mentioned previously, the exact values of t_{cN}, μ_m, and σ_m^2 are unknown and are modeled as random variables. Previously, a Bayesian network has been suggested to learn μ_m and σ_m^2, and the posterior samples of these two variables are stored in S-1. The Bayesian linear regression method has been suggested to learn t_{cN}, and its posterior samples have been stored in S-2. Starting with $l = 1$, the following procedure shows how to draw the samples of t_{aN} based on S-1 and S-2:

(1) Draw samples of μ_m and σ_m^2 at random from S-1, which are denoted as $\mu_m^{(l)}$ and $\sigma_m^{2(l)}$, respectively;
(2) Draw a sample of t_{cN} at random from S-2, which is denoted as $t_{cN}^{(l)}$;
(3) Draw the lth sample of t_{aN}, which is denoted as $t_{aN}^{(l)}$ here, based on Eq. (27), i.e. $f(t_{aN} \mid \mu_m^{(l)}, \sigma_m^{2(l)}, t_{cN}^{(l)})$;
(4) Let $l = l + 1$. Then return to step (1) unless the sufficient samples are collected.

7. Suggested procedure for SFT prediction

To facilitate implementation, the procedure for predicting SFT with the method suggested in this paper is summarized as follows:

(1) Learn the model uncertainty through the Bayesian network based on which the samples of μ_m and σ_m^2 can be obtained. In this step, the knowledge from slope failure data is learned.
(2) Learn the observational uncertainty associated with t_{cN} through Bayesian regression analysis of the monitoring data at the new slope to be analyzed, for which the samples of t_{cN} can be obtained. In this step, the knowledge from the monitoring data at the new slope is learned.

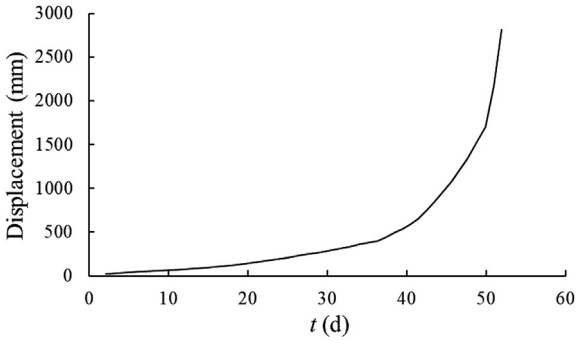

Fig. 3. Displacement data of the Abbotsford slope (Adapted from Hancox, 2008).

(3) Simulate the samples of the actual SFT with samples of μ_m, σ_m^2, and t_{cN} based on Eq. (27). With these samples, the histogram and empirical CDF of the SFT are obtained.

Notably, the calibration of the model and the observational uncertainty are separate. One might be interested if the two types of uncertainties can be calibrated simultaneously. Such an idea is also tested. However, it turns out that when the two types of uncertainties are calibrated together, the model is very difficult to converge. Hence, the two-stage calibration method is used in this paper.

8. An illustrative example of the abbotsford landslide

8.1. The Abbotsford landslide

The Abbotsford landslide is located in southwest Dunedin, New Zealand (Hancox, 2008). The slope lies on a spur that rises at an angle of 20°–25° for about 100 m and flattens toward the top. A sand quarry modifies the southern end of the slope, and a prehistoric slide lies to the north. The underlying rock of the landslide is comprised of mudstone, overlain by clayey to silty sand. Large cracks were found between Edwards and Mitchell streets to the west side of the slope in early July 1979. Fig. 3 shows the cumulative displacement across the crack in Mitchell street determined by the survey monitoring line from June 18 ($t = 0$) to August 8, 1979. In early July, the ground moved about 10 mm per day. Triggered by the rainfall, the ground movements accelerated to about 650 mm per day prior to the final movement. The compression rolls and cracks with a length of 30 m also developed in the prehistoric slide area in mid-July. The final movement of the slope lasted about 30 min from

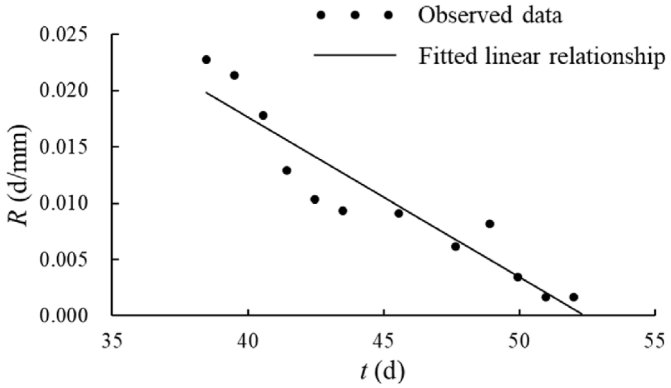

Fig. 4. Observed data of the reciprocal of the velocity of the Abbotsford slope after the OOA point.

Table 1
The landslide database compiled in this paper.

