PAPER • OPEN ACCESS

Effect of uncertainties on block volume estimation

To cite this article: M T Carriero *et al* 2023 *IOP Conf. Ser.: Earth Environ. Sci.* **1124** 012005

IOP Conf. Series: Earth and Environmental Science 1124 (2023) 012005 doi:10.1088/1755-1315/1124/1/012005

Effect of uncertainties on block volume estimation

M T Carriero[1], A M Ferrero[2], M R Migliazza[1], G Umili[2]

[1]Department of Structural, Geotechnical and Building Engineering, Polytechnic of Turin, Corso Duca degli Abruzzi 24, 10129 Turin, Italy

[2]Department of Earth Sciences, University of Turin, via Valperga Caluso 35, 10125 Turin, Italy

E-mail: gessica.umili@unito.it

Abstract. The combination of the aleatory nature of the rock mass structure and the epistemic errors related to the survey methods make rock mass characterization a challenge despite the remarkable evolution of the survey tools and the research on the subject. In particular, significant uncertainties affect block volume estimation: the need for simplification connected to the engineering approach to rockfall problems, for instance, risks to mask the ripple effect of uncertainties on the reliability of the results. Even considering a simplified shape of the block created by three sets of discontinuities (i.e., a prism), the uncertainties on the geometrical characteristics of the discontinuities (orientation, spacing, and persistence) greatly influence the resulting volume distribution. It is a fact that a single value of the volume cannot be representative of the rock mass: the In Situ Block Size Distribution (IBSD) should be built to describe the variability of block volumes. Many statistical distribution functions can be used for fitting spacing data (i.e., gamma, negative exponential, log-normal, Weibull). The choice of the function must follow a rigorous evaluation of the goodness of fit. This research aims to assess the influence of the uncertainties related to the discontinuities sets, with particular reference to spacing samples, on block volume estimation. Through numerical examples and a case study, this research shows that a reduction of uncertainty can be reached by rigorous statistical processing of the data.

Introduction

Rock block is an important engineering parameter influencing the behavior of rock masses around underground openings and surface excavations, extraction of blocks of commercial sizes of dimension stones and in rock fragmentation processes by blasting or mechanical excavation techniques [1]. The combination of the aleatory nature of the rock mass structure and the epistemic errors related to the survey methods make rock mass characterization a challenge despite the remarkable evolution of the survey tools and the research on the subject.

In particular, block volume estimation suffers from significant uncertainties: if, on the one hand, the engineering approach to rockfall problems needs for easing in the identification of the design parameters, on the other hand, a high level of simplification risks to mask the ripple effect of uncertainties on the reliability of the results. Even considering a simplified shape of the block created by three sets of discontinuities (i.e., a prism), the uncertainties on the geometrical characteristics of the discontinuities (orientation, spacing, and persistence) greatly influence the resulting volume distribution.

It is a fact that a single value of the volume cannot be representative of the rock mass: the In Situ Block Size Distribution (IBSD) should be built to describe the variability of block volumes. It represents the cumulative curve of the potentially detachable in situ blocks, and its construction considers the frequency distributions of spacing values [2].

The first statistical distribution function used for fitting spacing data was the negative exponential [3-5], which is a single parameter distribution.

Later, it was shown that both negative exponential and lognormal distributions can be used, depending on the minimum measurement size (MMS), namely the minimum length of discontinuities which are measured by operators on exposures [6]. By increasing the MMS, a great number of small discontinuities will be ignored in the field measurement, operating a sort of censoring: as an effect, the spacing of discontinuities will follow the lognormal law.

Many authors recommend [7, 8] use a Log-normal distribution for the analysis of discontinuity spacing estimates since it provides greater flexibility by considering the average discontinuity spacing and the variance of the discontinuity spacings. Gamma and Weibull family of distributions can also be used to fit joint spacing data [9].

