*Proc. of the 6th Intl. Symposium on Reliability Engineering and Risk Management (6ISRERM)*
*31 May – 1 June 2018, Singapore*
*Editor(s)* Xudong Qian, Sze Dai Pang, Ghim Ping Raymond Ong, Kok-Kwang Phoon


# BAYESIAN ESTIMATION OF EXCEEDANCE PROBABILITY OF DEBRIS FLOWS

JIAN HE[1], DIAN-QING LI[1], ZI-JUN CAO[1], MI TIAN[2], YI HONG[3],

*[1]State Key Laboratory of Water Resources and Hydropower Engineering Science, Institute of Engineering Risk and Disaster Prevention, Wuhan University, 8 Donghu South Road, Wuhan 430072, P. R. China.*
*E-mail: hejian@whu.edu.cn; dianqing@whu.edu.cn; zijuncao@whu.edu.cn*
*[2] School of Civil Engineering, Architecture and Environment, Hubei University of Technology/28 Nanli Road, Wuhan 430068, P. R. China.*
*E-mail: tianmi0525@126.com*
*[3]College of Civil Engineering and Architecture, Zhejiang University, Hangzhou, P. R. China.*
*E-mail: yi_hong@zju.edu.cn*

Quantitative risk assessment and management of debris flows necessitates estimation of exceedance probability of quantities (e.g., the total discharge $Q_{total}$ and the maximum impact pressure $P_{max}$) crucial to the hazard assessment and planning of mitigation strategies. This is a non-trivial task because various uncertainties exist in observation data of these quantities and the number of observation data is generally limited, particularly for extreme events (e.g., those with large $Q_{total}$ and $P_{max}$), which are of great interest in practice. This paper proposes a Bayesian approach to develop a probabilistic model for estimating exceedance probability of debris flows based on observation data of $Q_{total}$ and $P_{max}$. The probabilistic model obtained from the proposed approach provides not only the exceedance probability but also its associated uncertainty level. For illustration, the proposed approach is applied to developing the probabilistic model of $Q_{total}$ and $P_{max}$ for quantitative risk assessment at Jiangjia Ravine, China. Results show that ignoring the statistical uncertainty in the exceedance probability estimated from a limited number of observation data leads to unconservative results of risk assessment.

*Keywords*: Debris flow, risk assessment, Bayesian approach, exceedance probability, uncertainty.

## 1 Introduction

Risk assessment of debris flows is crucial to the design of mitigation strategies and countermeasures of debris flows. The risk level of debris flows can be quantitatively represented by the exceedance probability (EP), which relies on the probabilistic model of debris flow quantities (e.g., total discharge $Q_{total}$ and maximum impact pressure $P_{max}$), which is usually developed based on the observation data of debris flows. However, there are many unavoidable uncertainties and variabilities affecting site observation data, such as measurement errors, climate uncertainty and inherent variability in geotechnical materials. These uncertainties are propagated into estimates of model parameters, which affect probabilistic model identification and estimation of EP. The uncertainty of EP is often ignored in literature (e.g., Van Steijn, 1996;

Zimmermann et al., 1997; Liu et al., 2008; Hong et al., 2015), which may undermine the reliability of the risk assessment results. It is hence necessary to take the uncertainty into account in the risk assessment of debris flows.

To address this problem, this study proposes a Bayesian approach for development of the probabilistic model of debris flow quantities to estimate the EP and its associated uncertainty. The proposed approach is illustrated with observation data of debris flows at Jiangjia Ravine in China.

## 2 Exceedance Probability of $Q_{total}$ and $P_{max}$

### 3.1 *Field data*

Jiangjia Ravine is located in Yunnan, China (as shown in Fig. 1(a)), and is famous for its frequent debris flows. Dongchuan debris flow observation and research station, Chinese Academy of Sciences, has been recording debris flow events there since 1961. Debris flows in the records are classified into two types: continuous flow and surge flow. Hong et al. (2015) compiles $Q_{total}$ and $P_{max}$ data (see Fig. 1(b)) of 118 continuous flows and 139 surge flows in the period of 1967-2000 (Zhang and Xiong, 1997; Kang et al., 2006; Kang et al., 2007).