No.	Landslide	Material	Triggering factors	Actual SFT (d)	95% CI of SFT (d)		Source
					Maximum likelihood method	BML-based method	
1	Abbotsford	Rock	Rainfall and groundwater	51.88	[49.76, 53.85]	[48.40, 53.9]	Hancox (2008)
2	Agoyama	Rock	Human activities	55.06	[53.11, 56.15]	[53.26, 57.62]	Saito (1979)
3	Allori	Soil	Excavation	1706	[1535.08, 1865.33]	[1525.2, 1883.6]	D'Elia et al. (1998)
4	Asamushi	Rock	Weathering and thermal metamorphosis	6.93	[5.77, 8.33]	[5.88, 9.72]	Saito (1969)
5	Baige	Rock	Long-term slope deformation and more precipitation	2.74	[1.83, 4.34]	[2.01, 5.63]	Li et al. (2020)
6	Baishi	Rock	Rainfall and earthquake	200.98	[200.86, 203.72]	[200.59, 204.84]	Xu et al. (2011)
7	Bomba	Soil	Excavation	1036	[1035.24, 1044.8]	[1032.7, 1047.5]	D'Elia et al. (1998)
8	Ca'Lita (2004.11)	Rock and soil	Rainfall and weathering	140	[128.59, 148.31]	[119.69, 158.18]	Borgatti et al. (2006)
9	Ca'Lita (2005.04)	Rock and soil	Rainfall and weathering	147	[126.77, 160.62]	[126.31, 162.51]	Borgatti et al. (2006)
10	Cavallerizzo	Rock and soil	Rainfall and snowfall	6.12	[5.93, 8.87]	[5.34, 10.74]	Iovine et al. (2006)
11	Cowden	Soil	Marine erosion	163	[153.99, 179.91]	[143.26, 190.09]	Dixon et al. (2003)
12	Dangchuan#4	Loess	Irrigation	121.2	[122.06, 126.75]	[120.9, 125.98]	Xu et al. (2020)
13	Daye	Rock	Human activities and rainfall	1282	[1268.32, 1280.07]	[1267.1, 1282.07]	Xu et al. (2011)
14	Dosan	Rock	Erosion	3.12	[2.08, 4.61]	[1.92, 6.01]	Saito (1969)
15	Excavation C	Rock	–	207	[193.92, 207.05]	[188.42, 214.33]	Moretto et al. (2017)
16	Hanjiang	Rock and soil	Rainfall	184	[152.9, 215.6]	[148.78, 219.96]	Wang (2018)
17	Harmony 1C	Rock	Mining and blasting	41	[37.5, 42.22]	[37.09, 43.97]	Cahill and Lee (2006)
18	Huanglongxicun	Loess	Rainfall and human activities	5	[3.03, 5.87]	[2.98, 7.25]	Li et al. (2012)
19	Jimingsi	Rock	Human activities and rainfall	432	[429.32, 433.64]	[429.2, 435.04]	Xu et al. (2011)
20	Jinchuan	Rock	Excavation and blasting	814	[766.17, 859.9]	[762.3, 864.82]	Xu and Li (1986)
21	Jizukiyama	Rock and soil	Rainfall	11.72	[9.67, 12.02]	[9.55, 13.51]	Hayashi et al. (1988)
22	Kagemori	Rock	Rainfall	405	[397.45, 410]	[395.62, 412.46]	Segalini et al. (2018)
23	Kensal Green	Clay	Groundwater and excavation	4758	[4566.24, 4762.97]	[4549.8, 4789.1]	Skempton (1977)
24	La Saxe	Rock	Snow melt	21	[19.79, 22.33]	[19.83, 23.8]	Manconi and Giordan (2016)
25	Letlhakane	Rock	Mining activities	48	[46.79, 49.85]	[46.73, 51.34]	Kayesa (2006)
26	Longjing	Rock	Excavation	20.25	[19.47, 21.97]	[19.55, 23.29]	Fan et al. (2019)
27	Luscar	Rock	–	249	[247.65, 251.62]	[247.55, 253.11]	Cruden and Masoumzadeh (1987)
28	Maoxian	Rock	Rainfall	989	[972.44, 1005.13]	[967.3, 1013.5]	Intrieri et al. (2018)
29	Mt Beni	Rock	Rainstorm	248	[233.46, 253.03]	[233.02, 253.35]	Gigli et al. (2011)
30	Mt Owen	Soil	–	11.32	[10.7, 13.17]	[10.65, 14.7]	Harries et al. (2006)
31	Nevis Bluff	Rock	Excavation	286	[275.35, 298.49]	[274.82, 300.47]	Brown et al. (1980)
32	New Tredegar	Rock	Porewater	70	[68.38, 71.35]	[68.24, 72.79]	Carey et al. (2007)
33	Ohto	Rock	Rainstorm	86	[83.08, 85.77]	[82.84, 86.21]	Suwa et al. (2010)
34	Ooigawa	Soil	Erosion	9.17	[8.14, 10.7]	[8.16, 12.14]	Saito (1969)
35	Open cut mine in Africa	Rock	–	194.5	[174.97, 196.54]	[173.07, 198.62]	Harries et al. (2006)
36	Open pit mine in Chile	–	–	420	[394.65, 444.59]	[390.81, 448.59]	Xu et al. (2011)
37	Otomura	Rock	Rainfall	5.01	[3.96, 6.48]	[3.95, 8.15]	Fujisawa et al. (2010)
38	Preonzo	Rock	Rainfall	680	[674.88, 679.85]	[674.45, 680.34]	Loew et al. (2017)
39	Puigcercos	Rock	Progressive degradation of the stability conditions	2198	[2162.9, 2229.11]	[2142.5, 2250.5]	Royán et al. (2015)
40	Roesgrenda	Clay	Short-term rainfall	3.45	[2.52, 5.04]	[2.5, 6.54]	Okamoto et al. (2004)
41	Ruinon	Rock	Rainfall	193	[181.16, 204.2]	[180.36, 205.94]	Crosta and Agliardi (2002)
42	Selborne	Soil	Groundwater	602	[592.43, 602.84]	[589.33, 606.63]	Petley (2004)
43	Stromboli	Rock	Volcano eruption	30.1	[28.76, 33.15]	[28.8, 34.37]	Carlà et al. (2017)
44	Takabayama	Rock	Rainfall and snowfall	62.06	[61.65, 64.14]	[61.63, 65.56]	Saito (1979)
45	Tama	Rock	–	3.9	[2.92, 5.59]	[2.83, 7.03]	Saito (1979)
46	Tianhuangping	Rock	Excavation blasting and rainfall	39	[36.63, 38.97]	[36.3, 39.57]	Mei (2001)
47	Tom Price's North Deposit	Rock	Reduction in the thickness and quality of the covering	132	[129.9, 134.96]	[130.17, 135.93]	Venter et al. (2013)
48	Town of Peace River	Rock	Rainfall	176	[164.53, 183.6]	[165., 184.45]	Kim et al. (2010)
49	Tuveras	–	–	294	[292.04, 296.86]	[292.13, 298.05]	Rochet (1992)
50	Vajont	Rock	Groundwater	70	[69.12, 72.63]	[68.91, 73.63]	Sornette et al. (2004)
51	Volterra	Soil	Water accumulation	9.6	[8, 10.57]	[7.94, 12.12]	Carlà et al. (2017)
52	Welland	Clay	Water seepage	7.69	[6.57, 9.12]	[6.53, 10.54]	Kwan (1971)
53	West Angelas' Centre Pit North	Rock	Reduction in the thickness and quality of the covering	223	[219.98, 222.59]	[219.92, 223.37]	Venter et al. (2013)
54	Xintan	Rock	Rainfall and groundwater	2393	[2391.34, 2399.50]	[2382.8, 2407.3]	Xue et al. (2014)
55	Yunotani	Soil	Snow melt	7.74	[6.5, 9.11]	[6.46, 10.72]	Saito (1979)