The log-normal distribution is a continuous single-tailed probability distribution of a random variable whose logarithm is normally distributed, and its probability density function (pdf) can be expressed as:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-(ln(x)-\mu)^2/2\sigma^2} \tag{1}$$

where f(x) is the frequency of a discontinuity spacing x, μ and σ are the mean and standard deviation of the logarithm of the joint spacing, respectively.

The Gamma distribution is a continuous probability distribution described by two parameters: a shape parameter α and an inverse scale parameter β. Its pdf can be expressed as:

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \tag{2}$$

The exponential distribution is a special case of the Gamma distribution, with the standard deviation equal to the mean.

The Weibull distribution is a continuous probability distribution described by two parameters: a shape parameter k and a scale parameter λ. Its pdf can be expressed as:

$$f(x) = \begin{cases} \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k} & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{3}$$

The choice of the best distribution for a set of spacing measurements must follow a rigorous evaluation of the goodness of fit. The one-sample Kolmogorov–Smirnov, for example, quantifies a distance between the empirical cumulative distribution of the sample and the cumulative distribution function (cdf) of the reference distribution. Therefore, it is possible to test different distributions, create a ranking and choose the best one.

In doing so, the epistemic uncertainty on spacing is reduced, particularly the component related to the data processing, but not eliminated, therefore it will be propagated in volume calculation. Consequently, block volume distribution will contain both aleatory and epistemic errors, too.

Many studies [10, 11] proved that the pdf of the block size follows a lognormal distribution. However, studies usually report only the cdf of block volume. A particularly simple method for IBSD construction in the form of cdf was proposed by [12, 13]. Anyway, block volume distribution must be statistically robust, in order to associate with each value a reliable probability of not being exceeded [14].

This research aims to propose an approach to fill the gap between field surveys and data treatment for rock mass characterization. In particular, it assesses the influence of the uncertainties related to the discontinuities sets on block volume estimation through numerical examples and a case study. Evaluation of the goodness of fit of spacing and volume distributions is performed to show that the selection of the proper distribution is a fundamental step for the creation of the IBSD.

1. Fitting of spacing data

As described in [9], the cdf of a discontinuity set spacing, denoted as $F(x) = P\{X \leq x\}$, defines the probability that a given spacing value X is less than x.

A set of 100’000 input spacing values was randomly generated from a Normal distribution with a mean μ_s equal to 1 and standard deviation σ_s equal to 0.001 (0.1% of the mean): the corresponding cdf was obtained. It can be considered as the true distribution, in which the epistemic error is null and only the aleatory component of the uncertainty exists. The same process was done assuming different values of the standard deviation: 0.01, 0.1, 0.25, corresponding to 1%, 10%, and 25% of μ_s, respectively.

The one-sample Kolmogorov–Smirnov test in MATLAB's Statistics Toolbox was used to compare the samples cdf and the hypothesized cdf, among Normal, Gamma, Log-normal and Weibull distributions. The null hypothesis is that the data come from a specific distribution. The test accepts the null hypothesis if the p-value is higher than the significance level, which in this case is assumed to be equal to 0.01 (1%). Small values of p cast doubt on the validity of the null hypothesis.

Table 1 reports, for each generated sample, the logical value of h: if it is equal to 0, the null hypothesis is accepted, if it is equal to 1, it is rejected, that is, the distribution is not suitable for describing the data. The p-value of the distributions returned as a scalar value in the range [0,1] is shown, too. In the case of more than one accepted distribution, the best performance corresponds to the highest p-value.

Normal distribution was found to provide the best fitting of spacing samples, of course, since it is the true one from which data were generated. However, to demonstrate that a test is the only reliable way to identify the best fitting distribution, we repeated this simple exercise reducing the numerosity N of the data in each sample to 100. As expected, the number of data influences the test strongly: it is evident from the results in the last three columns in Table 1 that small samples lead to uncertainty in the identification of the best fitting distributions because the test is not able to reject the null hypothesis. Therefore, all the tested distributions are accepted, and similar p-values are not very helpful in assessing the best fitting distribution.