### 3.2 *Probabilistic models of $Q_{total}$ and $P_{max}$*

The EP of $Q_{total}$ and $P_{max}$ based on $\mathrm{N_D}$ observational data is defined as:

$$\mathrm{EP}\left(q,p\mid \boldsymbol{D},M\right)=\Pr\left[\left(Q_{total}>q\right)\cup\left(P_{max}>p\right)\mid \boldsymbol{D},M\right] \tag{1}$$

where $q$ and $p$ are threshold values of $Q_{total}$ and $P_{max}$, respectively; $\boldsymbol{D}$ is a 2-by-$\mathrm{N_D}$ vector comprised of observation data of $Q_{total}$ and $P_{max}$; $M$ is the joint probabilistic model of $Q_{total}$ and $P_{max}$. Considering the uncertainty in the model parameter vector $\boldsymbol{\omega}$, Eq. (1) can be rewritten as:

$$\Pr\left[\left(Q_{total}>q\right)\cup\left(P_{max}>p\right)\mid \boldsymbol{D},M\right]=1-\int_{\Omega}\mathrm{F_{Q,P}}\left(q,p;\boldsymbol{\omega}\right)\Pr\left(\boldsymbol{\omega}\mid \boldsymbol{D},M\right)\mathrm{d}\boldsymbol{\omega} \tag{2}$$

where $\mathrm{F_{Q,P}}(\cdot)$ is the joint cumulative distribution function (CDF) of $Q_{total}$ and $P_{max}$; $\Pr(\boldsymbol{\omega}|\boldsymbol{D},M)$ is the probability density function (PDF) of $\boldsymbol{\omega}$ conditional on $\boldsymbol{D}$ and $M$. To bypass the multidimensional integration in Eq. (2), the EP can be estimated using samples of model parameters simulated from $\Pr(\boldsymbol{\omega}|\boldsymbol{D},M)$:

$$\mathrm{EP}(q,p|\boldsymbol{D},M)\approx 1-\frac{1}{\mathrm{N_S}}\sum_{j=1}^{\mathrm{N_S}}\mathrm{F_{Q,P}}\left(q,p;\boldsymbol{\omega}^{(j)}\right) \tag{3}$$

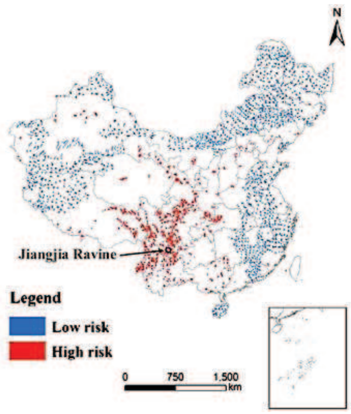
where $\mathrm{N_S}$ is the total number of $\boldsymbol{\omega}$ samples; $\boldsymbol{\omega}^{(j)}$ is the $j$-th sample of $\boldsymbol{\omega}$.

The CDF and PDF of $Q_{total}$ and $P_{max}$ can be constructed using copulas (Nelsen, 1999; Li et al., 2013):

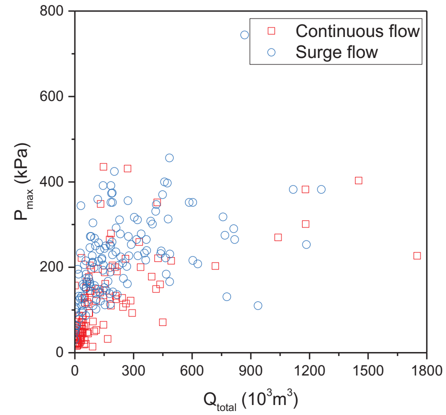
$$\mathrm{F_{Q,P}}\left(q,p;\boldsymbol{\omega}\right)=\mathrm{C}\left[\mathrm{F_Q}\left(q;\mu_\mathrm{Q},\sigma_\mathrm{Q}\right),\mathrm{F_P}\left(p;\mu_\mathrm{P},\sigma_\mathrm{P}\right);\tau\right]=\mathrm{C}(u,v;\tau) \tag{4}$$