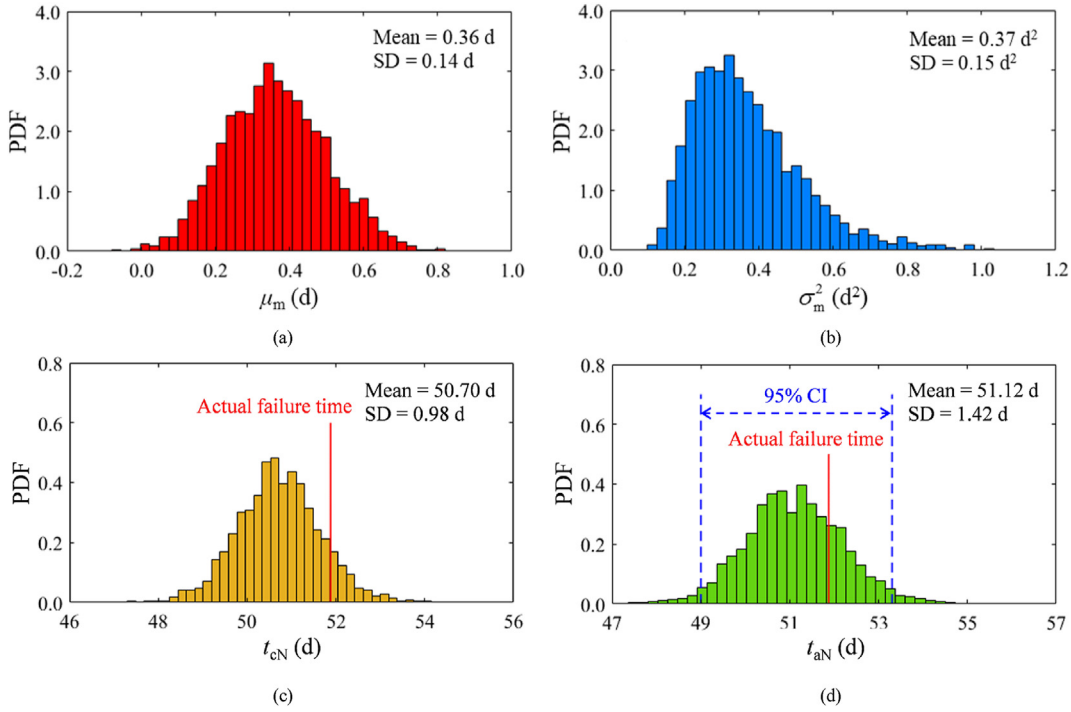

Fig. 5. Histograms of the samples of different variables in the SFT prediction of the Abbotsford slope: (a) μ_m; (b) σ_m^2; (c) t_{cN}; and (d) t_{aN}.

about 9 p.m. on August 8, 1979 ($t = 51.88$ d). The slope mass with a volume of 5 million m^3 slid southeast down about 50 m, which destroyed more than 20 houses and dammed Miller Creek. Following the criterion introduced in Dick et al. (2014) and Segalini et al. (2018), the OOA point is identified as $t = 39$ d. The observed data of the reciprocal of the velocity of this slope after the OOA point are shown in Fig. 4.

8.2. Landslide database

To characterize the model uncertainty of the INVM, the displacement data of other 54 landslides are collected. Together with the Abbotsford landslide, a landslide database with 55 slopes is formed. Table 1 summarizes the general information of the 55 landslides. As revealed from Table 1, both rock and soil landslides are contained in the database as both types of slopes can be analyzed through the linear INVM. The majority of the landslides are triggered by water-related factors such as rainfall, groundwater and snow melt, as well as production activities of human being.

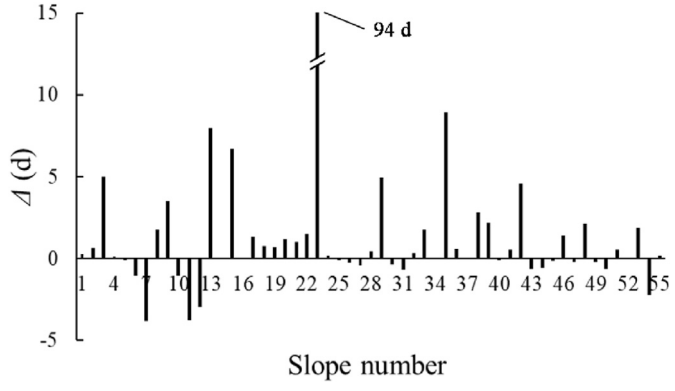

Fig. 6. Difference between the calculated SFT in the deterministic INVM and the actual one.

Next, the monitoring data of the 2nd-55th landslides in Table 1 will be used as the training dataset for model uncertainty characterization, through which the SFT of the Abbotsford slope will be analyzed.

8.3. Calibration of model uncertainty

First, the monitoring data of the 2nd-55th landslides in Table 1 are substituted into the Bayesian network, and the model uncertainty is learned through MCMCS. The length of the Markov chain is 10^5. It was found that the Markov chain of each variable can soon provide stable statistics with the number of samples increasing. To remove the influence of the initial point, the first 5×10^4 samples are discarded. Then the other 5×10^4 samples can be considered in the convergence stage and are collected as the samples of the posterior distributions. Fig. 5a and b shows the sample histograms of μ_m and σ_m^2, respectively. Computed with these samples, the means of μ_m and σ_m^2 are 0.36 d and 0.37 d^2, respectively; the SDs of μ_m and σ_m^2 are 0.14 d and 0.15 d^2, respectively. As is shown in the above analysis, the uncertainties associated with μ_m and σ_m^2 are not negligible.

8.4. Calibration of observational uncertainty

Substituting the data of Fig. 4 into Eqs. (23)-(26), one can then draw samples for σ_{oN}^2 and t_{cN} from the distribution defined by Eqs. (21) and (22). The histogram of t_{cN} is shown in Fig. 5c. The mean and the SD of the calculated t_{cN} are 50.7 d and 0.98 d, respectively. Note that the value of σ_m is similar to that of the SD of t_{cN}, indicating that in this example, the observational and the model uncertainties are of similar magnitudes, and hence both types of uncertainties are important for SFT prediction.