2. Block volume distribution

To investigate the possible effects of spacing uncertainties, a simple exercise is performed. We consider blocks created by the intersection of three discontinuity sets (K1, K2, K3). The three sets are assumed mutually perpendicular, therefore the effect of orientation in volume calculation is null and volume is simply the product of the three spacing values.

First, we consider the case in which spacing distribution is equal for the three sets (mean and standard deviation constant for K1, K2, K3). The spacing samples described in Section 1 are used to perform a Montecarlo simulation: the sample with σ_s equal to 0.001 is created by inverting the true distribution in correspondence of 1000 randomly generated values from 0 to 1, and it is called S1. The same is done again for creating S2 and S3. So, even if values in S1, S2 and S3 come from the same distribution, in general they are different. Therefore, by multiplying the three i-th values from S1, S2 and S3 we obtain the volume of a generic prismatic block.

The same process was done for each of the four generated spacing samples of Table 1, producing four volume distributions. Then the one-sample Kolmogorov–Smirnov test was used to compare the samples cdf and the hypothesized cdf, among Normal, Gamma, Log-normal and Weibull distributions.

Table 2, which has the same structure as Table 1, refers to volume distributions. Observing volume generated from samples of numerosity N = 1000, it is possible to notice that best fitting distribution changes: in particular, for σ_s equal or smaller than 0.1 Log-normal distribution well fits the data (high p-value), while increasing σ_s Gamma distribution performs better. The strong influence of sample numerosity is evident: again, reducing N to 100 the test is not able to reject the null hypothesis.

Table 3 shows the statistics of the spacing samples and those of the obtained volume samples. In this example, the standard deviation on volume is almost twice the one on spacing.

Table 1. The goodness of fit of spacing distributions evaluated through the Kolmogorov–Smirnov test (significance level α = 0.01) for N = 1000 and N = 100.

		N = 1000			**N = 100**		
σ_s	**Volume Distribution**	**h**	***p*-Value**	**Best Performance**	**h**	***p*-Value**	**Best Performance**
0.001	Normal	0	0.8007	Normal	0	0.4118	Weibull
	Gamma	0	0.7925		0	0.4048	
	Log-normal	0	0.7931		0	0.4094	
	Weibull	1	0.0032		0	0.5949	
0.01	Normal	0	0.4747	Normal	0	0.6190	Normal
	Gamma	0	0.4530		0	0.5956	
	Log-normal	0	0.4504		0	0.5797	
	Weibull	1	0.0011		0	0.2943	
0.1	Normal	0	0.9150	Normal	0	0.7876	Weibull
	Gamma	0	0.5564		0	0.6076	
	Log-normal	0	0.3245		0	0.5444	
	Weibull	1	0.0036		0	0.7905	
0.25	Normal	0	0.9059	Normal	0	0.9996	Normal
	Gamma	0	0.0575		0	0.7106	
	Log-normal	1	0.0014		0	0.4064	
	Weibull	0	0.4639		0	0.9929	

Table 2. The goodness of fit of volume distributions evaluated through the Kolmogorov–Smirnov test (significance level α = 0.01) for N = 1000 and N = 100.

		N = 1000			**N = 100**		
σ_s	**Volume Distribution**	**h**	***p*-Value**	**Best Performance**	**h**	***p*-Value**	**Best Performance**
0.001	Normal	0	0.9147	Log-normal	0	0.9753	Log-normal
	Gamma	0	0.9229		0	0.9688	
	Log-normal	0	0.9237		0	0.9754	
	Weibull	1	0.0000		0	0.4460	
0.01	Normal	0	0.8980	Log-normal	0	0.9775	Log-normal
	Gamma	0	0.9551		0	0.9717	
	Log-normal	0	0.9719		0	0.9781	
	Weibull	1	0.0000		0	0.4430	
0.1	Normal	0	0.0974	Gamma	0	0.8373	Log-normal
	Gamma	0	0.8390		0	0.9634	
	Log-normal	0	0.6840		0	0.9718	
	Weibull	1	0.0001		0	0.5160	
0.25	Normal	1	0.0001	Gamma	0	0.3260	Log-normal
	Gamma	0	0.8505		0	0.8322	
	Log-normal	0	0.0167		0	0.8637	
	Weibull	0	0.0157		0	0.5850	

Figures 1 and 2 show the comparison between cumulative frequency distributions obtained considering the different combinations in Table 3, for numerosity of the spacing sample for each of the three sets forming the block equal to 1000 and 100, respectively.