$$\mathrm{f_{Q,P}}\left(q,p;\boldsymbol{\omega}\right)=\mathrm{c}(u,v;\tau)\,\mathrm{f_Q}\left(q;\mu_\mathrm{Q},\sigma_\mathrm{Q}\right)\mathrm{f_P}\left(p;\mu_\mathrm{P},\sigma_\mathrm{P}\right) \tag{5}$$

where $\boldsymbol{\omega}=[\mu_\mathrm{Q},\sigma_\mathrm{Q},\mu_\mathrm{P},\sigma_\mathrm{P},\tau]$ is the vector of model parameters; $\mathrm{C}(\cdot)$ is the CDF of copula; $\mu_\mathrm{Q}$ and $\sigma_\mathrm{Q}$ are the mean and standard deviation of $Q_{total}$, respectively; $\mu_\mathrm{P}$ and $\sigma_\mathrm{P}$ are the mean and standard deviation of $P_{max}$, respectively; $u=\mathrm{F_Q}(q;\mu_\mathrm{Q},\sigma_\mathrm{Q})$ and $v=\mathrm{F_P}(p;\mu_\mathrm{P},\sigma_\mathrm{P})$, which are marginal CDFs of




(a)




(b)

Figure 1. (a) Spatial distribution of debris flow hazards in China (Liang et al., 2012; Hong et al., 2015); (b) Observation data of debris flows at Jiangjia Ravine during 1967-2000.

$Q_{total}$ and $P_{max}$, respectively; $\tau$ is the Kendall rank correlation coefficient; c(·) is the PDF of copula; $f_Q(\cdot)$ and $f_P(\cdot)$ are PDFs of $Q_{total}$ and $P_{max}$, respectively. Consider, for example, four commonly-used marginal distributions (i.e., Normal, Lognormal, Weibull and Gamma distributions) and five types of copulas (i.e., Gausssian, Plackett, Frank, Clayton and Gumbel copulas). Thens, 80 (i.e., 4×4×5) candidate joint probability distributions can be obtained by assembling a candidate distribution of $Q_{total}$, a candidate distribution of $P_{max}$, and a candidate copula in every possible combination, denoted as {$M_Q$, $M_P$, copula}. The next section presents the proposed method for probabilistic model identification.

## 3 Bayesian Identification of Probabilistic Models

The probabilistic model $M$ in Eq. (3) can be determined by comparing the occurrence probabilities of candidate models. For $N_M$ (i.e., 80) candidate models, the probability of the $i$-th model $M_i$ given the observation data $\boldsymbol{D}$ is defined as follow (Yuen, 2010):

$$\Pr\left(M_i \mid \boldsymbol{D}\right) = \frac{\Pr(\boldsymbol{D} \mid M_i)\Pr\left(M_i\right)}{\Pr\left(\boldsymbol{D}\right)}, i = 1, 2, \cdots, \mathrm{N_M} \tag{6}$$

where $\Pr(\boldsymbol{D}|M_i)$ is the probability of $\boldsymbol{D}$ given the model $M_i$, and it is referred to as model "evidence" in Bayesian framework; $\Pr(M_i)$ is the prior probability of $M_i$, which can be taken as a constant when there is no prevailing knowledge on models in the absence of data, i.e., $1/\mathrm{N_M}$; $\Pr(\boldsymbol{D}) = \sum_{i=1}^{\mathrm{N_M}} \Pr(\boldsymbol{D} \mid M_i)\Pr(M_i)$ is a normalizing constant for all candidate models. The model evidence is unknown and is needed for calculating the model probability.

Estimating the EP based on Eq. (3) also requires the posterior distribution of $\boldsymbol{\omega}$, which is written as:

$$\Pr\left(\boldsymbol{\omega} \mid \boldsymbol{D}, M_i\right) = \frac{\Pr\left(\boldsymbol{D} \mid \boldsymbol{\omega}, M_i\right)\Pr\left(\boldsymbol{\omega} \mid M_i\right)}{\Pr\left(\boldsymbol{D} \mid M_i\right)} \tag{7}$$

where $\Pr(\boldsymbol{D}|\boldsymbol{\omega},M_i)$ is the likelihood function; $\Pr(\boldsymbol{\omega}|M_i)$ is the prior distribution of model parameters, which can be taken as a joint uniform distribution when there is little prior knowledge on model parameters. Assume debris flow events in the $\mathrm{N_D}$ observations are mutually independent, and then the likelihood function can be calculated by:

$$\Pr\left(\boldsymbol{D} \mid \boldsymbol{\omega}, M_i\right) = \prod_{n=1}^{\mathrm{N_D}} \mathrm{f}_{\mathrm{Q,P}}\left(q_n, p_n; \boldsymbol{\omega}\right) \tag{8}$$

where $\mathrm{f}_{\mathrm{Q,P}}(\cdot)$ is the joint PDF of $Q_{\mathrm{total}}$ and $P_{\mathrm{max}}$ and it is given by Eq. (5); $q_n$ and $p_n$ are the total discharge and the maximum impact pressure of the $n$-th record, respectively.

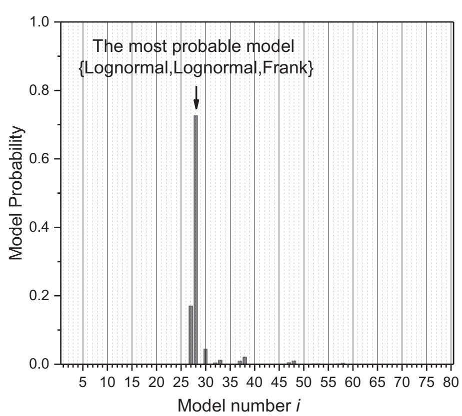
The calculation of model evidence involves multidimensional integration, which can be computationally expensive as the number of model parameters increases. To address this problem, a recently proposed Bayesian updating technique, called Bayesian updating with structural reliability methods (BUS) (Straub and Papaioannou, 2015) is employed. The BUS not only gives the model evidence, but also generates posterior samples of $\boldsymbol{\omega}$ for evaluating EP by Eq. (3). For the sake of conciseness, the details of BUS are not provided herein. Interested readers are referred to Straub and Papaioannou (2015) and Diazdelao et al. (2017).

## 4 Results

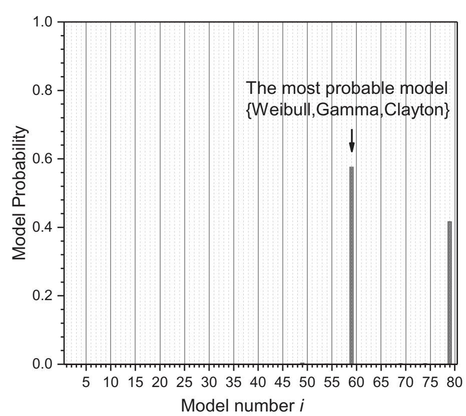
Fig. 2 shows the most probable model $M^*$ and probabilities of candidate models for continuous flow and surge flow. The most probable models for continuous flow and surge flow are {Lognormal, Lognormal, Frank} with the probability of 0.73 and {Weibull, Gamma, Clayton} with the probability of 0.58, respectively. The probabilities of other candidates are much smaller than those of the most probable model.

With the most probable model $M^*$ and posterior samples of $\boldsymbol{\omega}$, the EP of $Q_{total}$ and $P_{max}$ based on observation data can be estimated by Eq. (3), as shown in Fig. 3. Moreover, the coefficient of variation (COV) of the estimated EP can also be obtained using the proposed method, as shown in Fig. 4. For relatively large threshold values, the EP of debris flows exhibits large fluctuation due to scarcity of observation data of extreme events (i.e., debris flows with relatively large $Q_{total}$ and $P_{max}$). Since the EP for large threshold values provides valuable information for the determination of the design value of $Q_{total}$ and $P_{max}$ in the design of mitigation strategies and countermeasures, the large fluctuation of EP for extreme events might cause insufficient design and its associated uncertainty should be considered with caution.