In the maximum likelihood method, t_{cN} is assumed to follow the normal distribution. After obtaining the samples from the Bayesian method, it is interesting to assess if the normal assumption about t_{cN} is appropriate. Here, the Jarque-Bera test is conducted to verify

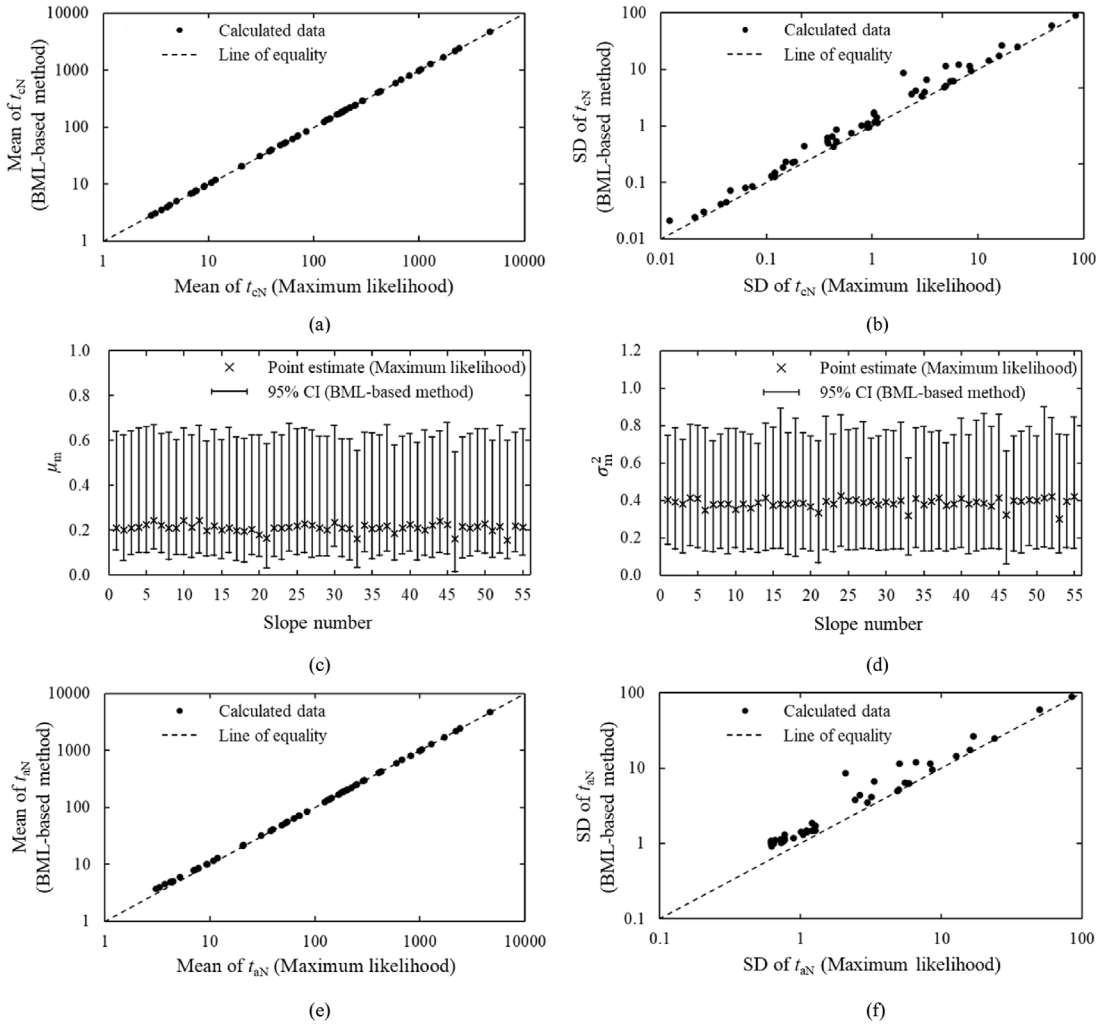

Fig. 7. Comparison of results from the maximum likelihood method and BML-based method: (a) Mean of t_{cN}; (b) SD of t_{cN}; (c) μ_m; (d) σ_m^2; (e) Mean of t_{aN}; and (f) SD of t_{aN}.

the normality of t_{cN}, which serves as a normality goodness-of-fit test suitable for large datasets (Jarque and Bera, 1987). The critical value for the Jarque-Bera test is 5.97 at the 0.05 significance level. In this example, the Jarque-Bera statistic is 47, which is far larger than the critical value. Therefore, the null hypothesis that t_{cN} is normal should be rejected when the significance level is 0.05.

8.5. SFT prediction

The histogram of t_{aN} is shown in Fig. 5d. For these samples, the Jarque-Bera statistic is equal to 6.85, slightly larger than the critical value when the significance level is 0.05. Thus, the null hypothesis of normality of t_{aN} is rejected. The mean and the SD of t_{aN} are 51.12 d and 1.42 d, respectively. Based on the samples of t_{aN}, the 95% CI of the SFT is [48.4 d, 53.9 d]. The observed SFT of the Abbotsford slope, 51.88 d, is within the 95% interval of t_{aN} derived from the suggested procedure.

For comparison, the predicted SFT is 51.63 d when the traditional deterministic INVM is used. When adopting the maximum likelihood method, the 95% CI of the SFT is [49.77 d, 53.88 d]. The 95% CI predicted via the maximum likelihood method is narrower than that predicted using the BML-based method in this study, indicating that the uncertainty relevant to the SFT is underestimated in the maximum likelihood method. Nevertheless, the SFT is located in the CI determined based on the BML-based method and the maximum likelihood method. In the following, a systematic comparison of the three methods will be conducted.

9. Comparison of methods for SFT prediction

To show the advantage of the suggested method, the SFT of each slope in Table 1 was analyzed using the traditional deterministic method, the maximum likelihood method, and the BML-based

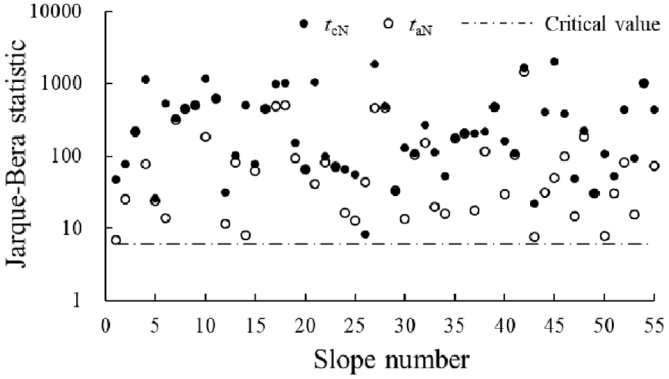

Fig. 8. Jarque-Bera statistics for t_{cN} and t_{aN}.