Table 3. Statistics of spacing samples (mean and standard deviation) and of obtained volume samples.

spacing data		calculated volume			
		N = 1000		N = 100	
μ_s (m)	σ_s (m)	μ_V (m^3)	σ_V (m^3)	μ_V (m^3)	σ_V (m^3)
1	0.001	1.0000	0.0017	1.0001	0.0015
1	0.01	1.0000	0.0172	1.0005	0.0150
1	0.1	1.0005	0.1734	1.0035	0.1502
1	0.25	1.0023	0.4465	1.0006	0.3803

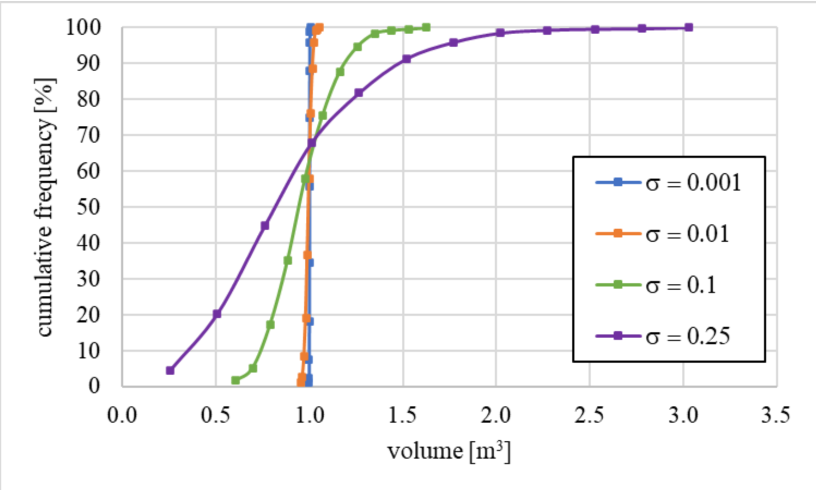

Figure 1. Cumulative frequency distributions of block volumes obtained for N = 1000.

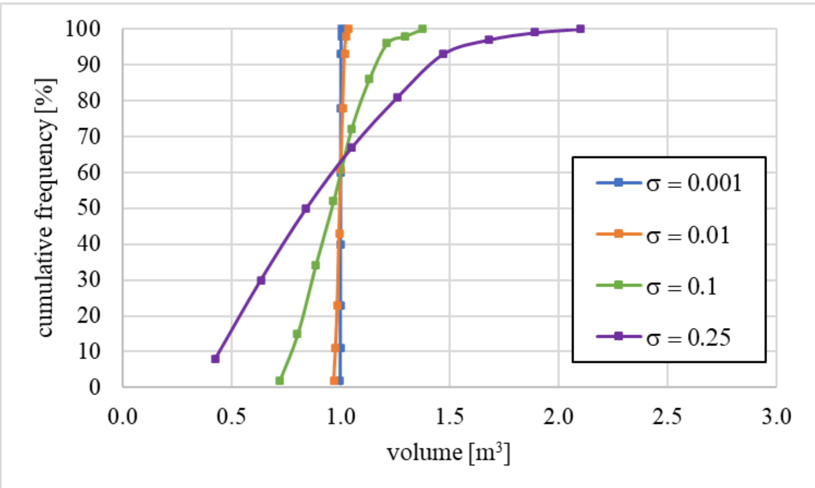

Figure 2. Cumulative frequency distributions of block volumes obtained for N = 100.