For comparison, the EP reported by Hong et al. (2015), where the probabilistic model is identified by K-S test and the model parameters are determined by the method of moments, is

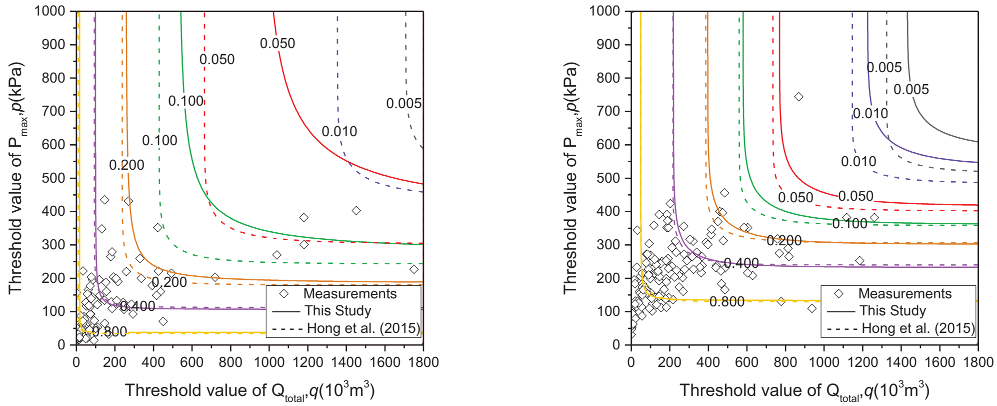



(a)



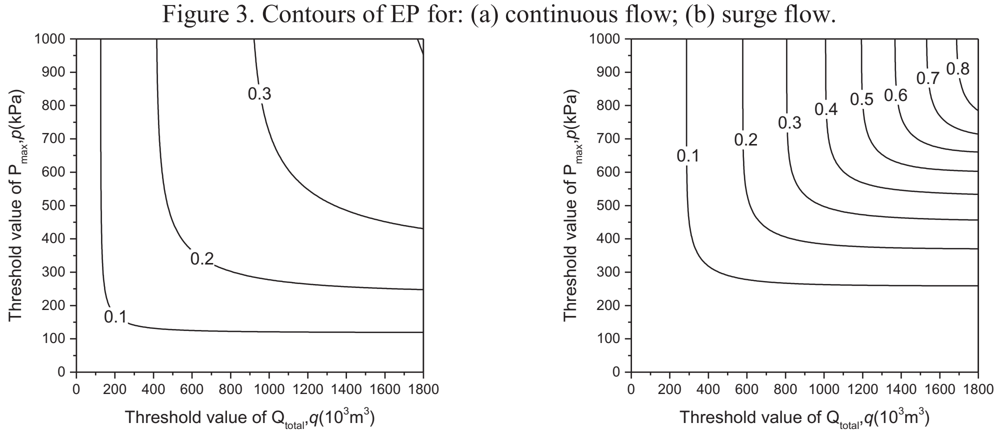

(b)

Figure 2. Model probabilities of candidate models for: (a) continuous flow; (b) surge flow.




(a) (b)

Figure 3. Contours of EP for: (a) continuous flow; (b) surge flow.




(a) (b)

Figure 4. COVs of EP for: (a) continuous flow; (b) surge flow.

also included in Fig. 3 by dashed lines. For both continuous flow and surge flow, the proposed method gives larger estimates of EP than Hong et al. (2015) at a given threshold value. Note that the proposed method considers the uncertainty in model parameters, which is ignored in Hong et al. (2015). This indicates that ignoring the uncertainty in model parameters may lead to underestimation of EP or unconservative designs for debris flow mitigation measures.

The EPs of continuous flow and surge flow exhibit quite different patterns, especially for extreme events. For relatively large threshold values (e.g., $q$=1600×$10^3$m$^3$, $p$=800kPa), the EP of continuous flow (e.g., 0.033 by the proposed method and 0.006 by Hong et al. (2015)) is larger than that of surge flow (i.e., 0.002 by the proposed method and 0.003 by Hong et al. (2015)), indicating the risk level of continuous flow is higher than that of surge flow. This pattern indicates that, to achieve the same safety degree, the design standard of countermeasures of continuous flow is higher than that of surge flow. Therefore, continuous flow plays a dominant role in the design of mitigation strategies.