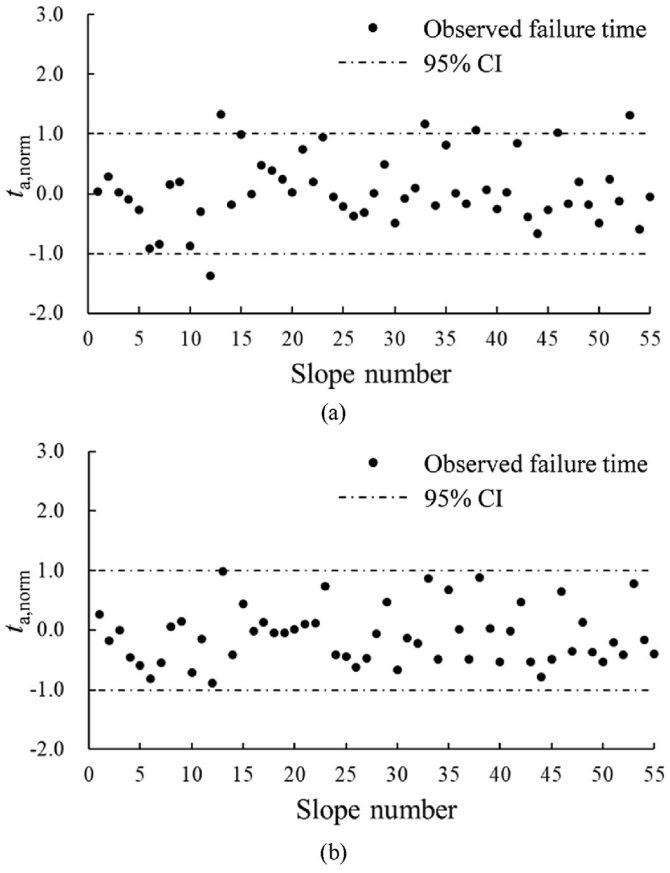

Fig. 9. Comparison of the observed SFT and the 95% CI of the SFT: (a) Maximum likelihood method; and (b) BML-based method.

method. When analyzing the SFT in Table 1 with the maximum likelihood method or the BML-based method, the rest slopes were used as the training dataset to learn the model uncertainty.

First, the traditional deterministic method was carried out to forecast the SFT. Let $t_{cN,d}$ denote the SFT calculated by the traditional deterministic method. Let $\Delta = t_a - t_{cN,d}$ denote the difference between the calculated SFT and the actual one. By definition, if $\Delta > 0$, the calculated SFT is earlier than the actual one, resulting in false warnings. If $\Delta < 0$, the calculated SFT is later than the actual one, resulting in missing alarms (i.e. warnings that are too late). Fig. 6 shows the values of Δ for each slope in Table 1. As can be seen from this figure, there are 33 out of 55 slopes with Δ being positive, and 22 out of 55 cases with Δ being negative. Among the 55 cases, there are 30 cases with the Δ values less than 1 d, indicating that the INVM is reasonable for SFT prediction. Nevertheless, it is also observed that the greatest difference between the actual and the predicted SFTs can be up to about 94 d. Such a difference can be caused by both the observational and the model uncertainties. If such uncertainties are not considered, the reliability of the prediction from the traditional INVM is largely unknown.

Then, the maximum likelihood method and BML-based method are used to analyze the SFT of each slope. Fig. 7a and b compares the mean and the SD of t_{cN} calculated through the maximum likelihood method and the BML-based method, respectively. As is shown in these two figures, the mean values of t_{cN} predicted based on the two methods are quite close. However, the SD predicted based on the maximum likelihood method are generally smaller than those calculated using the BML-based method, indicating that the observational uncertainty is underestimated in the maximum likelihood method.

Fig. 7c and d compares the values of μ_m and σ_m estimated by the maximum likelihood method and the BML-based method respectively. As mentioned previously, the uncertainties of μ_m and σ_m in the maximum likelihood method are not considered. Hence, the values of μ_m and σ_m predicted by the maximum likelihood method are represented as points in these two figures. For comparison, in the BML-based method, the uncertainties associated with μ_m and σ_m are considered. The 95% CI of μ_m and σ_m estimated by the BML-based method is shown in Fig. 7c and d. It is meaningful to notice that the values of μ_m and σ_m estimated by the maximum likelihood method are well within the 95% CI estimated based on the Bayesian method. On the other hand, the uncertainties associated with μ_m and σ_m are not negligible. Hence, the maximum likelihood method also underestimates the model uncertainty.

Fig. 7e and f compares the mean and the SD of t_{aN} calculated by the maximum likelihood method and the BML-based method. As can be seen in Fig. 7e, the mean values of t_{aN} calculated using the two methods are largely consistent. The values of SD calculated with the maximum likelihood method are generally smaller than those calculated by the BML-based method in this paper. This is reasonable, since the model uncertainty is underestimated by the maximum likelihood method as well the observational uncertainty.

To check the normality assumptions about t_{cN} and t_{aN} made in the maximum likelihood method, Fig. 8 shows the Jarque-Bera statistics for t_{cN} and t_{aN} for each slope in Table 1. As mentioned above, the critical value for the Jarque-Bera test for normality test is 5.97 at the significance level of 0.05. As can be seen from Fig. 8, the Jarque-Bera statistics for t_{cN} and t_{aN} for the slopes in Table 1 are generally greater than 5.97. Therefore, the null hypothesis of normality should be rejected.

Let t_{aL} and t_{aU} denote the lower and upper bounds of the 95% CI of the predicted SFT, respectively. To facilitate verification, a normalized actual SFT, $t_{a,norm}$, is defined as follows:

$$t_{a,norm} = \frac{t_a - \frac{1}{2}(t_{aL} + t_{aU})}{\frac{1}{2}(t_{aU} - t_{aL})} \tag{28}$$

By definition, if $t_{a,norm}$ is between -1 and 1, it indicates that the 95% CI of the predicted SFT covers the actual SFT. If $t_{a,norm}$ is smaller than -1, the actual SFT is earlier than the lower bound of the 95% CI. If $t_{a,norm}$ is greater than 1, the actual SFT is later than the upper bound of the 95% CI of the SFT. Fig. 9a shows the normalized actual SFT of all the 55 slopes when each of them is considered as the verification example for the case of maximum likelihood method. There are 6 slopes out of 55 slopes with the 95% CI of the predicted SFT not covering the observed SFT. Fig. 9b compares the observed SFT and the 95% CI of the SFT calculated by the suggested method in this study. In this case, the actual SFT of all the 55 slopes falls within their predicted 95% CI, respectively. Overall, while the maximum likelihood method can lead to reasonable prediction of the SFT, it could be unconservative due to the underestimation of the model and observational uncertainties. On the other hand, the BML-based method suggested in this paper is based on less assumptions and can provide more reliable predictions on the SFT.