3. Case study

In order to assess the influence of uncertainties related to discontinuities sets on block volume estimation, the case study of the Elva valley road (Northern Italy) was used as an example. The so-called

"Strada del Vallone" is located in the orographic left of the Maira Valley (Piedmont, Northern Italy) and directly connects the village of Elva (located at about 1637 m a.s.l.) with the Maira valley bottom road. This road stretches for about 9 km of the carriageway with overhanging rock walls, in which non-contact techniques have been applied to survey certain geometrical features of discontinuities, such as their orientation, spacing, and persistence [15].

Three main sets have been identified through a geostructural survey: a bedding plane (K0) and two conjugated sets (K1 and K2) perpendicular to the bedding plane. These planes remain perpendicular to each other but change their orientation along the road. Moreover, through these indirect techniques, it was possible to measure the spacing between the joints belonging to the three different sets. A sample of 1369 spacing data was collected for K0, while a sample of 934 data was obtained for K1 and K2 together. The mutual perpendicularity of the planes is in agreement with the assumptions made previously and therefore the effect of orientation in volume calculation is null and volume can be simply assumed as the product of the three spacing values.

The one-sample Kolmogorov–Smirnov test accepted only Log-normal distribution for both the spacing samples. Montecarlo simulation based on these distributions was performed to obtain a sample of 1000 volume data, whose best fitting distribution was found to be Log-normal. Figure 3 shows spacing frequency distributions of sets K0, K1, and K2 and their respective Log-normal best-fitting distributions.

For testing the effect of a wrong choice of spacing distributions, another Montecarlo simulation was performed: Normal distributions adapted to spacing samples were considered to obtain a sample of 1000 volume data. Table 4 reports the mean and standard deviation of the spacing and volume distributions obtained with the correct procedure (both spacing and volume best-fitting distributions are Log-normal): in this case study the volume standard deviation is almost twice the spacing one. Table 4 reports also the mean and standard deviation of the spacing and volume distributions obtained with the incorrect procedure: in this case Normal spacing distributions produce a Weibull volume distribution. It is possible to notice that statistics are very different in the two cases.

Figure 4 shows the cumulative frequency distribution of the volume sample obtained with the correct and the wrong procedure. The difference between the two cdfs is evident: in this case, the choice of the design block for a flexible barrier could be strongly over-estimated if the wrong cdf was considered. Considering for example the volume corresponding to the 90% cumulative frequency, namely the 10% probability of being exceeded, the wrong procedure gives a value three times larger than the one estimated through the correct procedure.

Table 4. Parameters of the best-fitting distributions of spacing samples and calculated volume sample (in round brackets equivalent values in logarithmic format); parameters of the wrong distributions of spacing samples and calculated volume sample.

	spacing data			calculated volume		
set	**pdf**	**μ_s (m)**	**σ_s (m)**	**pdf**	**μ_V (m^3)**	**σ_V (m^3)**
K0	Log-normal	0.271 (-1.307)	2.616 (0.962)	Log-normal	0.015 (-4.228)	4.530 (1.511)
K1 and K2	Log-normal	0.238 (-1.435)	2.304 (0.835)	Log-normal		
K0	Normal	0.428	0.464	Weibull	0.1	1.432
K1 and K2	Normal	0.349	0.407	Weibull		