## 5 Summary and Conclusions

This study proposed a Bayesian approach to develop the probabilistic model for modelling EP of debris flows. Copulas are used to construct a pool of candidate bivariate models of $Q_{total}$ and $P_{max}$. The proposed approach identifies the most probable model among a pool of candidate models, and quantifies the uncertainty of EP. The proposed approach was applied to Jiangjia

Ravine, China. Results show that EP of debris flows might be underestimated without proper consideration of uncertainty of EP, resulting in an unconservative risk assessment or insufficient risk mitigation measures. Comparison of EP between continuous flow and surge flow reveals that the risk level of continuous flow is higher than that of surge flow. Therefore, it is worth paying more attentions to continuous flow in the design of mitigation strategies and countermeasures at Jiangjia Ravine. More importantly, the uncertainty of EP was presented in the form of its COV. Results showed that the uncertainty of EP for large threshold values is inevitably large due to scarcity of observation data of extreme events, and the uncertainty of EP for surge flow is greater than that for continuous flow at the same threshold value. The uncertainty should be treated carefully in the design of mitigation strategies and countermeasures to avoid unconservative designs. Although equations in this study were derived based on $Q_{total}$ and $P_{max}$ data, the probabilistic model can be developed based on other quantities of debris flows (e.g., travel distance) using the proposed approach.

## Acknowledgment

This work was supported by the National Key R&D Program of China (Project No. 2017YFC1501301), the National Natural Science Foundation of China (Project Nos. 51409196, 51579190, and 51779189). The financial support is gratefully acknowledged.

## References


Diazdelao, F. A., Garbuno-Inigo, A., Au, S. K. and Yoshida, I., Bayesian Updating and Model Class Selection with Subset Simulation, *Comput. Meth. Appl. Mech. Eng.*, 317 (Supplement C), 1102-1121, Apr, 2017.

Hong, Y., Wang, J. P., Li, D. Q., Cao, Z. J., Ng, C. W., and Cui, P., Statistical and Probabilistic Analyses of Impact Pressure and Discharge of Debris Flow from 139 Events During 1961 and 2000 at Jiangjia Ravine, China, *Eng. Geol.*, 187, 122-134, Mar, 2015.

Kang, Z. C., Cui, P., Wei, F. Q., and He, S. F., *Data Collection of Dongchuan Debris Flow Observation and Research Station, Chinese Academy of Sciences (1961-1984)*, Science Press, Beijing, China, 2006. (In Chinese)

Kang, Z. C., Cui, P., Wei, F. Q., and He, S. F., *Data Collection of Dongchuan Debris Flow Observation and Research Station, Chinese Academy of Sciences (1995-2000)*, Science Press, Beijing, China, 2007. (In Chinese)

Li, D. Q., Tang, X. S., Phoon, K. K., Chen, Y. F., and Zhou, C. B., Bivariate Simulation Using Copula and Its Application to Probabilistic Pile Settlement Analysis, *Int. J. Numer. Anal. Methods Geomech.*, 37(6), 597-617, Nov, 2013.

Liang, W. J., Zhuang, D. F., Jiang, D., Pan, J. J., and Ren, H. Y., Assessment of Debris Flow Hazards Using a Bayesian Network, *Geomorphology*, 171, 94-100, Oct, 2012.

Liu, J. J., Li, Y., Su, P. C., and Cheng, Z. L., Magnitude–Frequency Relations in Debris Flow, *Environ. Geol.*, 55(6), 1345-1354, Sep, 2008.

Nelsen, R. B., *An Introduction to Copulas*, Springer, New York, 1999.

Straub, D., and Papaioannou, I., Bayesian Updating with Structural Reliability Methods, *J. Eng. Mech.*, 141(3), 04014134, Mar, 2015.

Takahashi, T., *Debris Flow: Mechanics, Prediction and Countermeasures*, 2nd ed., CRC press, Boca Raton, FL, 2014.

Yuen, K. V., Recent Developments of Bayesian Model Class Selection and Applications in Civil Engineering, *Struct. Saf.*, 32(5), 338-346, Sep, 2010.

Van Steijn, H. Debris-Flow Magnitude—Frequency Relationships for Mountainous Regions of Central and Northwest Europe. *Geomorphology*, 15(3-4), 259-273, Apr, 1996.

Zimmermann, M., Mani, P., and Romang, H., Magnitude-frequency aspects of alpine debris flows. *Eclogae Geol. Helv.*, 90(3), 415-420, Sep, 1997.

Zhang, J., and Xiong, G., *Data Collection of Dongchuan Debris Flow Observation and Research Station, Chinese Academy of Sciences (1987-1994)*, Science Press, Beijing, China, 1997. (In Chinese)