10. Concluding remarks

Due to model and observational uncertainties, accurate SFT prediction is challenging. In this paper, a BML-based method is proposed to predict the SFT, in which both the model and observational uncertainties are considered. A comprehensive slope database is compiled so that the model uncertainty can be learned through the INVM. Compared with the BML-based method, the

previous maximum likelihood method underestimates both the model and observational uncertainties, and hence also underestimates the uncertainty in the actual failure time. A comprehensive comparison among predictions from different methods shows that the prediction from the BML-based method accords better with the observations of slope failure phenomenon.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was substantially supported by the Shuguang Program from Shanghai Education Development Foundation and Shanghai Municipal Education Commission, China (Grant No. 19SG19), National Natural Science Foundation of China (Grant No. 42072302), and Fundamental Research Funds for the Central Universities, China.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jrmge.2021.09.010.

References

Aguilera, P.A., Fernández, A., Fernández, R., Rumí, R., Salmerón, A., 2011. Bayesian networks in environmental modelling. Environ. Model. Software 26 (12), 1376–1388.
Bartlett, M., Cussens, J., 2017. Integer linear programming for the Bayesian network structure learning problem. Artif. Intell. 244, 258–271.
Borgatti, L., Corsini, A., Barbieri, M., Sartini, G., Truffelli, G., Caputo, G., Puglisi, C., 2006. Large reactivated landslides in weak rock masses: a case study from the Northern Apennines (Italy). Landslides 3 (2), 115–124.
Brooks, S.P., Roberts, G.O., 1998. Assessing convergence of Markov chain Monte Carlo algorithms. Stat. Comput. 8 (4), 319–335.
Brown, I., Hittinger, M., Goodman, R., 1980. Finite element study of the Nevis Bluff (New Zealand) rock slope failure. Rock Mech. 12 (3–4), 231–245.
Cahill, J., Lee, M., 2006. Ground control at Leinster nickel operations. J. S. Afr. Inst. Min. Metall 106 (7), 471–478.
Carey, J.M., Moore, R., Petley, D., Siddle, H.J., 2007. Pre-failure behaviour of slope materials and their significance in the progressive failure of landslides. In: McInnes, R., Jakeways, J., Fairbank, H., Mathie, E. (Eds.), Landslides and Climate Change: Challenges and Solutions: Proceedings of International Conference on Landslides and Climate Change. Taylor & Francis Group, London, UK, pp. 207–215.
Carlà, T., Intrieri, E., Di Traglia, F., Nolesini, T., Gigli, G., Casagli, N., 2017. Guidelines on the use of inverse velocity method as a tool for setting alarm thresholds and forecasting landslides and structure collapses. Landslides 14 (2), 517–534.
Carlà, T., Macciotta, R., Hendry, M., Martin, D., Edwards, T., Evans, T., Farina, P., Intrieri, E., Casagli, N., 2018. Displacement of a landslide retaining wall and application of an enhanced failure forecasting approach. Landslides 15 (3), 489–505.
Castillo, I., Schmidt-Hieber, J., Van der Vaart, A., 2015. Bayesian linear regression with sparse priors. Ann. Stat. 43 (5), 1986–2018.
Ching, J.Y., Phoon, K.K., 2019. Constructing site-specific multivariate probability distribution model using Bayesian machine learning. J. Eng. Mech. 145 (1), 04018126.
Contreras, L.F., Brown, E.T., 2019. Slope reliability and back analysis of failure with geotechnical parameters estimated using Bayesian inference. J. Rock Mech. Geotech. Eng. 11 (3), 628–643.
Cowles, M.K., Carlin, B.P., 1996. Markov chain Monte Carlo convergence diagnostics: a comparative review. J. Am. Stat. Assoc. 91 (434), 883–904.
Crosta, G.B., Agliardi, F., 2002. How to obtain alert velocity thresholds for large rockslides. Phys. Chem. Earth, Parts A/B/C 27 (36), 1557–1565.
Cruden, D.M., Masoumzadeh, S., 1987. Accelerating creep of the slopes of a coal mine. Rock Mech. Rock Eng. 20 (2), 123–135.
Darwiche, A., 2009. Modeling and Reasoning with Bayesian Networks. Cambridge University Press, Cambridge, UK.
Del Castillo, E., 2007. Process Optimization: a Statistical Approach. Springer, New York, NY, USA.
Dick, G.J., Eberhardt, E., Cabrejo-Liévano, A.G., Stead, D., Rose, N.D., 2014. Development of an early-warning time-of-failure analysis methodology for open-pit mine slopes utilizing ground-based slope stability radar monitoring data. Can. Geotech. J. 52 (4), 515–529.
Dixon, N., Hill, R., Kavanagh, J., 2003. Acoustic emission monitoring of slope instability: development of an active waveguide system. Proc. Inst. Civil Eng. Geotech. Eng. 156 (2), 83–95.
D'Elia, B., Picarelli, L., Leroueil, S., Vaunat, J., 1998. Geotechnical characterisation of slope movements in structurally complex clay soils and stiff jointed clays. Riv. Ital. Geotec. 32 (3), 5–47.
Fan, X.M., Xu, Q., Liu, J., Subramanian, S.S., He, C.Y., Zhu, X., Zhou, L., 2019. Successful early warning and emergency response of a disastrous rockslide in Guizhou province, China. Landslides 16 (12), 2445–2457.