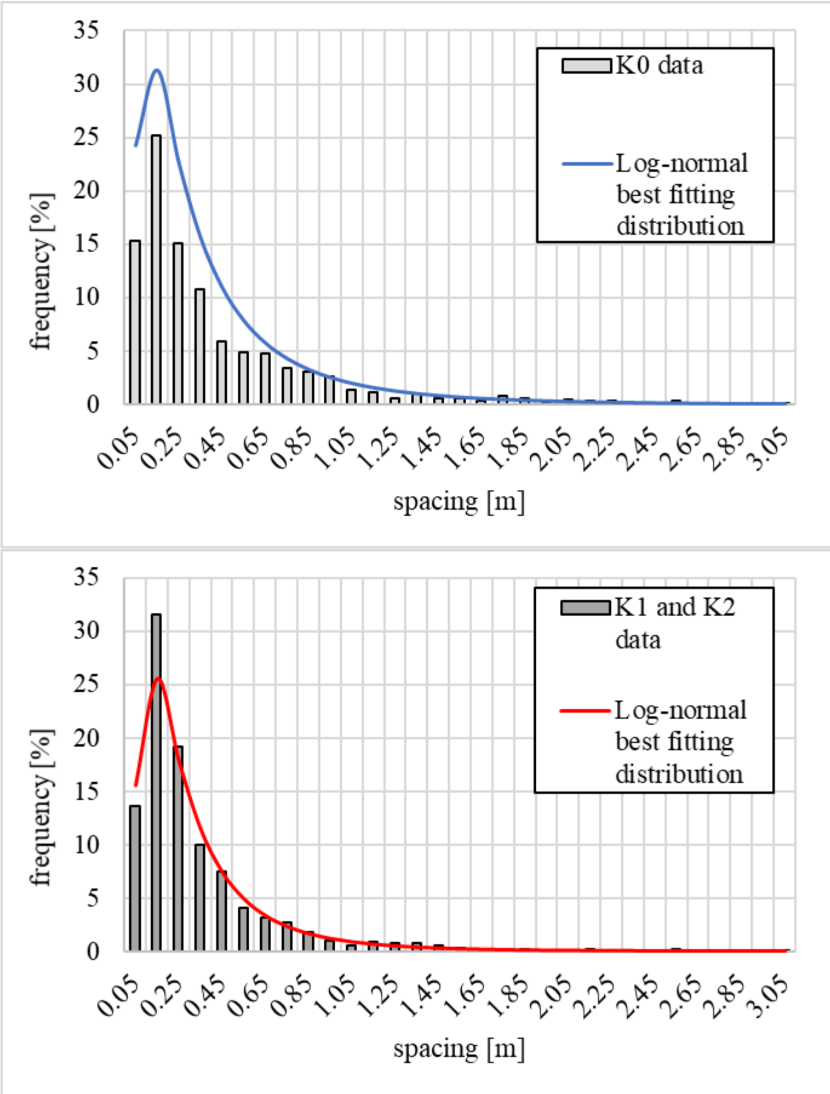

Figure 3. Spacing frequency distributions of set K0 (up), K1 and K2 (down).

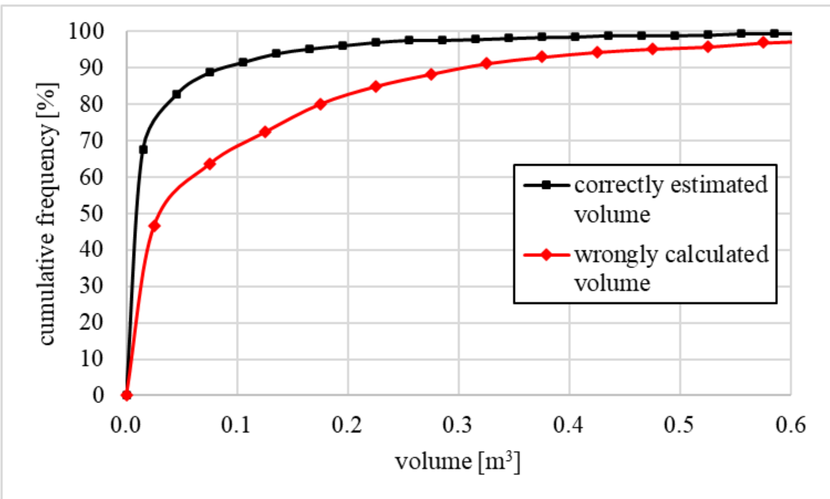

Figure 4. Cumulative frequency distribution of correctly and wrongly estimated block volumes.

Conclusions

This study, through simple numerical examples and a case study, contributed to demonstrate the influence of the uncertainties related to spacing samples on block volume estimation. Moreover, it showed that a reduction of the uncertainty can be reached by a rigorous statistical processing of the data.

Sample high numerosity, rigorous evaluation of the best fitting distribution for spacing samples by means of statistical tests, and a robust Montecarlo simulation are mandatory for reducing epistemic errors in block volume estimation.

References

[1] Stavropoulou M, Saratsis G, Xiroudakis G et al. 2021 Derivation of the Discontinuity Spacings Distribution from Frequency Data Along Scanlines and Boreholes. Rock Mechanics and Rock Engineering. Springer, 54(6), pp. 3095–3113. doi: 10.1007/s00603-021-02442-x

[2] Umili G, Bonetto S M R, Mosca P, Vagnon F and Ferrero AM 2020 In situ block size distribution aimed at the choice of the design block for rockfall barriers design: A case study along gardesana road *Geosci*. **10** 1–21

[3] Priest SD and Hudson J A 1976 Discontinuity spacings in rock. International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts, 13(5), pp. 135–148. doi: 10.1016/0148-9062(76)90818-4

[4] Priest SD, Hudson JA 1981 Estimation of discontinuity spacing and trace length using scanline survey. International Journal of Rock Mechanics and Mining Sciences and Geomechanics Abstract 18, 183-197.

[5] Einstein HH, Beacher BG 1983 Probabilistic and statistical methods in engi- neering Geology specific methods and examples part I: exploration. Rock Mechanics and Rock Engineering 16, 39-72.

[6] Ye J et al. 2012 Correction of the probabilistic density function of discontinuities spacing considering the statistical error based on negative exponential distribution. Journal of Structural Geology. Pergamon, 40, pp. 17–28. doi: 10.1016/j.jsg.2012.04.006.

[7] Odonne F, Lézin C, Massonnat G, Escadeillas G 2007 The relationship between joint aperture, spacing distribution, vertical dimension and carbonate stratification: an example from the Kimmeridgian limestones of Pointe-du-Chay (France). Journal of Structure Geology 29, 746-758.

[8] Annavarapu S, Kemeny J, Dessureault S 2012 Joint spacing distributions from oriented core data. International Journal of Rock Mechanics and Mining Sciences, 52, 40–45. https://doi.org/10.1016/j.ijrmms.2012.02.003

[9] Stavropoulou M 2014 Discontinuity frequency and block volume distribution in rock masses *Int. J. Rock Mech. Min. Sci.* **65** 62–74

[10] Song JJ, Lee CI, Seto M 2001 Stability analysis of rock blocks around a tunnel using a statistical joint modeling technique. Tunnell. Underground Space Technol. 16, 341–351.

[11] Kim BH et al. 2007 Estimation of Block Sizes for Rock Masses with Non-persistent Joints. Rock Mechanics and Rock Engineering. Springer, 40(2), pp. 169–192. doi: 10.1007/s00603-006-0093-8

[12] Lu P, Latham J-P 1999 Developments in the Assessment of In-situ Block Size Distributions of Rock Masses. Rock Mechanics and Rock Engineering. Springer New York, 32(1), pp. 29–49. doi: 10.1007/s006030050042

[13] Latham J-P, Van Meulen J, Dupray S 2006 Prediction of in-situ block size distributions with reference to armourstone for breakwaters. Engineering Geology. Elsevier, 86(1), pp. 18–36. doi: 10.1016/j.enggeo.2006.04.001

[14] Carriero MT, Ferrero AM, Migliazza MR, Umili G 2021 Comparison between methods for calculating the volume of rock blocks, in IOP Conference Series: Earth and Environmental Science. IOP Publishing, p. 012049. doi: 10.1088/1755-1315/833/1/012049

[15] Migliazza M, Carriero MT, Lingua A, Pontoglio E, Scavia C 2021 Rock Mass Characterization by UAV and Close-Range Photogrammetry: A Multiscale Approach Applied along the Vallone dell'Elva Road (Italy). Geosciences, 11, 436. https://doi.org/10.3390/geosciences11110436