Federico, A., Popescu, M., Elia, G., Fidelibus, C., Internò, G., Murianni, A., 2012. Prediction of time to slope failure: a general framework. Environ. Earth Sci. 66 (1), 245–256.
Federico, A., Popescu, M., Murianni, A., 2015. Temporal prediction of landslide occurrence: a possibility or a challenge. Ital. J. Eng. Geol. Environ. 1, 41–60.
Fujisawa, K., Marcato, G., Nomura, Y., Pasuto, A., 2010. Management of a typhoon-induced landslide in Otomura (Japan). Geomorphology 124 (3–4), 150–156.
Fukuzono, T., 1985. A new method for predicting the failure time of a slope. In: Proceedings of the 4th International Conference and Field Workshop on Landslides. Tokyo University Press, Tokyo, Japan, pp. 145–150.
Gelman, A., Stern, H.S., Carlin, J.B., Dunson, D.B., Vehtari, A., Rubin, D.B., 2013. Bayesian Data Analysis, third ed. Chapman and Hall/CRC, Boca Raton, FL, USA.
Geman, S., Geman, D., 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Trans. Pattern Anal. Mach. Intell. 6 (6), 721–741.
Gigli, G., Fanti, R., Canuti, P., Casagli, N., 2011. Integration of advanced monitoring and numerical modeling techniques for the complete risk scenario analysis of rockslides: the case of Mt. Beni (Florence, Italy). Eng. Geol. 120 (1–4), 48–59.
Giovanis, D.G., Papaioannou, I., Straub, D., Papadopoulos, V., 2017. Bayesian updating with subset simulation using artificial neural networks. Comput. Methods Appl. Mech. Eng. 319, 124–145.
Hancox, G.T., 2008. The 1979 Abbotsford Landslide, Dunedin, New Zealand: a retrospective look at its nature and causes. Landslides 5 (2), 177–188.
Harries, N., Noon, D., Rowley, K., 2006. Case studies of slope stability radar used in open cut mines. In: Proceedings of International Symposium on Stability of Rock Slopes in Open Pit Mining and Civil Engineering Situations. The South African Institute of Mining and Metallurgy, pp. 335–342.
Hayashi, S., Park, B.W., Komamura, F., Yamamori, T., 1988. On the forecast of time to failure of slope (II). Landslides 25 (3), 11-16_1.
Intrieri, E., Gigli, G., 2016. Landslide forecasting and factors influencing predictability. Nat. Hazard Earth Sys. 16 (12), 2501–2510.
Intrieri, E., Raspini, F., Fumagalli, A., Lu, P., Del Conte, S., Farina, P., Allievi, J., Ferretti, A., Casagli, N., 2018. The Maoxian landslide as seen from space: detecting precursors of failure with Sentinel-1 data. Landslides 15 (1), 123–133.
Iovine, G., Petrucci, O., Rizzo, V., Tansi, C., 2006. The March 7th 2005 Cavallerizzo (Cerzeto) landslide in Calabria—southern Italy. In: Proceedings of Engineering Geology for Tomorrow's Cities - the 10th IAEG Congress. The Geological Society of London, London, UK, pp. 1–12.
Jarque, C.M., Bera, A.K., 1987. A test for normality of observations and regression residuals. Int. Stat. Rev. 55 (2), 163–172.
Jiang, S.H., Huang, J., Qi, X.H., Zhou, C.B., 2020. Efficient probabilistic back analysis of spatially varying soil parameters for slope reliability assessment. Eng. Geol. 271, 105597.
Kardani, N., Zhou, A.N., Nazem, M., Shen, S.L., 2021. Improved prediction of slope stability using a hybrid stacking ensemble method based on finite element analysis and field data. J. Rock Mech. Geotech. Eng. 13 (1), 188–201.
Kass, R.E., Carlin, B.P., Gelman, A., Neal, R.M., 1998. Markov chain Monte Carlo in practice: a roundtable discussion. Am. Statistician 52 (2), 93–100.
Kayesa, G., 2006. Prediction of slope failure at Letlhakane mine with the geomos slope monitoring system. In: Proceedings of International Symposium on Stability of Rock Slopes in Open Pit Mining and Civil Engineering Situations. The South African Institute of Mining and Metallurgy, pp. 605–622.
Kim, T.H., Cruden, D.M., Martin, C.D., Froese, C.R., Morgan, A.J., 2010. Landslide movements and their characteristics, Town of Peace River, Alberta. In: Proceedings of the 63rd Canadian Geotechnical Conference and 6th Canadian Permafrost Conference, pp. 1622–1629.
Koller, D., Friedman, N., 2009. Probabilistic Graphical Models: Principles and Techniques. MIT press, Cambridge, MA, USA.
Korner-Nievergelt, F., Roth, T., Von Felten, S., Guélat, J., Almasi, B., Korner-Nievergelt, P., 2015. Bayesian Data Analysis in Ecology Using Linear Models with R, BUGS, and Stan. Academic Press, Amsterdam, Netherlands.
Kothari, U.C., Momayez, M., 2018. New approaches to monitoring, analyzing and predicting slope instabilities. J. Geol. Min. Res. 10 (1), 1–14.
Kwan, D., 1971. Observations of the failure of a vertical cut in clay at Welland, Ontario. Can. Geotech. J. 8 (2), 283–298.
Li, X.Z., Kong, J.M., Wang, Z.Y., 2012. Landslide displacement prediction based on combining method with optimal weight. Nat. Hazards 61 (2), 635–646.
Li, M.H., Zhang, L., Ding, C., Li, W.L., Luo, H., Liao, M.S., Xu, Q., 2020. Retrieval of historical surface displacements of the Baige landslide from time-series SAR observations for retrospective analysis of the collapse event. Remote Sens. Environ. 240, 111695.
Loew, S., Gschwind, S., Gischig, V., Keller-Signer, A., Valenti, G., 2017. Monitoring and early warning of the 2012 Preonzo catastrophic rockslope failure. Landslides 14 (1), 141–154.

Ma, Z.J., Mei, G., Piccialli, F., 2021. Machine learning for landslides prevention: a survey. Neural Comput. Appl. 33 (17), 10881–10907.
Manconi, A., Giordan, D., 2015. Landslide early warning based on failure forecast models: the example of the Mt. de La Saxe rockslide, northern Italy. Nat. Hazards Earth Syst. Sci. 15 (7), 1639–1644.
Manconi, A., Giordan, D., 2016. Landslide failure forecast in near-real-time. Geomatics, Nat. Hazards Risk 7 (2), 639–648.
Mazzanti, P., Bozzano, F., Cipriani, I., Prestininzi, A., 2015. New insights into the temporal prediction of landslides by a terrestrial SAR interferometry monitoring case study. Landslides 12 (1), 55–68.
Mei, Q., 2001. Forming conditions and sliding mechanism of switch yard slope at Tianhuangping power station. Chin. J. Rock Mech. Eng. 20 (1), 25–28 (in Chinese).
Moretto, S., Bozzano, F., Esposito, C., Mazzanti, P., Rocca, A., 2017. Assessment of landslide pre-failure monitoring and forecasting using satellite SAR interferometry. Geosciences 7 (2), 36.
Mufundirwa, A., Fujii, Y., Kodama, J., 2010. A new practical method for prediction of geomechanical failure-time. Int. J. Rock Mech. Min. 47 (7), 1079–1090.
Murphy, K.P., 2007. Conjugate Bayesian Analysis of the Gaussian Distribution. Computer Science, the University of British Columbia. https://www.cs.ubc.ca/∼murphyk/Papers/bayesGauss.pdf. (Accessed 25 April 2019).
Neapolitan, R.E., 2004. Learning Bayesian Networks. Pearson, Upper Saddle River, NJ, USA.
Okamoto, T., Larsen, J.O., Matsuura, S., Asano, S., Takeuchi, Y., Grande, L., 2004. Displacement properties of landslide masses at the initiation of failure in quick clay deposits and the effects of meteorological and hydrological factors. Eng. Geol. 72 (3–4), 233–251.
Pearl, J., 1988. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kauffmann Publishers, San Francisco, CA, USA.
Petley, D.N., 2004. The evolution of slope failures: mechanisms of rupture propagation. Nat. Hazard Earth Syst. 4 (1), 147–152.
Petley, D.N., Mantovani, F., Bulmer, M.H., Zannoni, A., 2005. The use of surface monitoring data for the interpretation of landslide movement patterns. Geomorphology 66 (1–4), 133–147.
Rochet, L., 1992. Auscultation—diagnostic—surveillance. B. Eng. Geol. Environ. 45 (1), 41–55 (in French).
Rose, N.D., Hungr, O., 2007. Forecasting potential rock slope failure in open pit mines using the inverse-velocity method. Int. J. Rock Mech. Min. 44 (2), 308–320.
Royán, M.J., Abellán, A., Vilaplana, J.M., 2015. Progressive failure leading to the 3 December 2013 rockfall at Puigcercós scarp (Catalonia, Spain). Landslides 12 (3), 585–595.
Saito, M., 1969. Forecasting time of slope failure by tertiary creep. In: Proceedings of the 7th International Conference on Soil Mechanics and Foundation Engineering, vol. 2, pp. 677–683.
Saito, M., 1979. Evidential study on forecasting occurrence of slope failure. Trans. Dept. Geomech., Armenian Acad. Sci, Yerevan, Armenian.
Segalini, A., Valletta, A., Carri, A., 2018. Landslide time-of-failure forecast and alert threshold assessment: a generalized criterion. Eng. Geol. 245, 72–80.
Shirzadi, A., Bui, D.T., Pham, B.T., Solaimani, K., Chapi, K., Kavian, A., Shahabi, H., Revhaug, I., 2017. Shallow landslide susceptibility assessment using a novel hybrid intelligence approach. Environ. Earth Sci. 76 (2), 60.
Sinharay, S., 2003. Assessing Convergence of the Markov Chain Monte Carlo Algorithms: a Review. ETS, Princeton, NJ, USA.
Skempton, A.W., 1977. Slope stability of cuttings in brown London clay. In: Proceedings of the 9th International Conference on Soil Mechanics and Foundation Engineering, vol. 3, pp. 261–270.
Smith, A.F., 1973. A general Bayesian linear model. J. R. Stat. Soc. Series B Stat. Methodol. 35 (1), 67–75.
Sornette, D., Helmstetter, A., Andersen, J.V., Gluzman, S., Grasso, J.R., Pisarenko, V., 2004. Towards landslide predictions: two case studies. Physica A 338 (3–4), 605–632.
Straub, D., Papaioannou, I., 2015. Bayesian updating with structural reliability methods. J. Eng. Mech. 141 (3), 04014134.
Suwa, H., Mizuno, T., Ishii, T., 2010. Prediction of a landslide and analysis of slide motion with reference to the 2004 Ohto slide in Nara, Japan. Geomorphology 124 (3–4), 157–163.
Venter, J., Kuzmanovic, A., Wessels, S.D.N., 2013. An evaluation of the CUSUM and inverse velocity methods of failure prediction based on two open pit instabilities in the Pilbara. In: Dight, P.M. (Ed.), Proceedings of the 2013 International Symposium on Slope Stability in Open Pit Mining and Civil Engineering. Australian Centre for Geomechanics, Perth, Australia, pp. 1061–1076.
Voight, B., 1988. A method for prediction of volcanic eruptions. Nature 332, 125–130.
Walter, G., Augustin, T., 2010. Bayesian linear regression — different conjugate models and their (in)sensitivity to prior-data conflict. In: Kneib, T., Tutz, G. (Eds.), Statistical Modelling and Regression Structures. Physica-Verlag, Heidelberg, Germany, pp. 59–78.
Wang, D., 2018. Landslide forecast based on numerical simulation and statistical physics. J. Harbin Univ. Commer. (Nat. Sci. Ed.) 34 (3), 333–337 (in Chinese).
Wang, H., 2020. Finding patterns in subsurface using Bayesian machine learning approach. Undergr. Space 5 (1), 84–92.
Xu, B., Li, M.R., 1986. Study on the compound failure type of landslide overturning and slipping in the western area of hanging wall in Jinchuan Open-pit Mine. In: Proceedings of Symposium on Typical Landslides in China. Science Press, Beijing, China, pp. 109–116 (in Chinese).
Xu, Q., Yuan, Y., Zeng, Y., Hack, R., 2011. Some new pre-warning criteria for creep slope failure. Sci. China Technol. Sci. 54 (1), 210–220.
Xu, Q., Peng, D., Zhang, S., Zhu, X., He, C., Qi, X., Zhao, K., Xiu, D., Ju, N., 2020. Successful implementations of a real-time and intelligent early warning system for loess landslides on the Heifangtai terrace, China. Eng. Geol. 278, 105817.
Xue, L., Qin, S., Li, P., Li, G., Oyediran, I.A., Pan, X., 2014. New quantitative displacement criteria for slope deformation process: from the onset of the accelerating creep to brittle rupture and final failure. Eng. Geol. 182, 79–87.
Zhang, J., Wang, Z.P., Zhang, G.D., Xue, Y.D., 2020a. Probabilistic prediction of slope failure time. Eng. Geol. 271, 105586.
Zhang, J., Hu, J., Li, X., Li, J., 2020b. Bayesian network based machine learning for design of pile foundations. Autom. ConStruct. 118, 103295.

Dr. Jie Zhang is a professor at the Department of Geotechnical Engineering of Tongji University, China. He received his PhD degree in Civil Engineering from The Hong Kong University of Science and Technology in 2009. He is currently the secretary of the Engineering Practice of Risk Assessment and Management Committee of the International Society of Soil Mechanics and Geotechnical Engineering (TC304, ISSMGE), and one of the founding managing editors of the journal of *Underground Space*. His research mainly focuses on probabilistic analysis and assessment of geohazards. He is the recipient of several academic awards, including the Outstanding Paper Award from *Computers and Geotechnics* (2017), the Young Researcher Award from Geotechnical Safety Network (GEOSNet) (2017), and the Natural Science Award from the Ministry of Education of China (2018).